Journal of Experimental Botany, Vol. 50, No. 331, pp. 165–174, February 1999

Journal of Experimental Botany

Cis-elements of protein transport to the plant vacuoles

Ken Matsuoka[1] and Jean-Marc Neuhaus[2,3]

[1]*Laboratory of Biochemistry, Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa-ku, Nagoya 464–8601, Japan*

[2]*Laboratoire de Biochimie, Université de Neuchâtel, rue Émile-Argand 9, CH-2007 Neuchâtel 7, Switzerland*

Received 3 March 1998; Accepted 16 July 1998

Abstract

Vacuolar proteins are synthesized and translocated into the endoplasmic reticulum and transported to the vacuoles through the secretory pathway. Three different types of vacuolar sorting signals have been identified, carried by N- or C-terminal propeptides or internal sequences. These signals are needed to target proteins to the different types of vacuoles that can coexist in a single plant cell. A conserved motif (NPIXL or NPIR) was identified within N-terminal propeptides, but can also function in a C-terminal propeptide and targets proteins in a receptor-mediated manner to a lytic vacuole. Binding to a family of putative sorting receptors for sequence-specific vacuolar sorting signals has been used as an assay to identify further peptides with other binding motifs. No motif was found in C-terminal sorting sequences, which need an accessible terminus, suggesting that they are recognized from the end by a still unknown receptor. The phosphatidylinositol kinase inhibitor wortmannin differentially affects sorting mediated by these two sorting sequences, suggesting different sorting mechanisms. Less is known about sorting mediated by internal protein sequences, which do not contain the conserved motif identified in N-terminal propeptides and may function by aggregation, leading to transport by coat-less dense vesicles to protein storage vacuoles. Even less is known about the sorting of tonoplast proteins, for which several sorting systems will also be needed.

Key words: Protein trafficking, propeptides, vacuole, tonoplast, sorting receptor.

Introduction

The plant cells possess one or several vacuoles that fulfil various functions. Transport of solutes through the tonoplast is required for the turgor pressure. Vacuoles store proteins, polysaccharides, organic acids, and pigments. Detoxified compounds are deposited in vacuoles, but also precursors of toxic compounds needed for defence (Wink, 1993). The vacuoles belong to the secretory pathway, since they are derived from the endoplasmic reticulum (ER), where both its soluble and tonoplast proteins are synthesized. The soluble proteins of the secretory pathway require a signal sequence for synthesis by membrane-bound ribosomes on the ER. After the protein entered the ER, its signal sequence is removed and the protein is released into the ER lumen. Membrane proteins are similarly inserted into the ER membrane. These proteins are exported from the ER and targeted to their final destinations by vesicular transport. The final location depends on targeting information within the polypeptides (Höfte *et al*., 1991). This review will focus on the structural determinants that are required for the sorting of vacuolar protein precursors and for their transport to the final compartment. The identification of three different types of sorting signals will be linked to the recent visualization of different vacuolar compartments within a single plant cell and to the differential effect of the lipid kinase inhibitor, wortmannin, on the sorting of two different vacuolar proteins. There will be a short discussion of tonoplast protein sorting, for which there is only scant information.

Targeting of soluble proteins to a vacuole requires specific peptide information

For the secretory pathways of animals the concept of a default pathway of secretion has been proposed. According to this 'bulk-flow' model, any soluble protein cotranslationally transported into the ER will be carried passively to the Golgi and from there to the surface of the cell and will eventually be released into the extracellu-

[3] To whom correspondence should be addressed. Fax +41 32 718 22 01. E-mail: jean-marc.neuhaus@bota.unine.ch

Abbreviations: CTPP, C-terminal propeptide; ER, endoplasmic reticulum; NTPP, N-terminal propeptide; VSS, vacuolar sorting sequence.

lar medium unless it carries sorting signals (Rothman and Wieland, 1996). These sorting signals will cause a receptor-mediated accumulation in transport carrier vesicles. This model applies both to the lysosomal targeting by mannose-6-phosphate receptors and to the recycling of escaped ER-resident proteins. In the ER itself, sorting signals may contribute both to the specific concentration of secreted proteins in COPII vesicles and cause the exclusion of ER-resident proteins from such carriers (Balch and Farquhar, 1995; Bannykh and Balch, 1997). Essentially, every transport step involving the formation of specific (coated) vesicles can be expected to cause the concentration of some proteins and the exclusion of others. The concept of a default pathway is more suitable for transport steps involving larger compartments, as in the cisternal maturation model of Golgi transport. This model supported by the direct observation of scale maturation in algae (Becker and Melkonian, 1996), and also the visualization of protein-GFP transport in animal cells (Lippincott-Schwartz and Smith, 1997). There is also a default pathway at single branch points, where a dominant sorting signal determines the fate of a protein. This is the case for the lysosomal sorting of soluble proteins carrying a mannose-6-phosphate in animal cells and for soluble vacuolar proteins carrying a specific propeptide in yeast (Horazdovsky *et al*., 1995). It also applies to plant vacuolar proteins for which such dominant signals have been identified.

Several plant vacuolar proteins have been analysed and in most cases specific determinants were necessary for proper targeting. When these determinants were deleted or mutated, the proteins were secreted into the extracellular space. The determinants could be fused to secreted proteins, transforming them into vacuolar proteins, confirming that secretion is a default pathway for soluble proteins and that vacuolar sorting signals are dominant determinants of protein localization. In all cases analysed the vacuolar sorting signal was found in the polypeptide sequence itself. Glycan side-chains did not contribute to the targeting, in contrast to lysosomal proteins in mammalian cells (Sonnewald *et al*., 1990; Wilkins *et al*., 1990).

Three different types of vacuolar sorting signals (VSS) have been identified

Many of the vacuolar sorting signals were found in N- or C-terminal propeptides, but not all propeptides of vacuolar proteins contain targeting information. In this case, or in vacuolar proteins lacking propeptides, the targeting information must be present within the mature polypeptide chain.

The C-terminal targeting propeptides of chitinases, glucanases or osmotins were identified by comparison of predicted protein sequences with those of related secreted proteins (Neuhaus *et al*., 1991; Sticher *et al*., 1992; Melchers *et al*., 1993). The C-terminal targeting propeptides of cereal lectins and the N-terminal propeptide of sporamin were identified by comparison of the precursor polypeptide with the mature protein (Matsuoka and Nakamura, 1991; Dombrowski *et al*., 1993; Saalbach *et al*., 1996). The N-terminal targeting propeptide of aleurain was identified by comparison with the very different propeptide of a related secreted protease (Holwerda *et al*., 1992). One internal targeting sequence was tentatively located in bean phytohaemagglutinin by fusion to invertase of a series of truncated polypeptides (von Schaewen and Chrispeels, 1993).

The sequence requirement for the vacuolar sorting signals found in the propeptides were determined by deletion and mutation, fusion to reporter proteins and comparison of the sequences. Some of these signals were also characterized by other criteria, such as the affinity to a putative sorting receptor or the sensitivity of targeting to the compound wortmannin in tobacco BY-2 cells.

Based on these results, it is possible to distinguish three types of vacuolar sorting signals, corresponding probably to three different sorting systems. (1) Sequence-specific VSS (ssVSS) can be located in N- or C-terminal propeptides as well as within mature proteins, contain conserved motifs and bind to one or several members of the putative vacuolar sorting receptors (VSR). Their sorting occurs by means of clathrin-coated vesicles and is insensitive to wortmannin. (2) C-terminal VSS (ctVSS) with low sequence specificity must be located at the very end of the polypeptide. Their sorting is sensitive to wortmannin. (3) Protein structure-dependent VSS (psVSS) may be determinants of aggregation and mediate transport to vacuoles by means of dense vesicles.

Sequence-specific vacuolar sorting signals (ssVSS)

To date, sorting signals have been extensively studied in the N-terminal propeptides (NTPPs) of sweet potato sporamin and barley aleurain (Matsuoka and Nakamura, 1991; Holwerda *et al*., 1992). These NTPPs contain a common sequence NPIRL/P, which can also be found in the NTPPs of proteins related to either of these two proteins, the Kunitz-type protease inhibitors of potato tuber (Ishikawa *et al*., 1994) and the family of plant vacuolar proteases, respectively (Fig. 1). Most potato protease inhibitors also have a similar motif NPLDV close to their C-terminus.

Although the deletion analysis indicated that the region of the NPIRL/P motif is critical for its function (Holwerda *et al*., 1992; Nakamura *et al*., 1993), a strict conservation of these amino acids may not be essential to constitute the vacuolar sorting signal. Detailed mutation analysis of this region in sporamin was carried out by replacing each of the five amino acids of the motif by

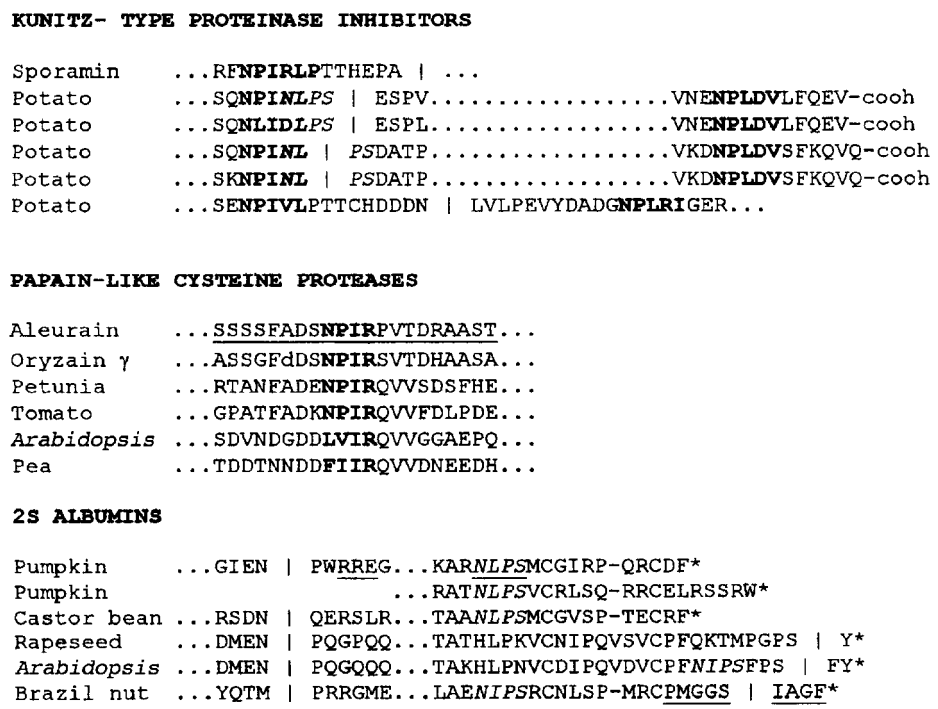

Fig. 1. Comparison of sequence-specific vacuolar sorting sequences (ssVSS) within precursors of vacuolar proteins. Sequences identified by binding *in vitro* to a BP-80 protein are underlined. The conserved NPIXL or NPIR motif is indicated in bold. The conserved NLPS motif is indicated in italics. Processing sites of the propeptides are indicated by |.

at least ten different amino acids and expressing the substitution mutants in tobacco BY-2 cells (K Matsuoka and K Nakamura, unpublished results). Based on this analysis, the following properties of amino acids are required to constitute the VSS: the first amino acid (N in sporamin) should not be small and hydrophobic (A or V). The second position (P) may not be acidic, since E and D cannot substitute for P without decreasing the sorting efficiency. The large and hydrophobic alkyl chain of the third position (I) is essential. This isoleucine can only be changed to leucine without losing the VSS function. Substitution with V or M decreased the sorting efficiency by about 50%, and F or any other tested amino acids caused almost complete secretion of sporamin. The fourth position (R) can be any amino acid, and this amino acid is not conserved in the NTPPs of potato protease inhibitors. The fifth position (L in the case of sporamin and potato protease inhibitors and P in aleurain) should contain a bulky and preferably hydrophobic side-chain, because the exchange with G, A or S abolished vacuolar targeting completely. Thus, many sequences that do not fit the strict NPIRL/P sequence can function as vacuolar sorting signals, although the signal should contain one core L or I. In aleurain, the sorting efficiency depends not only on the NPIRP sequence, but also on whole or part of the upstream and downstream flanking sequences (Fig. 1; Holwerda and Rogers, 1993).

The identification of a putative vacuolar sorting receptor of 80 kDa (BP-80) by affinity chromatography on a column with a coupled aleurain VSS peptide (Kirsch *et al.*, 1994) allowed the testing of potential VSS-containing peptides for their affinity to this sorting system, which were comparable to the results discussed above for sporamin. Binding of an 80 kDa protein from the same pea extract strongly enriched in clathrin-coated vesicles was the criterion (Kirsch *et al.*, 1996). In this system, binding was found with the VSS of aleurain but not with the propeptide of the related secreted endopeptidase. The propeptide of sporamin also bound, while the propeptide of barley lectin did not bind, as expected. Single and double replacements confirmed the importance of Ile within the VSS of sporamin. Its replacement by a Gly abolished binding, while substitution by a Met had no effect. Another 22 residue long peptide, derived from the C-terminus of the precursor of the 2S albumin from Brazil nut (Saalbach *et al.*, 1996), was also found to bind a BP-80. Progressive deletions from the N-terminus caused progressive loss, but then recovery of a BP-80 binding for an 11 amino acids long peptide containing the terminal nine amino acids. These nine terminal amino acids did not contain a NPIRP/L sequence, though they contained a ProMet and an Ile that could be involved in binding to the receptor. On the other hand, deletion of the four last amino acids (IAGF), which corresponded to the propeptide of this 2S albumin, abolished binding. This propeptide has been found to be necessary, though

not sufficient, for vacuolar targeting *in vivo* (Saalbach *et al*., 1996). It is also not excluded that different peptides bind different proteins of similar molecular weight, e.g. different isoforms of BP-80 that are known from different cDNA clones (Paris *et al*., 1997). Preliminary results seem to indicate that one peptide can deplete the extract from the protein that would bind another unrelated peptide. The peptides may also have widely different affinities not detected in this assay system, but at least the aleurain peptide was found to have a much higher affinity for BP-80 than the sporamin peptide (Kirsch *et al*., 1994).

In a similar work with extracts of vesicles from pumpkin, binding of a 72 and an 82 kDa protein (PV72 and PV82) was observed to the aleurain peptide as well as to two peptides derived from the pro-2S albumin from pumpkin itself (Shimada *et al*., 1997). One peptide included an internal propeptide while the other included the C-terminal region of the precursor. Binding to the internal peptide was traced down to a region containing a processing site (MRGIEN|PW**RR**E**G**), and a glycine scan identified two important residues, the R and E indicated in bold. Since this region is exposed on the surface of the albumin precursor, as evidenced by the processing site, it is quite possible that it is a VSS *in vivo*. The C-terminal peptide (KARNLPSMCGIRPQRCDF) contains an NLPS motif conserved in 2S albumins of other species (at the homologous position in castor bean and Brazil nut, but positioned further C-terminally in *Arabidopsis*). Replacement of the four residues by GGGG abolished the binding. Thus NLPS could still be another motif recognized by proteins of the BP-80 family (or VSR, Paris *et al*., 1997). It is also present in the N-terminal propeptides of several potato proteinase inhibitors related to sporamin, linked to the NPI sequence (NPINLPS, Ishikawa *et al*., 1994). It could also be an indication that the receptor can bind I/L-containing peptides in both orientations. Again, as for the Brazil nut albumin, the binding affinity of the pumpkin 2S albumin peptide is unknown and the relevance of this result is unclear as long as the motifs have not been shown to be involved in vacuolar sorting of complete proteins *in vivo*. On the other hand, a peptide representing the 17 C-terminal amino acids of the pro-2S albumin from *Arabidopsis*, including a motif NIPS, was reported not to have bound a BP-80 (Kirsch *et al*., 1996).

While the mutation analysis only detected one essential amino acid, I/L, the comparison of an increasing number of sequence-specific VSSs will probably allow the identification of more consensus positions. While many random sequences could replace a signal sequence or a mitochondrial transit peptide, it was still possible to derive consensus rules to identify natural signal or transit peptides (Baker and Schatz, 1987; Kaiser *et al*., 1987; von Heijne, 1990). There is also obviously no need for sequence-specific VSSs to be included within an N-terminal propeptide. The targeting of sporamin to the vacuole was correct even when the N-terminal propeptide was moved to a C-terminal position (Koide *et al*., 1997). Some proteins without propeptides may thus turn out to harbour their VSS within the mature polypeptide.

Transport to the vacuole of proteins with a ssVSS was found to be relatively resistant to wortmannin in tobacco BY-2 cells, compared with another group of vacuolar proteins described below (Matsuoka *et al*., 1995).

C-terminal vacuolar sorting signals (ctVSS)

Vacuolar proteins with a C-terminal propeptide were secreted when tobacco BY-2 cells were treated with wortmannin. This indicates a biochemical difference between two sorting systems and may be used as a diagnostic test for the type of VSS. However, the differential sensitivity to wortmannin of ssVSS and ctVSS can be observed in BY-2 cells, but not in many other plant cells where wortmannin caused secretion of both protein groups at a lower concentration (1 μM instead of 10–33 μM). The different levels of phosphatidylinositol (PI) 3-kinase activity may explain this difference. In BY-2 cells at log phase, PI 3-kinase activity was very high and the synthesis of PI 3-phosphate was not inhibited completely with relatively high concentrations of wortmannin (Matsuoka *et al*., 1995). By contrast, many plant cells and tissues express very low levels of PI 3-kinase activity and most of these activities are inhibited at lower concentrations of wortmannin (K Matsuoka, unpublished result).

By contrast with the sequence-specific VSSs, no common motif was found in the vacuolar sorting signals identified in the C-terminal propeptides of chitinases or cereal lectins. Several different segments of the barley lectin propeptide were each functional *in vivo*. A minimum length of three (!) residues still caused predominantly vacuolar localization of the lectin in tobacco and four Ala were as effective as LLVD or PIRP, but four Glu, four Lys or four Gly were ineffective (Dombrowski *et al*., 1993). Similarly, in the tobacco chitinase A, six residues were sufficient, many single replacements had little effect and several random sequences could also function as vacuolar sorting signals (Neuhaus *et al*., 1994). Importantly, rendering the sequence more hydrophilic or more hydrophobic did not strongly affect the sorting efficiency.

In both cases, the function of these vacuolar sorting signals could be most efficiently reduced by the addition of one or several Gly residues or of an N-glycosylation site to the end of the propeptide. This means that the ctVSS must be accessible from the end, similar to the ER retention signals -HDEL or -KDEL or the peroxisomal targeting signal -SKL, which can also be blocked by the addition of amino acids.

Sequence comparison of natural sequences may be more informative than analysis of mutants. As mentioned above, mutation analysis is a less sensitive method to identify consensus motifs than comparison of natural sequences. Therefore sequences of C-terminal propeptides were collected from the protein families for which at least one member had been shown to have a C-terminal VSS (Fig. 2). Comparison indicates a preference for Met at the last position, along with Leu, Ile and Val. Amino acids with a long aliphatic side-chain may also carry a terminal charge, as in Glu or Lys. The next two amino acids are most often hydrophilic and/or negatively charged. There is a slight preference for hydrophobic side-chains at the next more distal positions. There may also be species differences as suggested by the divergence of both chitinase and glucanase propeptides in bean, both due to a frame shift. Cereals also may have different preferences.

From these comparisons, the hypothetical receptor is expected to bind the propeptide at its terminal carboxyl group by interactions involving the backbone of the peptide. To encompass the C-terminal amino acid there must be a hydrophobic patch requiring at least a methyl group, but preferably two or three methylene groups, while the end group of the side-chain may not bind at all. Such binding pockets have been described in animal proteins (Saras and Heldin, 1996).

C-terminal extensions or truncations can arise easily by frameshift or point mutations without affecting the proper folding and the activity of proteins. Thanks to the low sequence specificity, secreted and vacuolar proteins can easily be reoriented to the other compartment. Vacuolar proteins using this sorting system are not digestive enzymes (at least not proteases) and are easily tolerated in the extracellular space, where related isoforms are often already present (e.g. chitinases, glucanases and other pathogenesis-related proteins).

The ER retention signal H/KDEL resembles a ctVSS and could also be able to mediate vacuolar sorting. Indeed, when overexpressed, a portion of sporamin with an HDEL extension escaped from the ER and was recovered in the vacuole (Gomord *et al.*, 1997). This may enable the plant cell to recycle escaped ER proteins but it may also be a mechanism controlling the transport of certain proteins to a non-lytic vacuole. An example could be the cysteine endoproteinase from mung bean, which is related to aleurain but harbours a terminal -KDEL and degrades seed globulins after germination (Okamoto *et al.*, 1994). In the same protease families, there are also proteases with an additional Cys-rich C-terminal domain which ends with what could well be a ctVSS, e.g. -LIDNM in oryzain β.

Some C-terminal propeptides may combine both sorting signals described so far. In the case of the Brazil nut 2S albumin, the C-terminal 9 amino acids (including the four amino acids long propeptide) constitute the vacuolar targeting signal (Saalbach *et al.*, 1991) and bind to a BP-80 (Kirsch *et al.*, 1996). It is possible that the propeptide of the tobacco osmotin AP-24 uses both mechanisms, as it contains a sequence FPLEM that fits to the minimum sequence requirement defined with sporamin (see above). As an artificial C-terminal propeptide, the N-terminal

CEREAL LECTINS

Rice	CYKGGDGMAAILAN*NQS*VSFEGIIESVA**E**LV	
Barley	CDG | VFA**E**AIAA*NST*LVA**E**	
Wheat	CDG | VFA**E**AITA*NST*LLQ**E**	
Wheat	CDA | VFAGAITA*NST*LLA**E**	
Wheat	CDG | VFA**E**AIAT*NST*LLA**E**	

CHITINASES

Maize	FNN NSASLAGTAAHA**E**A	
Maize	STA HPCWNRCSA**E**A	
Pea	FGS SLPLSSILLDTVAAA	
Alfalfa	FGS SLSLSSLFLNSI**D**T	
Bean	FGN SLLLSDLVTSQ	
Vigna	FGN SLLNLHPIV	
Avocado	FGV STNPLAASS	
Arabidopsis	FVN GLL**E**AAI	
Cotton	FGN GVSV**D**SM	
Ulmus	FGN GLLL**D**TM	
Tomato	FGN GLLV**D**IM	
Potato	FGN ALLV**D**TL	
Potato	FGN GLLV**D**TV	
Tobacco	FGN | GLLV**D**TM	
Grapevine	FGS GLLL**D**TI	
Poplar	FGY GLSGLK**D**TM	
Poplar	FEDN GLLK*M*VGTM	

GLUCANASES

Bean	FG AQRMQRLLLMSSMQHIPLRVTCKLEPSSQSLL	
Pea	FG GERRDGEIVEGDF*NGT*VSLKS**D**M	
Alfalfa	FG G-ERMGIVNGDF*NAT*I-SLKS**D**M	
N. plumbaginifolia	FG VSG-----SVET*NAT*A-SLIS**E**I	
N. glutinosa	FG VSGRV-DSSVET*NAT*A-SLIS**E**M	
Tobacco	FG | VSGGVWDSSVET*NAT*A-SLVS**E**M	
Tomato	FG VSERVWDI---T*NST*ASSLTS**E**I	
Potato	FG VSERVWDISAET*NST*TSSLIS**E**M	
Hevea	FG A-EKNWDISTEH*NAT*ILFLKS**D**M	

OSMOTINS

Thaumatococcus	CP TALEL**EDE**	
Tobacco	CPN GQAHP---NFPLEMP--GSD**E**VAK	
Tobacco	CPY GSAHNETTNFPLEMPTSSTH**E**VAK	
S. commersonnii	CPY GSTHNETTNFPLEMP-TSTL**E**VA	
Tomato	CPN GVADP---NFPLEMP-ASTD**E**VAK	
Arabidopsis	CPR SRLGATGSHQLPIKMVT**EE**N	

PR-4

Tobacco	CGDN MNVLVSPV**D**K**E**	
Potato	CGDN VNVPLLSV**V**DK**E**	
Tomato	CGDN VNVPLLSV**V**DR**E**	
Hevea	CGDS FNPLFSVMKSSVIN**E**	
Soybean	CGNE LDLTKPLLSIL**D**AP	
Arabidopsis	CGNE LIGQPDSRNMLVSAI**D**RV	

PR-1

Tobacco	RPFG DLEEQPFDSKLELPT**D**V	
Tomato	NVYG DLEEQKPDFDSKL**E**LPN	

Fig. 2. Sequence comparison of C-terminal propeptides of several protein families. Glycosylation sites are indicated in italics, negative charges near the C-terminus are indicated in bold. Processing sites of some propeptides identified by terminal sequencing of the mature proteins are indicated by |, for other propeptides they were placed by comparison with secreted members of the same protein family and are suggested by the spacing.

propeptide of prosporamin functioned as a vacuolar sorting signal, but targeting was still wortmannin-insensitive (Koide *et al.*, 1997). Replacement of the conserved Ile by Gly in this displaced propeptide did not cause secretion as it did in the original position, but rendered the vacuolar sorting wortmannin-sensitive as for a *bona fide* ctVSS.

The failure so far to identify a sorting receptor for ctVSSs could be a consequence of their low sequence specificity. It could, however, also be due to a receptor-free sorting mediated by aggregation, as discussed below for the third type of VSS.

Physical structure vacuolar sorting signals (psVSS)

Some storage proteins are thought to have a sorting signal within the mature polypeptide chain. This may be due to lack of information on processing, but several seed proteins are sorted by a different mechanism. Seed proteins of the vicilin and legumin families accumulate in dense vesicles that bud off from the *trans*-Golgi without the participation of clathrin coats and accumulate in protein storage vacuoles (Robinson *et al.*, 1998). Aggregation has been proposed as a non-receptor-mediated sorting mechanism (Vitale and Chrispeels, 1992) and is known to occur in a pH-dependent manner in animal cells. If aggregation occurs rapidly after synthesis, protein bodies can form from the ER as is the case for cereal prolamins (Levanony *et al.*, 1992). In legumes, oligomerization is an important determinant of transport competence (Vitale *et al.*, 1995) and affects the accessibility of hydrophobic patches on the surface of the protein. Pea prolegumin is more hydrophobic in the ER and Golgi than is legumin in the protein bodies (Hinz *et al.*, 1997). Processing in the Golgi may further uncover aggregation determinants. With such a mechanism, it is difficult to locate the sorting signals precisely.

One such internal signal was tentatively located in a surface loop of bean phytohaemagglutinin (PHA) by the use of fusions of truncated PHA to invertase (von Schaewen and Chrispeels, 1993). This internal loop could contain a ssVSS (GFLGL could fit). However, truncation may also have exposed the surface sequences that cause aggregation without necessarily playing any role in the sorting of the intact protein precursor. The role of the C-terminal propeptide of PHA has also not yet been addressed (Young *et al.*, 1995), as it could function either as a ctVSS or by modulating the exposition of an internal aggregation determinant. It is not excluded that BP-80 proteins could also play a role in transporting malfolded proteins to the lytic vacuole if they escaped the quality control in the ER, as does the CPY receptor Vps10p for malfolded proteins in yeast (Hong *et al.*, 1996).

Two or three sorting mechanisms, two or three target vacuoles?

The sorting system for sequence-specific VSSs has been compared with and is clearly distinct from the other two sorting systems, but the latter two have not been clearly compared yet. It is known that ssVSS-carrying proteins are exported from the Golgi apparatus to lytic vacuoles by clathrin-coated vesicles while psVSS-containing proteins are exported to protein bodies by apparently coat-less dense vesicles (Robinson *et al.*, 1998). It is also known that proteins with a ctVSS are sorted to a vacuole distinct from the lytic vacuole (Paris *et al.*, 1996) by a wortmannin-sensitive mechanism (Matsuoka *et al.*, 1995). This could even be seen in live cells by the use of a Green Fluorescent Protein (GFP) fused to the C-terminal VSS of tobacco chitinase. This GFP accumulated in the large central vacuole of chloroplast-rich mesophyll protoplasts, but mostly in a small vacuole in chloroplast-poor protoplasts, while the large vacuole remained intact. Staining with the acid-trapped dye Neutral Red indicated that the fluorescent vacuoles had a neutral pH while the non-fluorescent vacuoles were acidic. One vacuole type was large and the other small, depending on the cell type (Di Sansebastiano *et al.*, 1998). On the other hand, there is little information about the vesicles transporting proteins with a ctVSS and little idea of the effect of wortmannin on the transport of aggregated proteins by dense vesicles. It is thus unclear whether there are two or three distinct transport pathways to two or three distinct vacuole types. This is further complicated by the possibilities of convergence of pathways to a single vacuole and of fusion of different vacuole types that may explain the finding of sporamin and barley lectin in aggregates in the same central vacuole (Schroeder *et al.*, 1993).

One protein may well be transported to two different vacuoles, such as the barley aspartic proteinase that is found both in protein storage vacuoles and in lytic vacuoles (Paris *et al.*, 1996). Its precursor contains both an NPIR-containing propeptide and an internal insertion of a saposin-like insert that was found to be involved in vacuolar targeting (Törmäkangas, 1997). The sporamin with a C-terminally transplanted propeptide also indicates this possibility (Koide *et al.*, 1997). It would be interesting to know whether it is transported to two different vacuoles or whether the VSR-dependent sorting is dominant. There is also the more general question of the spatial and temporal order of action of these sorting systems within the Golgi. A sorting system acting earlier (in a more *cis* compartment) would prevent sorting by a later-acting system.

Processing of targeting propeptides

Many vacuolar proteins are processed after their sorting. In many cases the processing would occur after arrival in

the vacuole. Vacuolar endoproteases involved in processing have been identified in several plants (Hara-Nishimura *et al*., 1991). One of these enzymes preferentially cuts Asn-Gly bonds (Scott *et al*., 1992), a bond frequently found in precursors of seed storage proteins, but also in tobacco chitinase and AP-24 (Fig. 2). This site-specific processing may be followed by carboxypeptidase trimming, which may lead to ragged ends (Nagahora *et al*., 1992). It was found that mutations of the Asn-Gly site did not affect processing of the chitinase, which could be processed exclusively by a carboxypeptidase (Freydl, 1995).

In contrast to the processing of seed storage proteins, tobacco chitinase and AP-24, the processing of NTPPs during the transport are more complex in both sporamin and aleurain, as intermediate forms were found during the maturation of these proteins (Holwerda *et al*., 1992; Nakamura *et al*., 1993). The presence of these intermediate precursors raises the question, whether all proteolytic cleavages occur in the vacuole. One possibility is that the processing of the most N-terminal part of the propeptide, usually containing the ssVSS, occurs just after the transport from the Golgi apparatus to the newly identified intermediate compartment between the Golgi apparatus and vacuole, the 'prevacuole' (Paris *et al*., 1997) or 'Pep12 compartment' (da Silva Conceição *et al*., 1997). Further processing, such as the activation of the protease aleurain by the removal of the bulk of the N-terminal propeptide, a protease inhibitor domain, could occur after arrival to the vacuole from this intermediate compartment.

Sorting of tonoplast proteins

Analysis of targeting of tonoplast proteins has been much more difficult than for the soluble vacuolar proteins. This may be due to the lack of known simple type I or type II proteins that made many advances possible in animal systems. The most abundant tonoplast proteins are integral membrane proteins (TIPs) with six transmembrane segments and both N- and C-termini on the cytosolic face of the tonoplast. The different TIP isoforms are located in different vacuoles in pea and barley root cells and may thus be sorted by different mechanisms (Paris *et al*., 1996). As for the soluble proteins, they can also be found in one and the same vacuole. The C-terminal transmembrane and cytosolic tail of α-TIP was found to be sufficient to target a reporter protein to the tonoplast, and truncation of the cytosolic tail did not affect this result (Höfte and Chrispeels, 1992). This could mean that the transmembrane segment alone carries the sorting information or that a tonoplast (of which vacuole type?) is the default destination for membrane proteins as found for yeast (Roberts *et al*., 1992; Wilcox *et al*., 1992). The C-terminal tail of α-TIP prevented transit through the Golgi of a chimaeric VSR protein (Jiang and Rogers, 1998). Differential sensitivity to inhibitors like monensin or brefeldin A also hinted at different sorting pathways for soluble and membrane proteins (Gomez and Chrispeels, 1993). The transport of TIP may thus involve a completely different pathway than for soluble proteins, a pathway avoiding altogether the Golgi apparatus. The extreme abundance of TIPs in tonoplasts suggests in fact a structural role, with each TIP subtype defining a different vacuolar compartment (discussed in Neuhaus and Rogers, 1998).

Conclusion

Protein transport and sorting in plant cells remains a field with many open questions. The fundamental mechanisms can be expected to be similar to other eukaryotes, as evidenced by the identification of homologues to proteins involved in protein trafficking in the Expressed Sequence Tag (EST) programs. However, sorting to the plant vacuole has specific features that need further clarification. The existence of several different vacuolar compartments, several different vacuolar sorting signals and the different sorting pathways postulated for tonoplast proteins all indicate that with plants there is much more to do than confirming and extending models developed for other systems. One family of sorting receptors has been identified, but it may only sort a fraction of the vacuolar proteins. One or more additional receptors still await identification. We are therefore looking forward to exciting new results that will clarify the biogenesis of the vacuoles and the function of these important organelles.

Acknowledgements

The authors would like to Nadine Paris for useful discussions. The work described here is supported by the Swiss National Science Foundation, Grant No. 31–46926.96 and by grants from the Ministry of Education, Science and Culture, Japan.

References

Baker A, Schatz G. 1987. Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. *Proceedings of the National Academy of Sciences, USA* **84,** 3117–3121.

Balch WE, Farquhar GM. 1995. Beyond bulk flow. *Trends in Cell Biology* **5,** 16–19.

Bannykh SI, Balch WE. 1997. Membrane dynamics at the endoplasmic reticulum-Golgi interface. *Journal of Cell Biology* **138,** 1–4.

Becker B, Melkonian M. 1996. The secretory pathway of protists—spatial and functional organization and evolution. *Microbiological Reviews* **60,** 697 ff.

da Silva Conceição A, Marty-Mazars D, Bassham DC, Sanderfoot AA, Marty F, Raikhel NV. 1997. The syntaxin

homolog AtPEP12p resides on a late post-Golgi compartment in plants. *The Plant Cell* **9,** 571–582.

Di Sansebastiano GP, Paris N, Marc-Martin S, Neuhaus J-M. 1998. Specific accumulation of GFP in a non-acidic vacuolar compartment via a C-terminal propeptide-mediated sorting pathway. *The Plant Journal* **15,** 449–457.

Dombrowski JE, Schroeder MR, Bednarek SY, Raikhel NV. 1993. Determination of the functional elements within the vacuolar targeting signal of barley lectin. *The Plant Cell* **5,** 587–596.

Freydl E. 1995. Analysis of the targeting peptide of chitinase. PhD thesis, University of Basel.

Gomez L, Chrispeels MJ. 1993. Tonoplast and soluble vacuolar proteins are targeted by different mechanisms. *The Plant Cell* **5,** 1113–1124.

Gomord V, Denmat LA, Fitchette-Lainé AC, Satiat-Jeunemaitre B, Hawes C, Faye L. 1997. The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *The Plant Journal* **11,** 313–325.

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 Letters* **294,** 89–93.

Hinz G, Menze A, Hohl I, Vaux D. 1997. Isolation of prolegumin from developing pea seeds—its binding to endomembranes and assembly into prolegumin hexamers in the protein storage vacuole. *Journal of Experimental Botany* **48,** 139–149.

Höfte H, Chrispeels MJ. 1992. Protein sorting to the vacuolar membrane. *The Plant Cell* **4,** 995–1004.

Höfte H, Faye L, Dickinson C, Herman EM, Chrispeels MJ. 1991. The protein-body proteins phytohemagglutinin and tonoplast intrinsic protein are targeted to vacuoles in leaves of transgenic tobacco. *Planta* **184,** 431–437.

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

Holwerda BC, Rogers JC. 1993. Structure, functional properties and vacuolar targeting of the barley thiol protease, aleurain. *Journal of Experimental Botany* **44,** 321–329.

Hong E, Davidson AR, Kaiser CA. 1996. A pathway for targeting soluble misfolded proteins to the yeast vacuole. *Journal of Cell Biology* **135,** 623–633.

Horazdovsky BF, DeWald DB, Emr SD. 1995. Protein transport to the yeast vacuole. *Current Opinion in Cell Biology* **7,** 544–551.

Ishikawa A, Ohta S, Matsuoka K, Hattori T, Nakamura K. 1994. A family of potato genes that encode Kunitz-type proteinase inhibitors: structural comparisons and differential expression. *Plant Cell Physiology* **35,** 303–312.

Jiang L, Rogers JC. 1998. Integral membrane protein sorting to vacuoles in plant cells: evidence for two pathways. *Journal of Cell Biology* **143,** 1183–1199.

Kaiser CA, Preuss D, Grisafi P, Botstein D. 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* **235,** 312–317.

Kirsch T, Paris N, Butler JM, Beevers L, Rogers JC. 1994. Purification and initial characterization of a potential plant vacuolar targeting receptor. *Proceedings of the National Academy of Sciences, USA* **91,** 3403–3407.

Kirsch T, Saalbach G, Raikhel NV, Beevers L. 1996. Interaction of a potential vacuolar targeting receptor with amino- and carboxyl-terminal targeting determinants. *Plant Physiology* **111,** 469–474.

Koide Y, Hirano H, Matsuoka K, Nakamura K. 1997. The N-terminal propeptide of the precursor to sporamin acts as a vacuole-targeting signal even at the C-terminus of the mature part in tobacco cells. *Plant Physiology* **114,** 863–870.

Levanony H, Rubin R, Altschuler Y, Galili G. 1992. Evidence for a novel route of wheat storage proteins to vacuoles. *Journal of Cell Biology* **119,** 1117–1128.

Lippincott-Schwartz J, Smith CL. 1997. Insights into secretory and endocytic membrane traffic using green fluorescent protein chimeras. *Current Opinion in Neurobiology* **7,** 631–639.

Matsuoka K, Bassham DC, Raikhel NV, Nakamura K. 1995. Different sensitivity to wortmannin of two vacuolar sorting signals indicates the presence of distinct sorting machineries in tobacco cells. *Journal of Cell Biology* **130,** 1307–1318.

Matsuoka K, Nakamura K. 1991. Propeptide of a precursor to a plant vacuolar protein required for vacuolar targeting. *Proceedings of the National Academy of Sciences, USA* **88,** 834–838.

Melchers LS, Sela-Buurlage MB, Vloemans SA, Woloshuk CP, van Roekel JSC, Pen J, van den Elzen PJM, Cornelissen BJC. 1993. Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and β-1,3-glucanase in transgenic plants. *Plant Molecular Biology* **21,** 583–593.

Nagahora H, Ishikawa K, Niwa Y, Muraki M, Jigami Y. 1992. Expression and secretion of wheat germ agglutinin by *Saccharomyces cerevisiae*. *European Journal of Biochemistry* **210,** 989–997.

Nakamura K, Matsuoka K, Mukumoto F, Watanabe N. 1993. Processing and transport to the vacuole of a precursor to sweet potato sporamin in transformed tobacco cell line BY-2. *Journal of Experimental Botany* **44,** 331–338.

Neuhaus J-M, Pietrzak M, Boller T. 1994. Mutation analysis of the C-terminal vacuolar targeting peptide of tobacco chitinase: low specificity of the sorting system, and gradual transition between intracellular retention and secretion into the extracellular space. *The Plant Journal* **5,** 45–54.

Neuhaus J-M, Rogers JC. 1998. Sorting of proteins to vacuoles in plant cells. *Plant Molecular Biology* **38,** 127–144.

Neuhaus J-M, Sticher L, Meins Jr F, Boller T. 1991. A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. *Proceedings of the National Academy of Sciences, USA* **88,** 10 362–10 366.

Okamoto T, Nakayama H, Seta K, Isobe T, Minamikawa T. 1994. Post-translational processing of a carboxy-terminal propeptide containing a KDEL sequence of plant vacuolar cysteine endopeptidase (SH-EP). *FEBS Letters* **351,** 31–34.

Paris N, Rogers SW, Jiang L, Kirsch T, Beevers L, Phillips TE, Rogers JC. 1997. Molecular cloning and further characterization of a probable plant vacuolar sorting receptor. *Plant Physiology* **115,** 29–39.

Paris N, Stanley CM, Jones RL, Rogers JC. 1996. Plant cells contain two functionally distinct vacuolar compartments. *Cell* **85,** 563–572.

Roberts CJ, Nothwehr SF, Stevens TH. 1992. Membrane protein sorting in the yeast secretory pathway: Evidence that the vacuole may be the default compartment. *Journal of Cell Biology* **119,** 69–83.

Robinson DG, Bäumer M, Hinz G, Hohl I. 1998. Vesicle transfer of storage proteins to the vacuole: the role of the Golgi apparatus and multivesicular bodies. *Journal of Plant Physiology* **152,** 659–667.

Rothman JE, Wieland FT. 1996. Protein sorting by transport vesicles. *Science* **272,** 227–234.

Saalbach G, Jung R, Kunze G, Saalbach I, Adler K, Müntz K. 1991. Different legumin protein domains act as vacuolar targeting signals. *The Plant Cell* **3,** 695–708.

Saalbach G, Rosso M, Schumann U. 1996. The vacuolar targeting signal of the 2S albumin from Brazil nut resides at

the C-terminus and involves the C-terminal propeptide as an essential element. *Plant Physiology* **112,** 975–985.

Saras J, Heldin CH. 1996. PDZ domains bind carboxy-terminal sequences of target proteins. *Trends in Biochemical Sciences* **21,** 455–458.

Schroeder MR, Borkhsenious ON, Matsuoka K, Nakamura K, Raikhel NV. 1993. Colocalization of barley lectin and sporamin in vacuoles of transgenic tobacco plants. *Plant Physiology* **101,** 451–458.

Scott MP, Jung R, Müntz K, Nielsen NC. 1992. A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean. *Proceedings of the National Academy of Sciences, USA* **89,** 658–662.

Shimada T, Kuroyanagi M, Nishimura M, Hara-Nishimura I 1997. A pumpkin 72-kDa membrane protein of precursor-accumulating vesicles has characteristics of a vacuolar sorting receptor. *Plant and Cell Physiology* **38,** 1414–1420.

Sonnewald U, von Schaewen A, Willmitzer L. 1990. Expression of mutant patatin protein in transgenic tobacco plants: role of glycans and intracellular location. *The Plant Cell* **2,** 345–355.

Sticher L, Hinz U, Meyer AD, Meins F. 1992. Intracellular transport and processing of a tobacco vacuolar β-1,3-glucanase. *Planta* **188,** 559–565.

Törmäkangas K. 1997. Structure, expression and intracellular targeting of barley aspartic proteinase. PhD thesis, University of Helsinki.

Vitale A, Bielli A, Ceriotti A. 1995. The binding protein associates with monomeric phaseolin. *Plant Physiology* **107,** 1411–1418.

Vitale A, Chrispeels MJ. 1992. Sorting of proteins to the vacuoles of plant cells. *Bioessays*, **14,** 151–160.

von Heijne G. 1990. The signal peptide. *Journal of Membrane Biology* **115,** 195–201.

von Schaewen A, Chrispeels MJ. 1993. Identification of vacuolar sorting information in phytohemagglutinin, an unprocessed vacuolar protein. *Journal of Experimental Botany* **44,** 339–342.

Wilcox CA, Redding K, Wright R, Fuller RS. 1992. Mutation of a tyrosine localization signal in the cytosolic tail of yeast Kex2 protease disrupts Golgi retention and results in default transport to the vacuole. *Molecular Biology of the Cell* **3,** 1353–1371.

Wilkins TA, Bednarek SY, Raikhel NV. 1990. Role of propeptide glycan in post-translational processing and transport of barley lectin to vacuoles in transgenic tobacco. *The Plant Cell* **2,** 301–313.

Wink M. 1993. The plant vacuole—a multifunctional compartment. *Journal of Experimental Botany* **44,** 231–246.

Young NM, Watson DC, Yaguchi M, Adar R, Arango R, Rodriguez-Arango E, Sharon N, Blay PKS, Thibault P. 1995. C-terminal post-translational proteolysis of plant lectins and their recombinant forms expressed in *Escherichia coli*—characterization of 'ragged ends' by mass spectrometry. *Journal of Biological Chemistry* **270,** 2563–2570.

KUNITZ-TYPE PROTEINASE INHIBITORS

Sporamin	...RF**NPIRLP**TTHEPA|...
Potato	...SQ**NLID***N**L**PS*|ESPV.....VNE**NPLDV**LFQEV-cooh
Potato	...SQ**NLID***L**P**S*|ESPL.....VNE**NPLDV**LFQEV-cooh
Potato	...SQ**NPI***N**L*|*P**S*DATP.....VKD**NPLDV**SFKQVQ-cooh
Potato	...SQ**NPI***N**L*|*P**S*DATP.....VKD**NPLDV**SFKQVQ-cooh
Potato	...SE**NPIVL**PTTCHDDDN|LVLPEVYDADG**NPLRI**GER...

PAPAIN-LIKE CYSTEINE PROTEASES

Aleurain	...SSSSFADS**NPIR**PVTDRAAST...
Oryzain γ	...ASSGFdDSNPIRSVTDHAASA...
Petunia	...RTANFADE**NPIR**QVVSDSFHE...
Tomato	...GPATFADK**NPIR**QVVFDLPDE...
Arabidopsis	...SDVNDGDD**LVIR**QVVGGAEPQ...
Pea	...TDDTNNDD**FIIR**QVVDNEEDH...

2S ALBUMINS

Pumpkin	...GIEN|PWRREG...KAR*NLPS*MCGIRP-QRCDF*
Pumpkin	...RAT*NLPS*VCRLSQ-RRCELRSSRW*
Castor bean	...RSDN|QERSLR...TAA*NLPS*MCGVSP-TECRF*
Rapeseed	...DMEN|PQGPQQ...TATHLPKVCNIPQVSVCPFQKTMPGPS|Y*
Arabidopsis	...DMEN|PQGQQQ...TAKHLPNVCDIPQVDVCPF*NIPS*FPS|FY*
Brazil nut	...YQTM|PRRGME...LAE*NIPS*RCNLSP-MRCPMGGS|IAGF*

CEREAL LECTINS

Rice	CYKGGDGMAAILAN*NQS*VSFEGIIESVA**E**LV
Barley	CDG|VFAEAIAANSTLVA**E**
Wheat	CDG|VFAEAITA*NST*LLQ**E**
Wheat	CDA|VFAGAITA*NST*LLA**E**
Wheat	CDG|VFAEAIAT*NST*LLA**E**

CHITINASES

Maize	FNN NSASLAGTAAHA**E**A
Maize	STA HPCWNRCSA**E**A
Pea	FGS SLPLSSILLDTVAAA
Alfalfa	FGS SLSLSSLFLNSI**D**T
Bean	FGN SLLLSDLVTSQ
Vigna	FGN SLLNLHPIV
Avocado	FGV STNPLAASS

Arabidopsis	FVN GLL**E**AAI
Cotton	FGN GVSV**D**SM
Ulmus	FGN GLLL**D**TM
Tomato	FGN GLLV**D**IM
Potato	FGN ALLV**D**TL
Potato	FGN GLLV**D**TV
Tobacco	FGN|GLLV**D**TM
Grapevine	FGS GLLL**D**TI
Poplar	FGY GLSGLK**D**TM
Poplar	FEDN GLLKMVGTM
GLUCANASES	
Bean	FG AQRMQRLLLMSSMQHIPLRVTCKLEPSSQSLL
Pea	FG GERRDGEIVEGDF*NGT*VSLKS**D**M
Alfalfa	FG G-ERMGIVNGDF*NAT*I-SLKS**D**M
N.plumbaginifolia	FG VSG-SVET*NAT*A-SLIS**E**I
N.glutinosa	FG VSGRV-DSSVET*NAT*A-SLIS**E**M
Tobacco	FG|VSGGVWDSSVETNATA-SLVS**E**M
Tomato	FG VSERVWDI-T*NST*ASSLTS**E**I
Potato	FG VSERVWDISAET*NST*TSSLIS**E**M
Hevea	FG A-EKNWDISTEH*NAT*ILFLKS**D**M
OSMOTINS	
Thaumatococcus	CP TALEL**ED**E
Tobacco	CPN GQAHP-NFPLEMP-GSD**E**VAK
Tobacco	CPY GSAHNETTNFPLEMPTSSTH**E**VAK
S.commersonnii	CPY GSTHNETTNFPLEMP-TSTL**E**VA
Tomato	CPN GVADP-NFPLEMP-ASTD**E**VAK
Arabidopsis	CPR SRLGATGSHQLPIKMVT**EE**N
PR-4	
Tobacco	CGDN MNVLVSPV**D**K**E**
Potato	CGDN VNVPLLSVV**D**K**E**
Tomato	CGDN VNVPLLSVV**D**R**E**
Hevea	CGDS FNPLFSVMKSSVIN**E**
Soybean	CGNE LDLTKPLLSIL**D**AP
Arabidopsis	CGNE LIGQPDSRNMLVSAI**D**RV
PR-1	
Tobacco	RPFG DLEEQPFDSKLELPT**D**V
Tomato	NVYG DLEEQKPDFDSKL**E**LPN