# Studies of the Role of the Propeptides of the *Arabidopsis thaliana* 2S Albumin

Kathleen D'Hondt[1], Jose Van Damme, Christel Van Den Bossche, Suporn Leejeerajumnean, Riet De Rycke, Jan Derksen, Joël Vandekerckhove, and Enno Krebbers*

Plant Genetic Systems, Plateaustraat 22, B-9000 Gent, Belgium (K.D'H., J.V.D., E.K.); Laboratorium Fysiologische Chemie (K.D'H., C.V.D.B., S.L., J.V.), and Laboratorium voor Genetica (R.D.R.), Universiteit Gent, Ledeganckstraat 35, B-9000 Gent, Belgium; and Afdeling Botanie, Universiteit Nijmegen, Toernooiveld, NL-6525 ED Nijmegen, The Netherlands (J.D.)

To investigate the possible roles of the *Arabidopsis thaliana* 2S albumin propeptides with respect to sorting, processing, and stability of the protein in plant cells, five gene constructions deleting or modifying the propeptides were made based on one of the genes encoding the *Arabidopsis* 2S albumin. These constructions were introduced into tobacco (*Nicotiana tabacum*) plants. Using subcellular fractionation and immunocytochemistry on ripe seeds, it was demonstrated that none of the propeptides was necessary for the sorting of the protein. Detailed protein-chemical analysis of the mature gene products indicated that, for all of the modified 2S albumin precursors made, the proteins were stably folded and correctly processed. However, the latter is less efficient when the internal fragment between the small and the large subunit is missing or when this internal fragment is changed. In an attempt to establish a rapid assay system for modified 2S albumin precursors, yeast cells were transformed with the same gene constructs. It was demonstrated that the processing machinery in yeast cells differs from that in plants, and, in a perhaps related observation, differences in stability of a particular modified protein were observed.

The PSVs are specialized organelles in plant seeds corresponding to the vacuolar compartment of other eukaryotic cells. They are filled with reserve proteins, which are hydrolyzed upon germination to serve as a nutritional source for the seedlings (Higgins, 1984). Reserve proteins are targeted to the vacuolar compartment via the secretory pathway and carry specific targeting signals (Chrispeels, 1991). In contrast to lysosomal targeting (for review, see von Figura, 1991) in animal cells, so far no glycosylation-dependent targeting has been reported in plant or yeast cells. It is believed that the divergence of the proteins with differing final destinations occurs in the *trans*-Golgi network (Bednarek et al., 1990). Many seed proteins are synthesized as larger precursors, and, apart from the SP, other propeptides are removed during maturation. This processing is thought to occur in a post-Golgi compartment (Chrispeels et al., 1982; Higgins et al., 1983; Stinissen et al., 1985; Gayler et al., 1989).

The *Arabidopsis thaliana* 2S albumin seed storage proteins provide a useful model with which to study both vacuolar targeting and posttranslational processing. Similar proteins have been found in a variety of species (Youle and Huang, 1981), and they are also referred to as the 1.7S or napin-like proteins. In *Arabidopsis* they are encoded by a small gene family consisting of at least five members (Guerche et al., 1990; Van der Klei et al., 1993). These proteins are synthesized as 18-kD precursors from which, apart from the 21-amino acid SP, three other propeptides are removed during maturation of the protein: the ATPF, 16 amino acid residues; the IPF, 10 to 16 amino acids; and the CTPF, 1 to 2 amino acids. The mature protein consists of 3- and 9-kD subunits linked to each other by two disulfide bridges and is not glycosylated (Krebbers et al., 1988).

So far no role other than that of a nutritive source has been suggested for the 2S albumins, and yet they undergo a seemingly wasteful, intensive posttranslational processing. Within the gene family, the propeptides are more highly conserved than the mature parts of the proteins (Krebbers et al., 1988). In this work experiments were done to determine whether the 2S albumin propeptides might have a role in the intracellular targeting or stability of the protein. The question of targeting was addressed by constructing 2S albumin genes from which the regions encoding the propeptides were individually deleted. The same constructs made it possible to determine whether 2S albumins lacking the propeptides could be stably accumulated. Two additional gene constructions that changed the nature of the IPF were made to study its processing in more detail. All five gene constructs were used for the transformation of tobacco. The products were analyzed to determine whether the modified precursor proteins were stably accumulated, targeted to the PSV, and correctly processed. The same constructs were also introduced in *Saccharomyces cerevisiae*. These comparative studies revealed differences in the processing of 2S albumins between plants and yeast.

---

Abbreviations: ATPF, amino-terminal processed fragment; CTPF, carboxyl-terminal processed fragment; GA, Golgi apparatus; IPF, internal processed fragment; PSV, protein storage vacuole; SP, signal peptide.

## MATERIALS AND METHODS

DNA manipulations were done according to established procedures (Sambrook et al., 1989).

### Plasmid Constructions

De Clercq et al. (1990) previously inserted a 1-kb HindIII fragment containing the coding region of the Arabidopsis at2S1 gene into the mutagenesis vector pMAC5–8 (Stanssens et al., 1989) and created an NcoI site at the initiation codon. Because this vector allows further rounds of mutagenesis, the same clone was used as a basis for the modifications made in this study. Site specific oligonucleotide-directed mutagenesis was carried out as described by Stanssens et al. (1989). The mutagenized HindIII fragments were subsequently substituted for the unmodified fragment in pat2S1BG (De Clercq et al., 1990). The latter contained a genomic 3.6-kb BglII fragment including the at2S1-coding sequence as well as 1.85-kb upstream and 1.25-kb downstream flanking sequences. Each 3.6-kb BglII fragment carrying a modified at2S1 gene was placed between the T-DNA borders of the binary plant vector pGSC1703A (Cornelissen and Vandewiele, 1989; De Clercq et al., 1990).

The five modified at2S1 genes and the unmodified gene were also placed behind a yeast inducible promoter ($P_{GAL/CYC}$; Guarente et al., 1982) in the vector pEMBL$_{yex}$4 (Cesareni and Murray, 1987) as NcoI-HindIII fragments, the fusion thus being made at the initiation codon.

### Transformations

Transfer of the binary vector constructs to Agrobacterium tumefaciens strain C58C1Rif$^R$(pMP90) was done as described by Deblaere et al. (1987). Leaf disc transformation of Nicotiana tabacum cv SR1 plants was performed according to the method of Horsch et al. (1985). Saccharomyces cerevisiae strain DL1 (Van Loon et al., 1986) was transformed as described by Klebe et al. (1983).

### Cell Fractionation of Ripe Tobacco Seeds

PSVs were isolated from ripe tobacco seed using the anhydrous potassium-iodide-glycerol step gradient method as described by Sturm et al. (1988). Determination of aryl-$\alpha$-mannosidase (a vacuolar marker; Van Der Wilden and Chrispeels, 1983) activities showed that most of the activity was in the expected fraction, T4 (data not shown).

### RNA Analysis

Total RNA was extracted from seeds 15 to 20 d after flowering according to the method of Beachy et al. (1985). Total RNA (1 $\mu$g) was used for slot blots. Screening was done with a riboprobe specific for the 3' end of the at2S1 mRNA (pFG1, Guerche et al., 1990).

### Immunogold Labeling

Immunogold labeling on ripe seeds of the tobacco transformants was done as described by De Clercq et al. (1990). Statistical analysis of sections from plants expressing the ATPF and CTPF deletions (see "Results") was carried out as follows. In each of three seeds five cells were taken and 18 randomly chosen PSVs were analyzed. For each of these the number of gold particles was counted and correlated to the surface of the matrix. A result was considered positive only when the number of particles/surface area was 300% of the mean background level in the same cell. The results obtained were evaluated with a single-sided sign test (Sokal and Rohlf, 1981) with an error limit of 5% (Table I).

### Protein Purification

Protein extraction from tobacco seeds was done essentially as described by De Clercq et al. (1990) with the modification that 2S albumins were selectively extracted in 100 mM citric acid (pH 3–3.5) after the delipidation step. When further purification was required (for sequencing or for the analysis in Fig. 4) 350 mg of seeds were used. After the gel filtration and reversed phase HPLC steps (De Clercq et al., 1990), 2S albumin-containing fractions were collected and repurified using reversed phase HPLC on a $C_4$ $\mu$-bore column (2.1 × 250 mm, 5-$\mu$m particle size, 30-nm pore size; Vydac Separation Group, Hesperia, CA). The column was equilibrated in 0.1% TFA and eluted with a linear gradient to 70% acetonitrile in 0.1% TFA, applied over a period of 70 min and started 5 min after injection. The flow rate was 1 mL min$^{-1}$. Peptides were detected by $A_{214}$. 2S albumin-containing fractions were collected again, lyophilized, and subjected to oxidation at 0°C for 30 min in 100 $\mu$L of performic acid (prepared by mixing nine parts twice-redistilled formic acid [grade PA, UCB, Brussels, Belgium] with one part 30% hydrogen peroxide [Merck]; the mixture was allowed to stand for 2 h before use) to break the disulfide bridges. The two oxidized subunits were then separated on the $C_4$ $\mu$-bore column using the same procedure as described above. The purified sample of AT2S1 isoform used in Figure 4 was obtained using the procedure of Van der Klei et al. (1993).

### Amino Acid Sequence Determination

The amino acid sequences were determined using an Applied Biosystems, Inc. gas-phase sequenator model 470A equipped with an on-line 120 phenylthiohydantoin amino acid analyzer. The apparatus was run according to the manufacturer's instructions. No significant background sequences were observed, and it was never necessary to read sequences to the limits of acceptable signal:noise ratios to identify the sequence with certainty.

### Growth and Extraction of Yeast Cells

For induction of the $P_{GAL/CYC}$ promoter, transformed yeast cells were grown for 24 h in a rich Gal-containing medium (1% yeast extract [Oxoid]; 2% neutralized bacteriological peptone [Oxoid]; 2% Gal [Sigma]) until late log phase. Cells were harvested by centrifugation and lysed with glass beads as described by Coraggio et al. (1986). The supernatant was boiled before separation on SDS-PAGE (Laemmli, 1970).
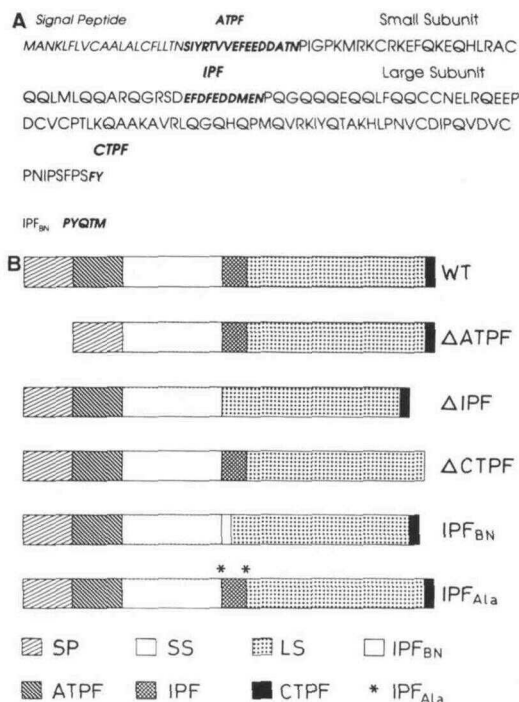
## Immunoblot Analysis

SDS-polyacrylamide gels (15 or 16%) (Laemmli et al., 1970) were run under denaturing but not reducing conditions unless otherwise noted. The proteins were transferred by electroblotting in Tris-Gly buffer onto nitrocellulose membranes (Amersham) (Towbin et al., 1979). 2S albumins were detected by an antiserum raised in rabbits (De Clercq et al., 1990). The membrane was blocked, and the antiserum was diluted (1:1000) in PBS with 0.5% Tween 20 (Batteiger et al., 1982; Wedege and Svenneby, 1986).

## RESULTS

### Experimental Strategy

To investigate the possible role of the *Arabidopsis* 2S albumin propeptides in the stable accumulation and/or intracellular targeting of the protein, three gene constructions (ΔATPF, ΔIPF, and ΔCTPF in Fig. 1) consisting of exact deletions of the information encoding the propeptides were made. It was anticipated that these constructs would also yield information concerning the importance of the propeptides for posttranslational processing. Two further constructs were specifically directed toward examining the effects of propeptide changes on processing. In the first of these (IPF$_{BN}$ in Fig. 1), the information encoding the IPF was replaced by the analogous sequence from the *Bertholletia excelsa* 2S albumin gene (Altenbach et al., 1987). The *Bertholletia* IPF differs from that of *Arabidopsis* in both length and sequence (*Arabidopsis*, EFDFEDDMEN; *Bertholletia*, PYQTM; Fig. 1). The final chimeric gene (IPF$_{Ala}$ in Fig. 1) encodes a 2S albumin precursor in which the amino acids at the IPF-processing sites defined by Krebbers et al. (1988) were changed such that the amino acids on both sites of the junctions are replaced by Ala's instead of Asp and Glu at the N terminus of the IPF and Asn and Pro at the C terminus of this fragment. All five modified *at2S1* genes were introduced into *N. tabacum*. The use of this heterologous system allowed immunological detection of the *Arabidopsis* 2S albumins without interference from endogenous proteins. De Clercq et al. (1990) previously demonstrated that *Arabidopsis* 2S albumins are correctly targeted and processed in this system, and the plants resulting from that study could be used as a positive control.

### The Chimeric Genes Are Expressed

To eliminate the possibility that any observed lack of expression would be due to transcriptional rather than posttranscriptional causes, the expression level of all five modified genes was first checked with RNA slot blots. Unripe seeds from all transformants were harvested at a stage of development when the expression level was expected to be high (De Clercq et al., 1990). No time course was done; therefore, the actual levels detected were minimal measures. With the exception of three plants, all of the transformants expressed the introduced 2S genes at reasonable levels (data not shown).

### The Modified 2S Albumins Are Stably Accumulated

To obtain a first approximation of the stability and processing of the proteins encoded by the modified *at2S1* genes, a small-scale protein extraction was done on ripe seeds harvested from all transformants. After the extracts were fractionated with SDS-PAGE followed by western analysis, 2S albumins could be detected in nearly all plants, representatives of which are shown in Figure 2. Levels of the same order of magnitude to that observed in the positive control plant (a tobacco plant characterized by De Clercq et al. [1990] expressing a wild-type *Arabidopsis* 2S albumin; lane WT, Fig. 2) were observed in extracts from seeds of plants transformed with each of the five constructs. Some variation was observed in individual transformants (data not shown). Because the product could be detected by the antibodies and appeared to have the same molecular mass as the protein encoded by the wild-type gene, no gross defects in processing have occurred. To determine whether processing was precise, amino-terminal sequence analysis of the N termini of the large subunits of the purified proteins was carried out (see below).



**A** Signal Peptide      ATPF      Small Subunit

MANKLFLVCAALALCFLLTN**SIYRTVVEFEEDDATN**PIGPKMRKCRKEFQKEQHLRAC

**IPF**      Large Subunit

QQLMLQQARQGRSD**EFDFEDDMEN**PQGQQQEQQLFQQCCNELRQEEP
DCVCPTLKQAAKAVRLQGGQHQPMQVRKIYQTAKHLPNVCDIPQVDVC

**CTPF**

PNIPSFPS**FY**

IPF$_{BN}$   **PYQTM**

**B**                                              WT

                                                  ΔATPF

                                                  ΔIPF

                                                  ΔCTPF

                                                  IPF$_{BN}$

                                        * *       IPF$_{Ala}$

▨ SP     ☐ SS     ⊞ LS     ☐ IPF$_{BN}$
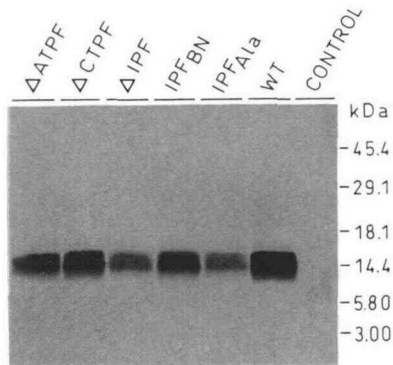▧ ATPF   ▨ IPF     ■ CTPF    * IPF$_{Ala}$

**Figure 1.** Schematic representation of the proteins encoded by the chimeric genes used in this study. A, The amino acid sequence of the 2S albumin precursor encoded by the *at2S1* gene is shown, with the SP and the propeptides indicated in italics. The sequence of the *B. excelsa* IPF is also shown. B, A schematic representation of the structure of the 2S albumin precursor is shown in the top panel (WT), followed by representations of the alterations made. SS, Small subunit; LS, large subunit; IPF, IPF of the *Arabidopsis* 2S albumin; ΔATPF, ATPF deletion construction; ΔIPF, IPF deletion construction; ΔCTPF, CTPF deletion construction; IPF$_{BN}$, the IPF of the *Bertholletia* 2S albumin; IPF$_{Ala}$, modification of the *Arabidopsis* IPF borders (*) as described in the text.

**Figure 2.** Immunoblot analysis of extracts from seeds of individual transgenic tobacco plants. Total extracts were fractionated on a 15% polyacrylamide gel and immunoblotted using a polyclonal antiserum specific for native *Arabidopsis* 2S albumins. The notations above each lane correspond to those in Figure 1; control (negative), an extract from seeds of an untransformed plant. The locations of size markers, indicated in kD, are given on the right.

## Processing Analysis

To determine whether the modifications made in the 2S albumin precursor allowed correct posttranslational processing, the proteins whose precursors had the most drastic changes, the ATPF and IPF deletions, were subjected to N-terminal sequence analysis. *Arabidopsis* 2S albumins were purified from the seeds of tobacco plants carrying the ATPF and IPF deletion constructions. In Figure 3A the elution profile of the protein extracted from seeds of the ΔATPF transformants is shown. The peak indicated by the arrow was N-terminally sequenced after repurification over a $\mu$-bore $C_4$ column (data not shown). A double sequence was obtained, corresponding to the N termini of both the small and the large subunit. A similar protein sample obtained after the repurification step was subjected to performic acid oxidation and separated again over the $\mu$-bore $C_4$ column (Fig. 3B). The peaks corresponded to the small and the large subunit as indicated in the figure. The results obtained upon sequencing of the small subunit matched the sequence of the N terminus of the small subunit of wild-type at2S1-encoded protein (Krebbers et al., 1988). Together, the two sets of sequences indicate that processing was correct at the N terminus of both subunits.

The chromatogram obtained after the first HPLC run of the protein sample extracted from the ΔIPF transformants is shown in Figure 3C. Material from both of the peaks indicated with arrows was immunoreactive with antibodies to *Arabidopsis* 2S albumins in the immuno-dot blot assay. The peaks were pooled and repurified with the $\mu$-bore column. The peaks labeled 1 and 2 in Figure 3D both reacted with the antiserum and were subjected to oxidation. The separation profiles after the oxidized samples were passed over the $\mu$-bore column are given in Figure 3, E and F, respectively. The sequences obtained from the peaks marked SS and LS matched those of the normally processed at2S1, whereas that from the peak indicated with the arrow (Fig. 3F) matched the sequence of the N terminus of the small subunit. On the basis
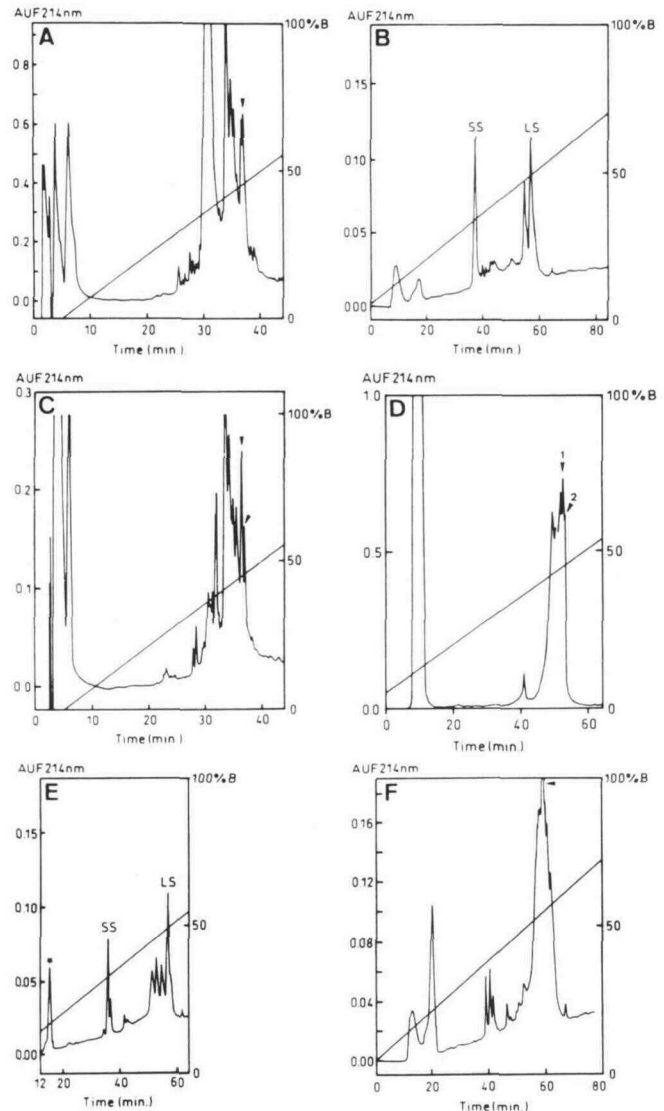


**Figure 3.** Purification of the 2S albumins encoded by the ATPF and IPF deletion constructions. Gel filtration fractions reacting positively with *Arabidopsis* 2S albumin antibodies were combined, loaded on a $C_4$ reversed phase column, and eluted with a linear gradient of acetonitrile as indicated on the chromatograms. Polypeptides are detected by $A_{214}$. A, Purification over a $C_4$ column of the 2S albumin encoded by the ATPF deletion construction. The fractions reacting positively with the antiserum (arrow) were collected and separated a second time over a $C_4$ $\mu$-bore column. The 2S-containing fractions were then oxidized with performic acid and the subunits separated over the $C_4$ $\mu$-bore column (B). The N termini of both subunits were determined. C, Purification over the reversed phase $C_4$ column of the 2S albumin encoded by the IPF deletion construction. The fractions reacting with the antiserum were pooled and separated again over a $C_4$ $\mu$-bore column (D). Both immunoreactive peaks (1 and 2) were individually collected and performic acid oxidized. Although peak 1 could be separated into its subunits (E) from which the N termini were determined, peak 2 could not be separated (F). SS, Small subunit; LS, large subunit; AUF, absorption units full scale. The samples were loaded in 0.1% TFA and eluted with an increasing gradient of 70% acetonitrile in 0.1% TFA (B). The gradients are shown on the chromatograms. The asterisk (*) is not significant.

of the elution time it was concluded that the latter peak represents a form of the precursor whose subunits have not been separated. Taken together, these data indicate that, although at least a portion of the precursor encoded by the ΔIPF construction is correctly processed, a significant portion of the ΔIPF precursor, represented by the peak indicated with the arrow (Fig. 3F) is uncleaved, suggesting that the separation of the two subunits is less efficient in the absence of the IPF.

Due to the low expression levels, it was difficult to subject the proteins encoded by the three other modified genes to sequence analysis, certainly for the C-terminal sequencing that would be required to analyze processing of the protein encoded by the CTPF deletion construct. Instead, these proteins were analyzed using SDS-PAGE. This analysis depends on the fact that, if IPF processing takes place, the two subunits can be separated from each other after the disulfide bridges are broken either by reduction or by performic acid oxidation. The large subunit resulting from such breakage can be detected with an antiserum raised against oxidized mature *Arabidopsis* 2S albumins. 2S albumins encoded by the three additional *at2S1* gene modifications were purified in the same way as the proteins encoded by the ATPF and IPF deletion constructions. The fractions obtained after reversed phase separation of the oxidized samples were analyzed on SDS-PAGE and immunoblotted (Fig. 4). The 2S albumin encoded by the IPF deletion construction was also subjected to this procedure. Samples from tobacco plants expressing the unmodified *Arabidopsis* 2S albumin (lane WT), as well as those from all the other plants, gave a pair of bands representing the large subunit of normally processed 2S albumin (lower arrowhead). Even in plants expressing an unmodified 2S albumin, the large subunit is detected as two bands of different intensities, possibly reflecting differential processing (see "Discussion"). This interpretation is supported by the fact that an individual isoform purified from *Arabidopsis* seeds
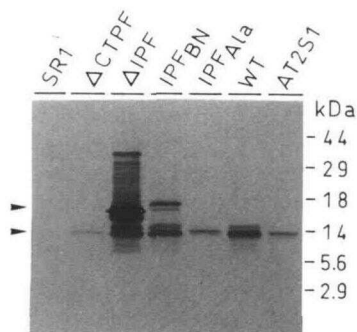


**Figure 4.** Processing analysis through immunoblotting. Chromatography-purified, oxidized 2S albumins from tobacco seeds were fractionated on a 15% polyacrylamide gel and immunoblotted using a polyclonal antiserum raised against oxidized *Arabidopsis* 2S albumins. The notations above each lane correspond to those in Figure 1. Δ, Deletion of the fragment indicated; WT, extract from a transformant expressing an unmodified *Arabidopsis* 2S albumin gene; SR1, an extract from seeds of an untransformed plant; AT2S1, purified isoform 1, out of a complete mixture of *Arabidopsis* 2S albumins. The size markers are indicated in kD on the right. The arrows indicate relevant bands discussed in the text.

(lane AT2S1) does not show the extra band. (Note that all the proteins migrate slower than expected due to the oxidation procedure.) As expected from the sequencing data, in the extract from plants carrying the IPF deletion construction, a larger protein could be detected (upper arrowhead, Fig. 4), whose shift in molecular mass, relative to that of the oxidized large subunit, is that which would be predicted for the unprocessed form. In addition, two bands of a higher molecular mass were detected on the immunoblot in the case of the IPF$_{BN}$ construction, suggesting that in this case as well the processing of the precursor at the IPF was affected; again, the size difference of the higher bands relative to the lower pair is consistent with this interpretation. No protein bands in the region of the upper arrowhead were observed in the other extracts, suggesting that in these cases (IPF$_{Ala}$ and ΔCTPF) the large and small subunits have been cleaved despite the modifications introduced. The origin of the smear and the much higher molecular mass bands (above 20 kD) in the ΔIPF lane are unclear but were reproducible; presumably the modification induces other changes in the isoform affecting the electrophoretic behavior of the oxidized protein.

## Intracellular Localization

To study the impact of the changes that were made in the 2S albumin precursor on their transport to the PSV, two techniques were used. PSVs were isolated using cell fractionation (Sturm et al., 1988). Immunodetection showed that in all cases a large portion of the 2S albumin cofractionated with the PSV, a result consistent with vacuolar targeting of the modified proteins (data not shown). However, significant quantities of 2S albumin were also detected in other fractions. Therefore, the results were confirmed using immunocytochemistry. For each of the five constructs, immunogold EM demonstrated the presence of *Arabidopsis* 2S albumins in the PSV (Fig. 5). The seeds of two classes of transformants (the ATPF and CTPF deletions) showed a large variation in labeling, but statistical analysis of a number of PSVs demonstrated a clear difference in labeling over the matrix region of the PSV between transformants and the negative controls (Table I). No differences were observed between the matrix and the crystalloid regions in the other three classes of transformants; whether this difference among classes of transformants is relevant or related to technical factors was not further studied. In no case was significant labeling observed over the extracellular space. The two approaches together indicate that at least a significant fraction of the 2S albumins encoded by the altered 2S albumin genes are correctly targeted to the PSV in transgenic tobacco seeds and, thus, that none of the propeptides contains a signal that is (alone) necessary for targeting the protein to the PSV. Nevertheless, neither of the two techniques makes it possible to determine the efficiency of the process.

## Expression of 2S Albumins in Yeast

To determine whether yeast could be used as a tool to study the expression, targeting, and/or stability of modified seed proteins, the coding regions of the constructions described above, as well as that of the wild-type gene, were

**Figure 5.** Immunogold electron microscopic localization of *Arabidopsis* 2S albumins encoded by modified *at2S1* genes in the PSVs of transgenic tobacco seeds. Thin sections were made from the seeds of the transformants and incubated with the 2S albumin antiserum, followed by incubation with 15-nm protein A gold particles (De Clercq et al., 1990). Bar, 1 μm; control, section from an untransformed plant. The latter is printed at a smaller magnification to show that background labeling in the regions surrounding the PSV is low.



ATPF deletion

IPF deletion

CTPF deletion

IPF_BN

IPF_Ala

Control
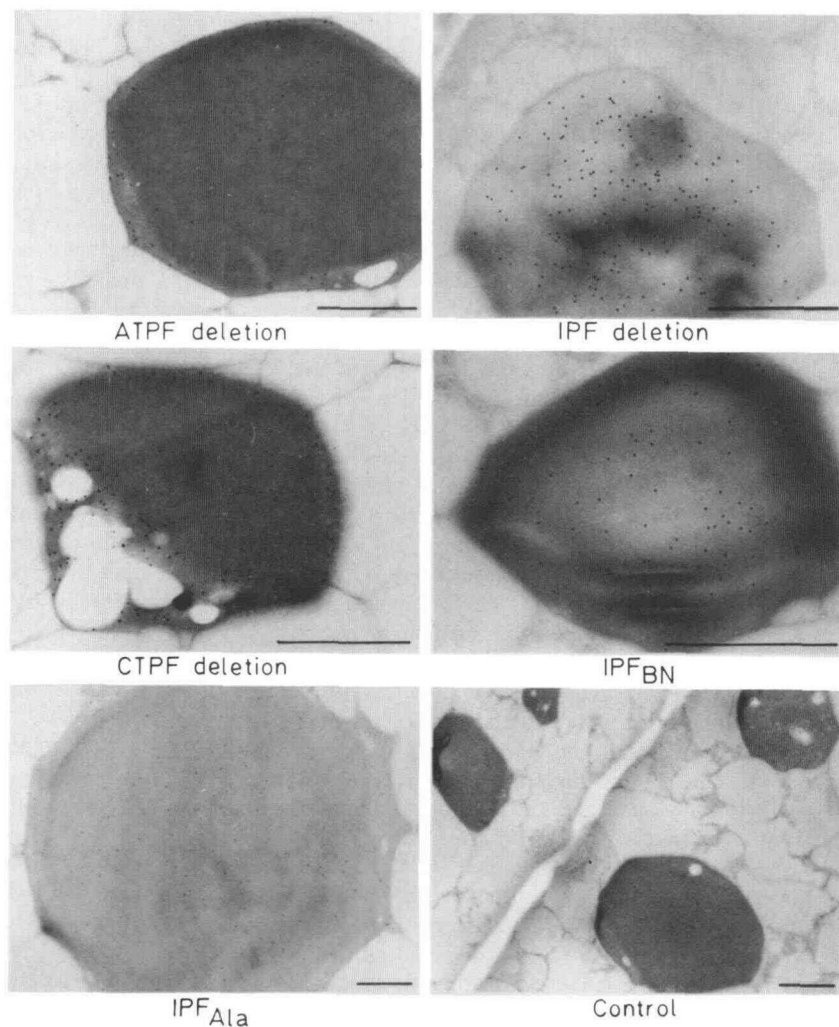
**Table I.** *Statistical analysis of immunogold-labeling experiments*

Results of a single-sided sign test on sections prepared from plants expressing the constructs indicated in the stub. The symbols correspond to the ones used in Figure 1; −control refers to untransformed seeds; P indicates the probability that the distribution of plus- and minus-designated vacuoles could arise from a random distribution. Eighteen vacuoles were statistically analyzed for the transformants indicated. A vacuole was assigned to the positive class if its matrix contained 3-fold the number of gold particles/area as the background level in the matrix region; vacuoles not meeting this standard were assigned to the negative class.

| Transformant | Number of Positive PSVs | Number of Negative PSVs | P |
|---|---|---|---|
| ΔATPF | 14 | 4 | 0.0154 |
| ΔCTPF | 15 | 3 | 0.0038 |
| WT | 13 | 5 | 0.0481 |
| −Control | 3 | 15 | 0.993 |

cloned behind a yeast inducible promoter and introduced into yeast cells (Klebe et al., 1983). Yeast cultures were grown and expression induced as described in "Materials and Methods". Protein extracts were prepared from such cultures and fractionated using SDS-PAGE, and an immunoblot was made (Fig. 6). All of the modified 2S albumins were expressed but, on the basis of migration distance, probably not processed to the mature form (compare lanes 2S and WT, Fig. 6). The expression levels obtained were too low to allow sequence analysis or vacuolar targeting studies. More important, in the case of the ATPF deletion construct, differences in protein stability relative to the results obtained in tobacco were detected. As can be seen when comparing slots ΔATPF* and ΔATPF (Fig. 6), the 2S albumin encoded by this construct was not resistant to boiling after extraction in a Tris pH 7.5 buffer, which is not the case when the analogous protein is purified from tobacco seeds (data not shown). Yeast cells thus appear not to be a representative heterologous expression system for modified 2S albumins.

## DISCUSSION

After the SP has been cleaved off about 20% of the precursor sequence of the *Arabidopsis* 2S albumin precursor
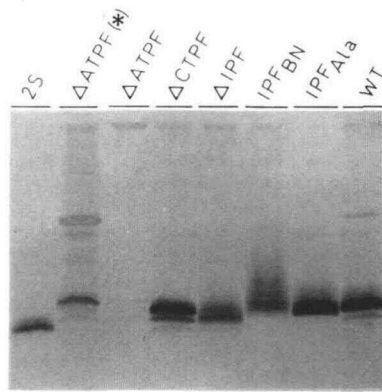
**Figure 6.** Expression of wild-type and modified *at2S1* genes in yeast cells. Protein extracts from yeast cells were separated on 16% SDS-PAGE, blotted, and analyzed with antibodies to native *Arabidopsis* 2S albumins. 2S, Purified 2S albumin (5 ng) from *A. thaliana*. The type of yeast transformants are indicated as in Figure 1; all protein extracts, except the one marked with an asterisk (*), were boiled before sample buffer was added. The 2S albumin in the left-most lane runs at a molecular mass of 15 kD in the gel system used.

is removed during posttranslational processing, a seemingly wasteful process given the lack of a known enzymic activity for these storage proteins. The objective of the experiments described here was to study what the possible role of the 2S albumin propeptides might be, particularly with respect to stability, intracellular targeting, and posttranslational processing. It is perhaps surprising that none of the five modifications made appear to have drastic effects.

One possible effect of deleting a crucial peptide might have been the complete degradation of the translation product, either in the ER or further downstream in the secretory pathway, due to improper folding or disulfide bridge formation. Although small quantitative changes would not have been detected in this study, no gross changes in the accumulation of the introduced 2S albumin could be detected in the transgenic tobacco plants when the ATPF, IPF, or CTPF was deleted. E. Murén and L. Rask (unpublished data) obtained similar results when expressing modified *Brassica napus* 2S albumin genes (napin) from which portions of the ATPF- or IPF-encoding regions have been deleted in tobacco plants; no gross quantitative variations or changes in size were observed. Hoffman et al. (1988) have shown that seemingly small changes in the sequence of a seed protein can result in its complete degradation during intracellular transport.

More significant effects were observed when the processing of the modified proteins was examined, although here again they were not far reaching. Processing of the ATPF deletion protein appeared to be normal, although some heterogeneity was observed; this is probably itself a normal phenomenon (see below). Although it was not possible to carry out C-terminal sequencing, cleavage of the small from the large subunit as measured by gel electrophoresis appeared to be normal when the two-amino acid CTPF was deleted, as might have been expected. However, when the IPF was deleted or replaced by the homologous fragment of the Brazil nut 2S

albumin, a significant portion (peak 2 in Fig. 3D, slots ΔIPF and IPF$_{BN}$ in Fig. 4) of these precursors were not processed. This suggests that, although the precursors can still be processed in the absence of the IPF or in the presence of a completely different IPF, the efficiency of processing is influenced.

In a parallel study it has been shown that there exists an endoprotease in seeds from both *B. napus*, which is closely related to *Arabidopsis*, and tobacco capable of cleaving the *Arabidopsis* IPF in vitro. The better characterized of these activities (from *Brassica*) has been shown to cleave at two sites in the middle of the IPF (K. D'Hondt, D. Bosch, J. Van Damme, M. Goethals, J. Vandekerckhove, E. Krebbers, unpublished data). If the protease activity identified is active in vivo as well, it is not surprising to find that the precursor modified at the junction sites between the IPF and the two subunits (IPF$_{ALA}$) is cleaved; because it was not possible to sequence the N terminus of that product, it is unknown whether the final "trimming" that would be required is correctly carried out. In the case of the IPF deletion and the IPF$_{BN}$ substitution, the protease presumably either finds, with low efficiency, a cryptic site such that the N-terminal sequence of the large subunit is normal and the small subunit is not altered significantly in size or its role is partially overtaken by another protease. In any case several proteases seem to be involved in the maturation of the 2S albumin precursors. If the protease shown to be active in vitro is also active in vivo, when the initial cleavages have been made in the middle of the IPF additional enzymes, either exo- or endoproteases would be necessary for the removal of the rest of the propeptide. Replacement of the *Arabidopsis* IPF by that of Brazil nut does appear to affect cleavage in vivo (slot IPF$_{BN}$, Fig. 4), which is consistent with the results obtained with synthetic peptides corresponding to this chimera, which could not be processed in vitro by the endoprotease under study (K. D'Hondt, D. Bosch, J. Van Damme, M. Goethals, J. Vandekerckhove, E. Krebbers, unpublished data).

Scott et al. (1991) identified a (presumably) Cys proteinase necessary for the processing of soybean glycinin that cuts the precursor specifically behind Asn in an Asn-Gly site. The Asn seems to be conserved in the *Arabidopsis* 2S albumin precursor between the ATPF and the small subunit and between the IPF and the large subunit, and thus it is possible that the same type of proteinase is also involved in removal of the propeptides from the 2S albumin precursor; Hara-Nishimura et al. (1991) reported in vitro cleavage of a pumpkin 2S albumin precursor by a thiol protease. However, the Asn is present in neither the Brazil nut 2S albumin, the chimeric structure represented by construction IPF$_{BN}$, nor the partially processed precursor lacking an IPF. A variety of proteases is presumably responsible for the cleavage of these different forms; elucidation of their specificities and regulation should prove fascinating.

A minor heterogeneity in the processing at the ATPF-small subunit junction was observed. Part of the small subunit encoded by the ATPF deletion construction was processed six amino acids further downstream. It is unlikely that this alternative processing is caused by the modifications that were brought into the precursor. A peak comigrating with such a shorter small subunit was also found during purifica-

tion of the subunits encoded by the IPF deletion construction (data not shown). Similar heterogeneity is observed in *B. napus* (Ericson et al., 1986; Lonnerdahl and Janson, 1972). Therefore, it is likely that the heterogeneity in processing observed at the ATPF-small subunit junction is probably a naturally occurring phenomenon, implying some plasticity in the processing system.

Results obtained by immunocytochemistry show that labeled PSVs were detected in seeds of every type of transformant. This suggests that none of the propeptides are necessary for the vacuolar targeting of the 2S albumins. Targeting information could reside in the propeptides in a redundant fashion, for example in both the ATPF and IPF, because no double-deletion mutants were made. The presence of both necessary and sufficient vacuolar targeting information in one of the propeptides would have made the situation similar to that described by Bednarek and Raikhel (1991), Neuhaus et al. (1991), and Holwerda et al. (1992), all of whom have defined C- or N-terminal propeptides necessary and sufficient for the targeting of different vacuolar proteins. No redundant targeting information appears to reside elsewhere in these proteins. Matsuoka and Nakamura (1991) identified an N-terminal propeptide necessary for the targeting of the sweet potato tuber protein sporamin to the vacuolar compartment. Until now no common feature could be identified among the known vacuolar targeting signals or with any of the propeptides of the 2S albumin precursor. Saalbach et al. (1992) recently reported that a 20-amino acid peptide fragment from the C terminus of the *B. excelsa* 2S albumin (which includes 16 amino acids of the large subunit), a region displaying little sequence homology with the same region of the *Arabidopsis* 2S albumin, can prevent the secretion of invertase fusions from tobacco protoplasts.

The difficulties and delays involved in studying seed proteins in transgenic plants has led to a search for alternative biological systems in which to study protein folding, targeting, and the stability of modified proteins. The possibility of working with yeast initially generated enthusiasm. However, it has become apparent that the vacuolar targeting and processing machinery is not identical with that of plants. In this study it has been demonstrated that the endoproteases involved in 2S albumin maturation are not present or the 2S albumins and proteases involved in maturation of these proteins are localized in different compartments. It further seems that there are differences in stability, and hence most likely in folding, at least in the case of the modified 2S albumin lacking an ATPF. Saalbach et al. (1990) found differential stability of modified globulins in yeast and transgenic plants, and Chrispeels and Tague (1991) reported that there are differences with respect to plants in the recognition of plant vacuolar targeting signals. It thus appears that yeast is for many purposes not a suitable model system for the biosynthesis of plant seed storage proteins.

## ACKNOWLEDGMENTS

## LITERATURE CITED

**Altenbach SB, Pearson KW, Leung FW, Sun SSM** (1987) Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. Plant Mol Biol **8:** 239–250

**Batteiger B, Jones RB, Newhall WJV** (1982) The use of Tween-20 as a blocking agent in the immunological detection of proteins transferred to nitrocellulose membranes. J Immunol Methods **55:** 297–307

**Beachy RN, Chen Z-L, Horsch RB, Rogers SG, Hoffmann NJ, Fraley RT** (1985) Accumulation and assembly of soybean beta-conglycinin in seeds of transformed petunia plants. EMBO **4:** 3047–3053

**Bednarek SY, Raikhel NV** (1991) The barley lectin carboxyl-terminal propeptide is a vacuolar protein sorting determinant in plants. Plant Cell **3:** 1195–1206

**Bednarek SY, Wilkins TA, Dombrowski JE, Raikhel NV** (1990) A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to the vacuoles of tobacco. Plant Cell **2:** 1145–1155

**Cesareni G, Murray J** (1987) Plasmid vectors carrying the replication origin of filamentous single-strand phages. *In* JK Setlow, ed, Genetic Engineering, Vol 9. Plenum Publishing, New York, pp 135–154

**Chrispeels MJ** (1991) Sorting of proteins in the secretory system. Annu Rev Plant Physiol Plant Mol Biol **42:** 21–53

**Chrispeels MJ, Higgins TJV, Spencer D** (1982) Assembly of storage protein oligomers in the endoplasmic reticulum and processing of the polypeptides in the protein bodies of developing pea cotyledons. J Cell Biol **93:** 306–313

**Chrispeels MJ, Tague BW** (1991) Protein sorting in the secretory system of plant cells. Int Rev Cytol **125:** 1–45

**Coraggio I, Compagno C, Martegani E, Ranzi BM, Sala E, Alberghina L, Viotti A** (1986) Transcription and expression of zein sequences in yeast under natural plant or yeast promoters. EMBO J **5:** 459–465

**Cornelissen M, Vandewiele M** (1989) Nuclear transcriptional activity of the tobacco plastid *pbs*A promoter. Nucleic Acids Res **17:** 19–29

**Deblaere R, Reynaerts A, Höfte H, Hernalsteens J-P, Leemans J, Van Montagu M** (1987) Vectors for cloning in plant cells. Methods Enzymol **153:** 277–292

**De Clercq A, Vandewiele M, Van Damme J, Guerche P, Van Montagu M, Vandekerckhove J, Krebbers E** (1990) Stable accumulation of modified 2S albumin seed storage proteins with higher methionine contents in transgenic plants. Plant Physiol **94:** 970–979

**Ericson ML, Rodin J, Lenman M, Glimelius K, Josefsson L-G, Rask L** (1986) Structure of the rapeseed 1.7S storage protein, napin, and its precursor. J Biol Chem **261:** 14576–14581

**Gayler KR, Wachsmann F, Kolivas S, Nott R, Johnson ED** (1989) Isolation and characterization of protein bodies in *Lupinus angustifolius*. Plant Physiol **91:** 1425–1431

**Guarente L, Yocum RR, Gifford P** (1982) A Gal 10-Cyc 1 hybrid yeast promoter identifies the Gal 4 regulatory region as an upstream site. Proc Natl Acad Sci USA **79:** 7410–7414

**Guerche P, Tire C, Grossi De Sa F, De Clercq A, Van Montagu M, Krebbers E** (1990) Differential expression of the *Arabidopsis* 2S albumin genes and the effect of increasing gene family size. Plant Cell **2:** 469–478

**Hara-Nishimura I, Inoue K, Nishimura M** (1991) A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. FEBS Lett **294:** 89–93

**Higgins TJV** (1984) Synthesis and regulation of major proteins in seeds. Annu Rev Plant Physiol **35:** 191–221

**Higgins TJV, Chandler PM, Zurawski G, Button SC, Spencer D** (1983) The biosynthesis and primary structure of pea seed lectin. J Biol Chem **258:** 9544–9549

**Hoffman LM, Donaldson DD, Herman EM** (1988) A modified storage protein is synthesized, processed, and degraded in the seeds of transgenic plants. Plant Mol Biol **11:** 717–729

Holwerda BC, Padgett HS, Rogers JC (1992) Proaleurain vacuolar targeting is mediated by short contiguous peptide interactions. Plant Cell 4: 307–318

Horsch R, Hoffman N, Eicholtz D, Rogers S, Fraley R (1985) A simple and general method for transferring genes into plants. Science 227: 1229–1231

Klebe RJ, Harris JV, Sharp ZD, Douglas MC (1983) A general method for polyethylene glycol induced genetic transformation of bacteria and yeast. Gene 25: 333–341

Krebbers E, Herdies L, De Clercq A, Seurinck J, Leemans J, Van Damme J, Segura M, Gheysen G, Van Montagu M, Vandekerckhove J (1988) Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. Plant Physiol 87: 859–866

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

Lonnerdahl B, Janson J-C (1972) Studies on *Brassica* seed protein. I. The low molecular weight proteins in rapeseed. Isolation and characterization. Biochim Biophys Acta 278: 175–183

Matsuoka K, Nakamura K (1991) Propeptide of a precursor to a plant vacuolar protein required for vacuolar targeting. Proc Natl Acad Sci USA 88: 834–838

Neuhaus J-M, Sticher L, Meins F, Boller T (1991) A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. Proc Natl Acad Sci USA 88: 10362–10366

Saalbach G, Jung R, Kunze G, Manteuffel R, Saalbach I, Muntz K (1990) Expression of modified legume storage protein genes in different systems and studies on intracellular targeting of *Vicia faba* legumin in yeast. *In* Proceedings of Genetic Engineering of Crop Plants, 49th Nottingham Easter School, Sutton Bonington. Butterworth, London, pp 151–158

Saalbach G, Schumann U, Kunze G, Saalbach I, Müntz K (1992) A short C-terminal segment of the Brazil nut 2S albumin targets yeast invertase into the vacuoles of tobacco mesophyll cells. Poster presentation at the Federation of European Societies of Plant Physiology Meeting, Antwerpen, Belgium

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Scott MP, Jung R, Muntz K, Nielsen N (1991) A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. Proc Natl Acad Sci USA 89: 658–662

Sokal RR, Rohlf FJ (1981) Biometry. The Principles and Practice of Statistics in Biological Research, Ed 2. WH Freeman, New York

Stanssens P, Opsomer C, McKeown YM, Kramer W, Zabeau M, Fritz H-J (1989) Efficient oligonucleotide-directed construction of mutations in expression vectors by the gapped duplex DNA method using alternating selectable markers. Nucleic Acids Res 17: 4441–4453

Stinissen HM, Peumans WJ, Chrispeels MJ (1985) Posttranslational processing of proteins in vacuoles and protein bodies is inhibited by monensin. Plant Physiol 77: 495–498

Sturm A, Voelker TA, Herman EM, Chrispeels MJ (1988) Correct glycosylation, Golgi-processing, and targeting to protein bodies of the vacuolar protein phytohemagglutinin in transgenic tobacco. Planta 175: 170–183

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. Proc Natl Acad Sci USA 76: 4350–4354

Van der Klei H, Van Damme J, Casteels P, Krebbers E (1993) A fifth 2S albumin isoform is present in *Arabidopsis thaliana*. Plant Physiol 101: 1415–1416

Van der Wilden W, Chrispeels MJ (1983) Characterization of the isozymes of alfa-mannosidase located in the cell wall, protein bodies, and endoplasmic reticulum of *Phaseolus vulgaris* cotyledons. Plant Physiol 71: 82–87

Van Loon APGM, Pesold-Hurt B, Schatz G (1986) A yeast mutant lacking mitochondrial manganese-superoxide dismutase is hypersensitive to oxygen. Proc Natl Acad Sci USA 83: 3820–3824

von Figura K (1991) Molecular recognition and targeting of lysosomal proteins. Curr Opin Cell Biol 3: 642–646

Wedege E, Svenneby G (1986) Effects of the blocking agents bovine serum albumin and Tween-20 in different buffers on immunoblotting of brain proteins and marker proteins. J Immunol Methods 88: 233–237

Youle R, Huang AHC (1981) Occurrence of low molecular weight and high cysteine containing albumin storage protein in oilseeds of diverse species. Am J Bot 68: 44–48