The Plant Cell, Vol. 2, 533–546, June 1990

A Short Domain of the Plant Vacuolar Protein Phytohemagglutinin Targets Invertase to the Yeast Vacuole

Brian W. Tague,[1] Craig D. Dickinson, and Maarten J. Chrispeels[2]

Department of Biology and Center for Molecular Genetics, C-016, University of California, San Diego, La Jolla, California 92093-0116

Phytohemagglutinin (PHA), the seed lectin of the common bean, accumulates in protein storage vacuoles of storage parenchyma cells in cotyledons. When expressed in yeast, PHA is efficiently targeted to the yeast vacuole [Tague and Chrispeels (1987). J. Cell Biol. 105, 1971–1979]. To identify vacuolar sorting information in PHA, a series of 3′ deletions of the PHA gene were fused in-frame to a truncated yeast invertase gene. An amino-terminal portion of PHA composed of a 20-residue signal sequence and 43 residues of the mature protein efficiently targeted invertase to the yeast vacuole. Internal deletions in a short PHA-invertase fusion showed that targeting information exists between residues 14 and 23 of mature PHA. Based on examinations of three-dimensional structures of related lectins, only a portion of these residues would be available on the surface of PHA for interaction with a putative receptor. Amino acid replacements at these positions in a PHA-invertase hybrid caused secretion of the invertase. The results indicate the presence of a vacuolar targeting domain in PHA that is centered at position 19 of the mature protein. This sequence of PHA also shows sequence identity to a vacuolar sorting domain characterized in yeast carboxypeptidase Y. Single amino acid alterations in a short PHA-invertase hybrid protein that caused the highest levels of secretion introduced a glycosylation site at position 21 of PHA. This observation suggests that glycan addition may interfere with recognition of a sorting determinant. These same amino acid changes did not dramatically increase secretion in a long PHA-invertase fusion or in PHA itself. Thus, a second domain of PHA may function in concert with the first one to bring about correct targeting of PHA.

INTRODUCTION

There is remarkable conservation of both general and specific features of the secretory pathway in eukaryotic cells. The general topology of secretory protein transport—endoplasmic reticulum to Golgi apparatus to the cell surface or vacuole/lysosome—is the same in all eukaryotic cells examined (Chrispeels, 1985; Schekman, 1985; Rose and Doms, 1988; Bradshaw, 1989; Jones and Robinson, 1989; Klausner, 1989). Many of the specific components necessary for this transport are also highly conserved (Dunphy et al., 1986; Pâquet et al., 1986; Segev et al., 1988; Thomas et al., 1988) such as the NSF protein/SEC18 gene product involved in vesicle fusion events (Beckers et al., 1989; Diaz et al., 1989; Wilson et al., 1989). The initial signal for entry into the secretory pathway, the signal sequence, has been well characterized (Walter and Lingappa, 1986). Evidence indicates that in the absence of additional targeting information or if that information is perturbed, a protein that enters the secretory pathway is secreted; that is, secretion directed by the signal sequence is the default pathway (Kelly, 1985; Moore and Kelly, 1986; Stevens et al., 1986; Pfeffer and Rothman, 1987; Wieland et al., 1987; Dorel et al., 1989). For proteins to be diverted from this pathway to the lysosome or vacuole or to be retained in the endoplasmic reticulum (ER) or Golgi, additional targeting information is needed.

This additional targeting information ultimately resides in the amino acid sequence of the protein. In some mammalian systems, the mechanism for lysosomal sorting is the well-characterized mannose 6-phosphate/mannose 6-phosphate receptor system (von Figura and Hasilik, 1986; Kornfeld, 1987; Kornfeld and Mellman, 1989). The addition of mannose 6-phosphate to certain glycans of lysosomal proteins is directed by a protein determinant (Lang et al., 1984). It has been shown, furthermore, that alternate, perhaps glycan-independent, pathways must exist (Gabel et al., 1983; Waheed et al., 1988). Sorting of hydrolases to vacuoles in yeast (for review, see Rothman et al., 1989b) is mediated by a protein determinant for carboxypeptidase

[1] Current address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114.

[2] To whom correspondence should be addressed.

Y (CPY) (Johnson et al., 1987; Valls et al., 1987) and for proteinase A (Klionsky et al., 1988). Evidence indicates that plants also utilize a carbohydrate-independent, presumably proteinaceous, vacuolar sorting signal (Bollini et al., 1985; Voelker et al., 1989).

By using the conservation of the secretory system as a basis, we have examined the possibility of using yeast as a system for identifying a plant vacuolar sorting signal. A microbiological system such as yeast would overcome the long times needed for analysis in transgenic plants and aid in the identification of mutations in a vacuolar sorting signal. Our work has focused on the vacuolar seed lectin, phytohemagglutinin (PHA), from the bean *Phaseolus vulgaris*. PHA is a mixed tetrameric glycoprotein (120 kD) of two subunits, PHA-E (erythroagglutinating) and PHA-L (leucoagglutinating). PHA polypeptides are co-translationally inserted into the ER, where they are core glycosylated twice per peptide and where they form oligomers. Subsequent transport to the Golgi apparatus results in the modification of one oligosaccharide to a complex type. PHA is then deposited in the protein storage vacuoles of the storage parenchyma cells of cotyledons (Chrispeels and Bollini, 1982; Chrispeels, 1983a, 1983b; Vitale et al., 1984). We have previously shown that when PHA-L (252 amino acids plus a 20-amino acid signal sequence, hereafter referred to as simply PHA) is expressed in *Saccharomyces cerevisiae* under control of the yeast acid phosphatase (*PHO5*) promoter, this plant vacuolar protein is transported to the yeast vacuole (Tague and Chrispeels, 1987).

In this report gene fusions were used to identify a domain of PHA responsible for vacuolar sorting in yeast. Gene fusions to the yeast secretory protein invertase have proved to be very powerful for the identification of yeast mitochondrial transit peptides (Emr et al., 1986), vacuolar sorting signals (Johnson et al., 1987; Klionsky et al., 1988), and for ER retention signals (Pelham et al., 1988). Two forms of invertase are produced in wild-type yeast from a single gene (*SUC*) by utilizing different mRNA initiation sites (Carlson and Botstein, 1982; Perlman et al., 1982). One form is constitutively produced, unglycosylated, and cytoplasmic. Under low glucose derepression, a second mRNA is synthesized that includes sequences coding for a signal sequence, leading to secretion of a highly glycosylated form of invertase. Invertase is an excellent marker because it can be easily assayed by a number of techniques; glycosylation can be used as a marker for entry into the secretory system; and invertase is both stable and active in the cytoplasm and periplasmic space, the normal locations, and in the yeast vacuole (Johnson et al., 1987).

We have generated a number of fusion genes between 3′ deletions of the gene for PHA and a truncated gene coding for yeast invertase. Analysis of the invertase activity directed by these fusions shows that the 43 NH_2-terminal amino acids of PHA are sufficient to sort invertase to the yeast vacuole. Derivatives of these fusions indicate that sequences in this region that bear homology to the carboxypeptidase Y sorting signal are involved in the sorting. Changes that result in high levels of invertase secretion in a short fusion construct, when introduced into a long fusion or intact PHA, do not result in higher levels of secretion. These results suggest the presence of multiple targeting domains in the PHA protein.

RESULTS

Construction of PHA-Invertase Fusion Genes

PHA-L, expressed in yeast under control of the yeast acid phosphatase (*PHO5*) promoter, is efficiently targeted to the yeast vacuole (Tague and Chrispeels, 1987). To identify vacuolar sorting information in the PHA protein, a series of gene fusions with the yeast invertase were constructed. The strategy used to produce the PHA-invertase fusion plasmids is shown in Figure 1 (see also Methods). We used unidirectional exonuclease III digestion of the PHA

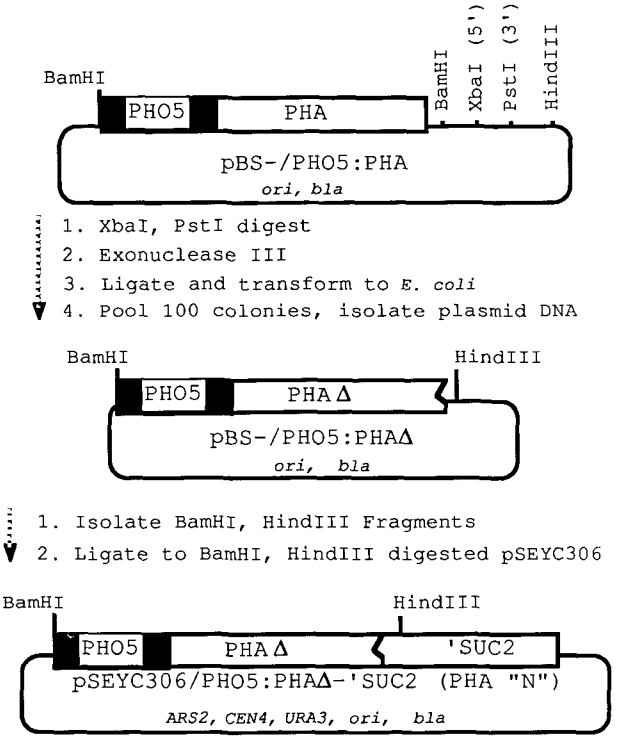

Figure 1. Construction of PHA-Invertase Fusion Plasmids.

The fusion plasmids are referred to as PHA N, where N is the final amino acid of PHA in the fusion constructs. Plasmids contain the following components as noted on the figure: *PHO5*, yeast acid phosphatase promoter; PHA, phytohemagglutinin-L gene; *ori*, bacterial origin of replication; *bla*, bacterial β-lactamase gene; *ARS2*, yeast autonomous replicating sequence; *CEN4*, yeast centromere sequence; and *URA3*, gene for selection in *ura3* yeast. See Methods for details.

gene to create a nested set of deletions from the 3′ end. Clones derived from the exonuclease III digestions were treated as pools to create PHA-invertase fusions with the vector pSEYC306 (Johnson et al., 1987). Each pool corresponded to a discrete time of exonuclease III digestion. Rather than analyze individual fusion constructs for correct in-frame fusion, PHA-invertase deletion constructs from each pool were transformed into yeast (strain 1403-1A). Individual yeast colonies were initially screened for invertase activity using a filter assay (Klionsky et al., 1988), which also gave qualitative information on the location of invertase activity (internal versus external).

Screening for Invertase Activity by a Filter Assay

To identify in-frame PHA-invertase fusions, we used a filter assay for invertase activity (Klionsky et al., 1988). Approximately 100 yeast colonies from each fusion plasmid transformation were picked to master SCM plates lacking uracil and grown overnight. These colonies were replica plated to minimal-fructose, low-phosphate plates to induce the *PHO5* promoter. After overnight induction, the plates were screened for invertase activity.

DNA was prepared from invertase-producing yeast clones from each pool and their plasmids were rescued in *Escherichia coli*. The recovered fusion plasmids were then analyzed by DNA sequencing as described in Methods. The PHA-invertase fusions chosen for further analysis were numbered according to the final amino acid of PHA in the fusion protein (initiating methionine equals number 1). For instance, the PHA-invertase fusion protein produced from plasmid PHA44 contained a 20-amino acid PHA signal peptide, 24 amino acids of mature PHA, followed by the yeast invertase, which lacks its own signal peptide. All of the fusion proteins contained the additional sequence Gly-Met-Gln-Ala-Phe between the PHA and the invertase sequence that resulted from the polylinker.

The filter assay screen for obtaining invertase-producing yeast involved replica plating the colonies to filter paper that had been soaked in sucrose and the glucose oxidase-horseradish peroxidase (HRP) assay mixture. Shown in Figure 2 are the results of this assay using the final fusion construct bearing yeast strains examined in this work. A number of controls are shown in the top row of each panel. In no instance did strain 1403-1A carrying the parent plasmid pSEYC306 show invertase activity. When whole cells were replica plated to the assay filter (Figure 2A), secreted invertase was detected, as shown for strain 20B12, a wild-type *SUC* strain, and for strain 1403-1A carrying the plasmids P4I-23 (Klionsky et al., 1988) and pCYI-20 (Johnson et al., 1987), both of which direct the constitutive production of secreted invertase. Pretreatment of the plates (Figure 2B) by a 5-min inversion of the plate over chloroform vapor to permeabilize the cells allowed detection of total (i.e., internal and external) invertase as shown for strain 1403-1A carrying plasmids P4I-

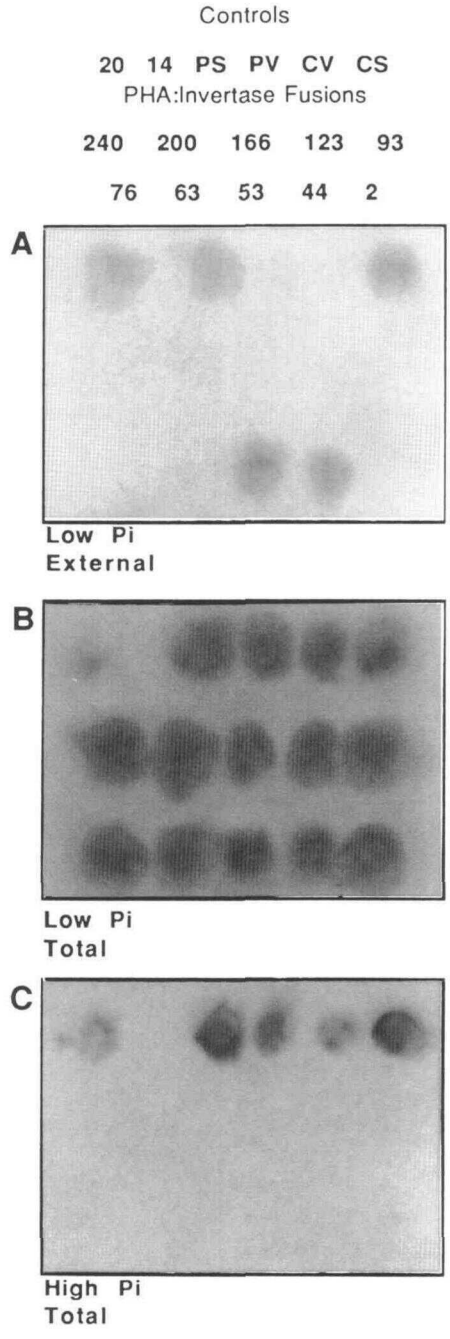

Figure 2. Filter Screen of PHA-Invertase Fusion Plasmids.

Individual yeast colonies harboring PHA-invertase fusion plasmids were screened for invertase activity as described in Methods. Control strains are: 20, 20B12 (*SUC*); 14, 1403-1A (*suc*) harboring parent plasmid pSEYC306; PS (P4I-23) and PV (P4I-137) (Klionsky et al., 1988); CV (pCYI-50) and CS (pCYI-20) (Johnson et al., 1987). PHA-invertase fusions are numbered according to the final amino acid of PHA in the fusion construct (initiating methionine numbered 1).

(A) Low P_i external invertase.

(B) Low P_i total invertase.

(C) High P_i total invertase.

137 (Klionsky et al., 1988) or pCYI-50 (Johnson et al., 1987), both of which directed the constitutive production of vacuolar invertase in yeast.

The filter assay provided preliminary indications of the location of the invertase resulting from various PHA-invertase fusions. Those fusions that contained more than 63 amino acids, including the 20-amino acid signal sequence, directed the synthesis of internal invertase. Fifty-three or 44 amino acids of PHA directed detectable secreted invertase (compare total, 2B and external, 2A). A hybrid with only 2 amino acids of PHA and that, therefore, lacked the PHA signal peptide, resulted in invertase activity that was internal and presumably cytoplasmic. As an additional control, it was shown that the invertase production was dependent on the activity of the phosphate-inducible *PHO5* promoter and was independent of glucose concentration (Figure 2C). Invertase activity was detected with the control stains that constitutively produce invertase fusion proteins and for the low glucose-derepressed wild-type strain 20B12. Yeast harboring the PHA-invertase fusion plasmids grown on high inorganic phosphate did not produce invertase activity, internally or externally, as expected. Although in theory 33% of the yeast colonies screened should produce invertase, in practice 15% to 25% of the URA (able to grow without uracil) yeast were positive for invertase activity. The invertase filter assay identified yeast producing invertase and gave qualitative information on the location of this activity.

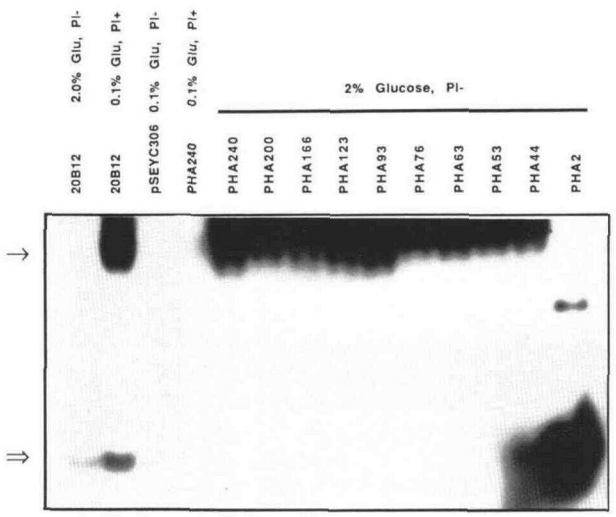

Figure 3. Native Invertase Activity Gel.

Yeast cells harboring fusion plasmids were analyzed for the glycosylation of invertase on a native acrylamide gel as described in Methods. First two lanes show wild-type *SUC* strain 20B12 grown under repressed (2.0% glucose) and derepressed (0.1% glucose) conditions. Remaining lanes show invertase activity from strain 1403-1A harboring the indicated plasmids and cultured in the indicated conditions (2.0% or 0.1% glucose; Pi^-, 10 μg/mL inorganic phosphate; Pi^+, 1500 μg/mL inorganic phosphate).

Activity Gels of PHA-Invertase Fusions

A second assay, native invertase activity gels, provided additional information about invertase. Two forms of invertase are produced in wild-type yeast cells: an unglycosylated, constitutive cytoplasmic form and a highly glycosylated glucose-repressible secreted form. These two forms can be separated on a native acrylamide gel and the activity visualized (Gabriel and Wang, 1969; Carlson et al., 1981), as shown in Figure 3. Yeast strain 20B12 displays the unglycosylated form when cultured in 2.0% glucose and both forms when cultured in 0.1% glucose. No activity was present in the parent strain carrying plasmid pSEYC306 cultured in conditions which induce both the wild-type *SUC* gene (0.1% glucose) and the *PHO5* (low P_i) promoters. The PHA-invertase fusion plasmids produced no activity in low-glucose, high-phosphate medium (shown for PHA240). In low phosphate all of the fusion plasmid bearing yeast except PHA2 produced highly glycosylated invertase activity. PHA44 produced a small amount of unglycosylated activity, most likely due to the in-frame ATG from the multiple cloning site between PHA and invertase in these constructs. PHA2 produced an unglycosylated invertase, indicating that it had not entered the secretory system. The higher molecular weight activity in PHA2 has not been further characterized but the discrete nature suggests that it is not heavily glycosylated and is most likely a higher order oligomer of the enzyme (Williams et al., 1985; Tammi et al., 1987). The glycosylation of the larger PHA-invertase fusions that contained the PHA signal peptide indicated entry into the yeast secretory pathway. Given that PHA itself is transported to the yeast vacuole, we have taken internal glycosylated invertase activity directed by PHA-invertase fusion plasmids as initial evidence of a vacuolar location.

Quantitation of Invertase Activity

The fusion-bearing strains were next quantitatively analyzed for the location of invertase activity using the coupled glucose oxidase-HRP procedure (Goldstein and Lampen, 1975). When whole cells were assayed by this method, only secreted invertase was detected because the plasma membrane of yeast is impermeable to sucrose. Disruption of cells by vortexing in the presence of glass beads allowed the detection of total invertase. When grown in low phosphate medium (see Methods), many of the strains produced remarkably high levels of invertase activity (up to 1000 U), and up to 25% of the activity was secreted in strains producing the large fusions. It has been shown that overproducing the vacuolar protein CPY by increasing gene copy number leads to elevated levels of secretion

presumably due to overloading of the sorting machinery (Stevens et al., 1986). We reasoned that the high level of invertase secretion may be due to the high total production of invertase activity driven by the strong *PHO5* promoter.

The rate of production and the total accumulation of acid phosphatase from the *PHO5* promoter can be modulated by the presence of phosphate in the medium (Toh-E et al., 1973). This modulation by phosphate was used to decrease the level of total invertase produced in the PHA93 stain. As shown in Table 1, adding phosphate to the medium decreased the total amount of invertase activity. More importantly, the amount of invertase secretion decreased with decreasing invertase production, down to 10% in the presence of 40 μg/mL phosphate. Therefore, the quantitative invertase assays were performed on strains that had been grown in the presence of 40 μg/mL phosphate. Changes in phosphate concentration did not affect the localization of PHA2 or PHA44 invertase activity (data not shown).

As shown in Table 2, when at least 63 amino acids of PHA-L were linked to invertase, 10% or less of the total invertase was secreted. In these cases, the invertase was highly glycosylated, indicating that the invertase was probably transported to the yeast vacuole. When only 53 amino acids of PHA were fused to invertase, greater than 20% of the activity was secreted; this figure does not change by further lowering the total invertase activity (data not shown). As shown below, the remainder of the activity in PHA53 was associated with the vacuole, indicating that this fusion still contained sorting information. The PHA44 construct secreted about 80% of the total activity; the remaining activity was not vacuolar but was probably cytoplasmic, as described above. The secretion of invertase from strain PHA44 indicated that the sequences important for vacuolar localization had been lost or impaired in this construct. Finally, PHA2 directed little external invertase activity as expected for cytoplasmically localized enzyme.

A second sensitive test for invertase secretion is the ability of yeast strains to utilize sucrose. Because the plasma membrane is impermeable to sucrose, invertase must be secreted to utilize this sugar as the sole carbon source. Yeast strains producing the PHA-invertase fusions were tested on bromcresol purple indicator plates (Sherman et al., 1986) containing 2% sucrose and 10 μg/mL antimycin A. On plates containing 40 μg/mL phosphate, all of the strains except PHA2 showed a *SUC* phenotype, as might be expected due to the amounts of invertase secreted as detailed in Table 2. We found, however, that at 80 μg/mL phosphate, secretion was low enough that only PHA44 and PHA53 displayed a *SUC* phenotype.

Table 1. Effect of Added Phosphate on Invertase Secretion

	Invertase Activity[a]		
P_i (μg/mL)[b]	Total	External	% Secretion
10	912	200	22
20	719	137	19
30	566	79	14
40	366	37	10

[a] Invertase activity is in nanomoles of glucose produced per minute per OD_{600} yeast cells at 30°C.
[b] The low-phosphate medium contains 10 μg/mL P_i (see Methods); KH_2PO_4 was added to reach the desired level of total P_i.

Table 2. Invertase Activity Produced by PHA-Invertase Fusions

	Invertase Activity			SUC
Plasmid	Total	External	% Secretion	Phenotype[a]
PHA240	484	40	8	−
PHA200	489	38	8	−
PHA166	362	21	6	−
PHA123	518	39	8	−
PHA93	366	37	10	−
PHA76	679	40	6	−
PHA63	611	44	7	−
PHA53	637	148	23	+
PHA44	442	374	79	+
PHA2	754	12	2	−
pSEYC306[b]	2	0	0	−
PHA240[c]	2	0	0	ND[d]

[a] SUC phenotype determined on minimal 2% sucrose plates containing bromcresol purple, 10 μg/mL antimycin A, and 80 μg/mL P_i.
[b] Grown in 0.1% glucose, 10 μg/mL P_i.
[c] Grown in 0.1% glucose, 1500 μg/mL P_i.
[d] ND, not determined.

A Short Domain of PHA Sufficient for Vacuolar Sorting

To determine whether the larger PHA-invertase fusions that were internal were in fact localized to the vacuoles, the vacuoles from a number of the yeast strains were isolated and analyzed (Figure 4). Vacuoles were isolated by lysing spheroplasts with gentle osmotic shock and floating vacuoles up through a Ficoll density gradient. In all cases 25% to 30% of the α-mannosidase and carboxypeptidase Y (two yeast vacuole markers) activities were recovered in the vacuole fraction; these fractions contained relatively little of the marker enzyme activity for the ER (NADPH-cytochrome *c* reductase) or for the cytoplasm (α-glucosidase). For yeast strains carrying fusion plasmids coding for more than 63 amino acids of PHA, the invertase activity co-fractionated with the vacuole marker activities, indicating that for these strains greater than 90% of invertase was directed to the yeast vacuole. Plasmid PHA53 also directed vacuolar invertase. In this case about 20%

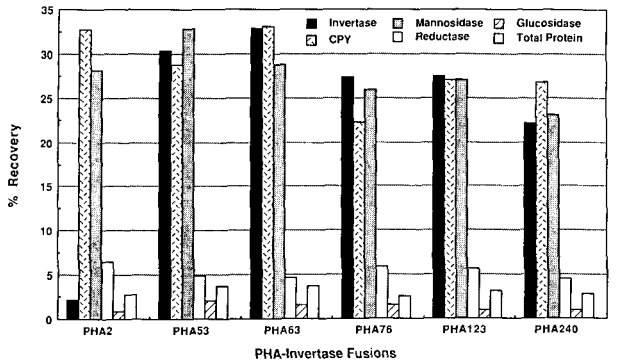

Figure 4. Vacuolar Localization of Fusion Proteins.

Vacuoles were isolated by Ficoll buoyant density centrifugation (see Methods) from fusion plasmid bearing strains listed at bottom of figure. The vacuole fractions were analyzed for the following marker enzyme activities: invertase (PHA-invertase fusions); carboxypeptidase Y (CPY) and α-mannosidase (vacuole markers); NADPH-cytochrome *c* reductase (ER marker), α-glucosidase (cytoplasmic marker), and total protein. Plotted is the percent of total activity of the gradient recovered in the vacuole fraction.

of the whole cell activity was lost during spheroplast formation (results not shown). This suggested a periplasmic or cytoplasmic location of 20% of the total invertase, consistent with the results shown in Table 2. The 80% of the activity that remained in spheroplasts cofractionated with vacuoles, indicating that the vacuolar sorting signal contained in PHA53 or the recognition of this signal was slightly impaired. Little invertase was found in the vacuoles of strain PHA2; the invertase activity cofractionated with α-glucosidase confirming a cytoplasmic location (results not shown). Vacuoles were also isolated from the PHA44 strain. About 85% of the invertase activity was lost during spheroplasting as expected for a secreted form of invertase. Only about 4% of the whole cell invertase from PHA44 was found in the purified vacuole fraction (results not shown).

Fusion Plasmid Derivatives

To further define vacuolar sorting information contained in the PHA63 fusion, internal deletions were introduced in the PHA sequences of the gene encoding PHA63 protein. The construction of these derivatives is described in Methods. Three separate derivatives of PHA63 contained deletions of residues 24 to 40 (PHA63Δ24-40), residues 44 to 52 (PHA63Δ44-52), and residues 23 to 33 (PHA63Δ23-33). Each of these constructs was introduced into strain 1403-1A, and the plasmid-bearing strains were analyzed for the location of invertase (internal versus external). Deletion of amino acids 23 to 33 or of amino acids 44 to 52 had little effect on the level of invertase secretion as compared with PHA63. Deletion of amino acids 24 to 40 had a dramatic effect on the secretion level: 7% for PHA63, 93% for PHA63Δ24-40. Although the deletion derivatives were not analyzed directly for vacuolar location of internal invertase activity, these results, together with that for PHA53, indicate that vacuolar sorting information occurs within amino acids 34 to 43.

Figure 5 summarizes the data for the shorter NH_2-terminal constructs and derivatives and gives the sequences of PHA in the fusions. At the top of the figure is the predicted structure of PHA (see Discussion) based on the crystal structure of three legume seed lectins. The amino acids in this region are in the form of β-sheet structure and two loops that are predicted to be at the surface of the folded PHA protein. Analysis of the crystal structures revealed that a portion of the short domain implicated by the fusion protein analysis is external on the protein where it is potentially available for interaction with a sorting receptor. This further narrowed the search for the targeting domain to residues 38 to 42 of PHA.

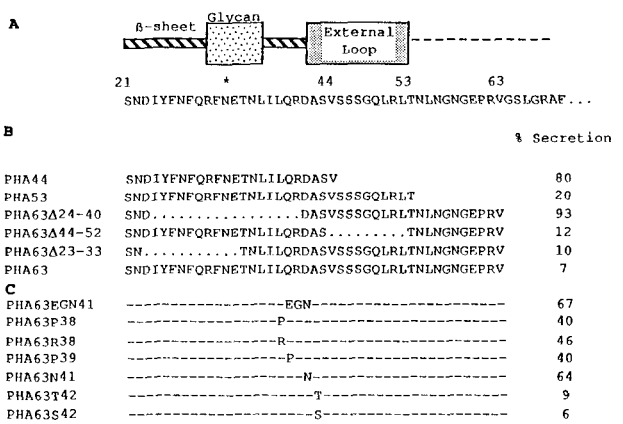

Figure 5. Structure of PHA and Sequences of Fusion Plasmids.

(A) Amino acid sequence and predicted structure of mature PHA based on the crystal structures of three homologous legume lectins (see Discussion). Site of asparagine-linked oligosaccharide side-chain is marked by an asterisk (*).

(B) Sequences of PHA fusion proteins. Final amino acids in the fusion constructs directed by plasmids PHA63, PHA53, PHA44, and PHA63 derivatives were determined by DNA sequencing. Location of PHA-invertase fusions were determined by quantitative invertase assays and/or co-fractionation of invertase activity with vacuolar markers.

(C) Amino acid modifications directed to the putative vacuolar targeting determinant of the fusion protein PHA63 and the effect of these changes on secretion of invertase activity. The changes were introduced by site-directed mutagenesis as described in Methods.

To test directly the importance of this sequence in sorting, the region was mutagenized by random oligonucleotide mutagenesis of the fusion PHA63. The results of this experiment are shown in Figure 5C. Changes of both Leu^{40} and Gln^{41} resulted in up to 40% secretion of invertase as compared with 7% for the wild-type PHA63, whereas changes of Ser^{45} had little effect. The most dramatic increase in invertase secretion (64%) was seen in Asp^{43} to Asn. Interestingly, this change introduced a glycosylation site (Asn-Ala-Ser^{45}) that may block recognition of this sequence because of the presence of a glycan. Similar results were obtained in a PHA63 derivative in which Gln-Arg-Asp^{41} was changed to Glu-Gly-Asn (EGN) also introducing a glycosylation site. This change resulted in 67% invertase secretion as compared with 7% for the original construct. The results demonstrate that sorting information exists at this location in PHA.

Further evidence that this sequence contains targeting information comes from sequence identity between this region of PHA and the yeast carboxypeptidase Y vacuolar sorting determinant. The region of CPY containing the sorting determinant is shown in Figure 6 and contains the sequence LQR^{25} [numbering from the initiator methionine (Valls et al., 1987)], which overlaps with the sorting determinant QRPL^{27} (Rothman et al., 1989b). This LQR sequence is also found in PHA-L at position 38. The same sequence can be found in a number of other plant vacuolar proteins such as phaseolin, pea legumin, and ricin (Figure 6). Although the sequence is not strictly conserved in all legume lectins, the Gln (Q^{39} in PHA-L) is found in all legume lectins sequenced to date (Etzler, 1985; Yarwood et al., 1985; Strosberg et al., 1986), a number of which are shown in Figure 6. Mutational analysis of the CPY sorting domain indicates that the Gln in the QRPL^{27} sequence is required for correct sorting (Valls et al., 1987).

```
                                     ****
                   .    .       .    .         . .          ..
PHAL         SNDIYFNFQRF--NETNLI  LQRDASV-SSSGQLRLTNLNGNGEPRVG-SLGR...
PHAE         ASQTSFSFQRF--NETNLI  LQRDATV-SSKGQLRLTNVNDNGEPTLS-SLGR...
Soybean      AETVSFSWNKFVPKQPNMI  LQGDAIVTSSG-KLQLNKVDENGTPKPS-SLGR...
ConA     STHETNALHFMFNQFSKDQKDLI  LQGDA-TTGTEGNLRLTRVSSNGSPQGS-SVGR...
PeaLec       TETTSFLITKFSPDQQNLI  FQGDGYTT-KE-KLTLTKAVKN-------TVGR...
Favin       TDEITSFSIPKFRPDQPNLI  FQGGGYTT-KE-KLTLTKAVKN-------TVGR...

              (N-proximal)
Phaseolin     SDNSWNTLFKNQYGHIRV  LQRFDQQSKRLQNL...
              (C-proximal)
Legumin       TSSVINDLPLDVVAATFK  LQRDEARQLKSNN...
              (C-proximal)
Ricin...      LSTAIQESNQGARASPIQ  LQRDGSKFSVYDV...

                                1 (pro)
CPY (yeast)                     IS  LQRPLGLDKDVLL...
```

Figure 6. Sequence Homology between CPY and Plant Vacuolar Proteins.

Sequences for lectins [PHA-L, PHA-E, soybean lectin, concanavalin A, pea lectin, favin, and ricin (Etzler, 1985; Hoffman and Donaldson, 1985; Lamb et al., 1985; Strosberg et al., 1986)] and two storage proteins (Slightom et al., 1983; Argos et al., 1985) as well as the CPY vacuolar sorting domain (Valls et al., 1987) are shown. The sequences for the legume lectins begin at the first residue after the signal sequence cleavage site. Residues conserved in all legume lectins sequenced to date are indicated by a dot over the PHA-L sequence. The four asterisks indicate the location of a conserved motif.

Multiple Sorting Domains in PHA?

The QRD^{41} to EGN^{41} mutation, which caused a dramatic increase in secretion in a short fusion (PHA63), was also introduced into a longer fusion PHA240 for analysis in yeast. With PHA240, however, the change was not as dramatic: 8% of the invertase was secreted in the original PHA240 construct; the EGN^{41} change resulted in an increase in invertase secretion to 28%. This change was also introduced into the (unfused) PHA gene and expressed in yeast. To examine whether or not PHA was secreted, yeast cells expressing this mutant gene were spheroplasted; proteins were prepared from the spheroplasts and from the periplasmic space and were analyzed by protein gel blotting, as shown in Figure 7. It is apparent that the glycosylation site introduced by the mutagenesis was utilized by the cellular glycosylation machinery because of the slower mobility of the PHA-EGN^{41} compared with PHA expressed in yeast. The same result was obtained for the PHA-N^{41} protein, which contained only the Asp^{41} to Asn change (not shown). However, as is the case with the expression of wild-type PHA in yeast (Tague and Chrispeels, 1987), very little PHA was detected in the periplasmic space. This demonstrated that the PHA with the EGN^{41} mutation was not secreted. Figure 7 also shows that the Gln^{39} to Pro change does not result in secretion of PHA. The results indicate that the CPY homologous domain is sufficient, but not strictly required, for vacuolar sorting of PHA in yeast. We infer from this that PHA contains more than one vacuolar sorting determinant.

DISCUSSION

Why Analyze a Plant Protein in Yeast?

We have used the conservation of the eukaryotic secretory pathway to develop a system for identifying and analyzing the vacuolar sorting signal of a plant protein, phytohemagglutinin, using the yeast *S. cerevisiae*. The reasons for using yeast are twofold. To do the same analysis in plants would take much longer because of the necessity of performing the analysis in seeds rather than in vegetative tissue. The large central vacuoles of leaf cells are fragile and difficult to isolate in high yield and contain very potent proteases (Boller and Kende, 1979; Lin and Wittenbach, 1981; Sherman et al., 1986). For example, Beachy and coworkers have expressed the soybean storage vacuole protein conglycinin in transgenic petunia under control of

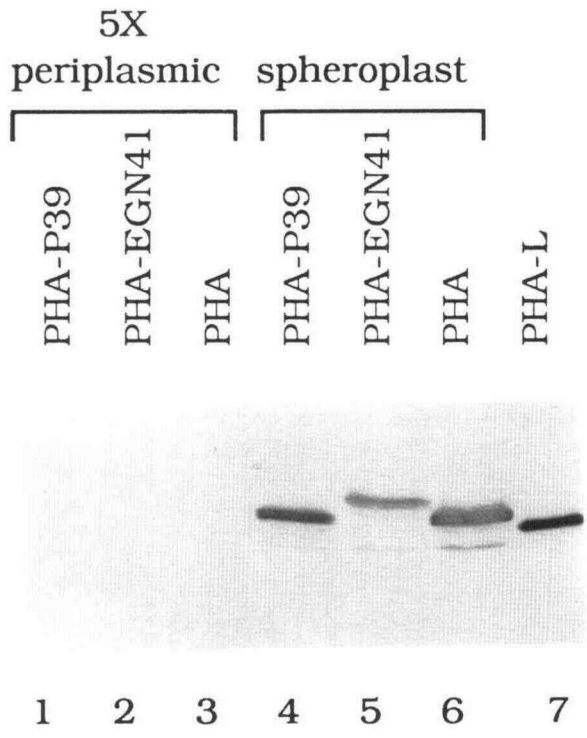

Figure 7. Immunoblot Showing Localization of PHA and Modified Forms of PHA in Yeast.

Yeast strain 29B5 carrying pYE7:PHA (Tague and Chrispeels, 1987) (lanes 3 and 6), pYE7:PHA-P39 with the Gln^{39} to Pro change (lanes 1 and 4), and pYE7:PHA-EGN41 with the Gln-Arg-Asp^{41} to Glu-Gly-Asn change (lanes 2 and 5) were induced by growth at 22°C for 18 hr. Total proteins from spheroplast fractions (lanes 4 through 6) and periplasmic fractions (lanes 1 through 3) were prepared from these cells as described previously (Tague and Chrispeels, 1987). 0.5 OD units of spheroplast fractions and 2.5 OD units of periplasmic fractions were loaded onto a 12.5% SDS gel and analyzed by immunoblotting. Lane 7 contains 50 ng of *P. vulgaris* PHA-L.

the non-tissue-specific cauliflower mosaic virus 35S promoter (Beachy et al., 1986). Although the mRNA levels for conglycinin were similar in leaf and seed tissue, 20 to 50 times less protein was found in the leaf samples, and this protein was not of the mature size. These results were interpreted to indicate lower stability of conglycinin in leaf vacuoles (della-Cioppa et al., 1987). In contrast, the nature of the protein storage vacuoles of seeds makes them the choice for analyzing plant vacuolar sorting because they contain little or no protease activity (Harris and Chrispeels, 1975; Baumgartner et al., 1978). However, with the most developed and consistent transgenic plant system, tobacco, 4 to 6 months are needed to obtain seeds. Yeast offers the possibility of a rapid identification of plant vacuolar sorting signals. Time is not the only problem. At present a more finely detailed analysis of sorting signal mutants by screening or selection is difficult if not impossible in plants. Yeast offers all the obvious advantages of a microbiological system, and many details of vacuolar biogenesis in yeast have been studied (Rothman et al., 1989b).

A Short Domain of PHA Is Sufficient for Correct Targeting

Construction and initial analysis of fusion protein-producing vectors was rapid with the yeast system. Random PHA-invertase fusions were constructed by producing deletions of the PHA coding sequences and inserting these sequences into the vector pSEYC306, which carried a truncated invertase gene (Figure 1). The deletion plasmids were then treated as pools of plasmids, and yeast colonies containing in-frame fusions were identified by the invertase filter assay method (Klionsky et al., 1988; Figure 2). This screen was additionally useful for its ability to differentiate between external and total (i.e., internal) invertase activity. In this way it was possible to determine that sequences close to the NH_2 terminus of PHA were sufficient to maintain internal invertase activity (Figures 2A and 2B).

Evidence for entry into the secretory pathway was obtained by using an invertase gel activity assay to demonstrate glycosylation of invertase (Figure 3). As expected, fusion proteins carrying at least the 20-amino acid signal sequence of PHA underwent the extensive glycosylation characteristic of yeast invertase. The combination of the random construction of fusions and these two invertase assays demonstrated that sequences very close to the NH_2 terminus of PHA were responsible for producing internal glycosylated invertase, and this was taken as initial qualitative evidence of vacuolar location.

To quantify the location of invertase activity, assays were performed on liquid cultures (Table 2) and on purified vacuoles (Figure 4). These results confirmed the results obtained in the initial screening. If at least 63 amino acids of PHA (43 amino acids of the mature protein) were fused to invertase, greater than 90% of the invertase activity cofractionated with the yeast vacuole. The sorting to the vacuole was slightly impaired if only 53 amino acids were fused to invertase and was essentially abolished when only 44 amino acids (24 of the mature protein) were included.

Elevated Levels of Expression Result in Missorting

The sorting of PHA-invertase fusions was similar to endogenous yeast sorting, at least in terms of overproducing the protein, which led to elevated levels of secretion for both the PHA-invertase fusions (Table 1) and for the endogenous yeast vacuolar protein, CPY (Stevens et al., 1986). As we have argued previously (Tague and Chrispeels, 1987), we believe that the sorting of PHA and now PHA

fusions to the yeast vacuole is an active, directed process, particularly given that secretion appears to be the default pathway. The PHA signal may differ in its efficiency of sorting: 6% to 10% of the invertase activity is secreted in the vacuolar fusion constructs, whereas for CPY-invertase (Johnson et al., 1987) and proteinase A-invertase (Klionsky et al., 1988) fusions the secretion level is closer to 2% to 4%. Inefficiency may be expected when targeting a heterologous protein but it should be noted that the average total invertase activity from the *PHO5* promoter-driven fusions (Table 2) is higher than that from the two yeast proteinase promoters (Johnson et al., 1987; Klionsky et al., 1988). Thus, the higher level of secretion with the PHA fusions may be due to the higher total production (Table 1), consistent with the ability to demonstrate a *suc* phenotype of the vacuolar sorting fusions when the yeast were grown on higher levels of inorganic phosphate (Table 2).

Comparisons with Related Lectins

There are partial or complete sequence data for about 30 legume seed lectins (Etzler, 1985; Yarwood et al., 1985; Strosberg et al., 1986). The lectins are homologous, with sequence identity ranging from 40% to 90%. Additionally, about 35 amino acids out of an average of 250 in the mature protein are absolutely conserved. Three of these lectins have been crystallized and their structures have been determined. These lectins are concanavalin A [ConA, from *Canavalis ensiformis* (Becker et al., 1975; Hardman and Goldstein, 1977)], favin [from *Vicia faba* (Reeke and Becker, 1986)], and pea lectin [from *Pisum sativum* (Einspahr et al., 1986)]. Pea lectin and favin are closely related (76% identity) but are more distantly related to ConA (39% and 37% identity, respectively). Nevertheless, the root-mean-square distance between corresponding alpha carbons in the tertiary structures is 1.4 Å (Reeke and Becker, 1986) for the favin-ConA comparison and only 0.87 Å (Einspahr et al., 1986) for the pea lectin-ConA comparison. It is expected that all legume seed lectins, including PHA (about 40% identity with each of these three lectins) will share the same general tertiary structure.

The predicted structure of PHA is shown at the top of Figure 5. At the mature NH_2 terminus, the structure is β-sheet followed by two loops. It is unlikely that the β-sheet (residues 1 to 9 of the mature protein) that is involved in extensive monomer-monomer contacts is the information-carrying portion. The first loop structure carries the high-mannose N-linked glycan attached to asparagine 12 (Sturm and Chrispeels, 1986). The structure of this loop may be altered when glycosylation is inhibited with tunicamycin or is abolished by substituting threonine 14 of the glycosylation site with alanine by site-directed mutagenesis. Neither tunicamycin treatment (Bollini et al., 1985) nor site-directed mutagenesis (Voelker et al., 1989) affects the vacuolar sorting of PHA in bean or in transgenic tobacco, respectively. The second loop is predicted to be on the surface of the PHA molecule on the basis of the crystal structure of the other legume lectins and it is in the area implicated in sorting by the fusion protein analysis.

Localization of a Targeting Determinant in PHA

The 63 NH_2-terminal residues of PHA present in PHA63 were found to be sufficient for the sorting of invertase to the yeast vacuole. Analysis of further deletions within PHA63 indicated that targeting information occurred in PHA in the region between residues 34 and 43. Based on crystal structure examination of related lectins, only a subset of these residues would be exposed on the surface of PHA and available for interaction with a receptor. Therefore, the sequences between PHA residues 38 and 43 were postulated to contain sorting information. This prediction was tested by targeting single residue changes to this region of PHA63. Changes at Leu^{38}, Gln^{39}, and Asp^{41} caused dramatic increases in the levels of secretion. The effect of the single residue changes indicates that this region is recognized as a bona fide yeast vacuolar sorting sequence and that the sorting is not simply dependent, for example, on the length of the PHA protein fused to invertase.

Although some of the single amino acid changes in PHA63 caused dramatic increases in secretion, none of those tested completely inhibited vacuolar targeting. It may be that there is not a high sequence specificity required for a vacuolar targeting determinant but that more general structural features are recognized by the sorting apparatus. This would distinguish a vacuolar sorting determinant from the highly conserved ER retention signal (Pelham et al., 1988).

An unexpected result obtained by random mutagenesis of the putative vacuolar targeting determinant in PHA63 was the creation of a glycan addition site. The modification resulted in the highest level of secretion observed for any of the single residue changes. That the site was recognized by the glycosylation machinery was demonstrated by testing the modification in PHA (Figure 7). It is possible that the addition of a glycan at the location of the targeting determinant inhibited its recognition by the sorting machinery. This result suggests that a function for glycans in secreted proteins may be, in some cases, to mask "accidental" vacuolar targeting signals. Such a view would be consistent with a low specificity requirement in a targeting determinant and the possible occurrence of multiple targeting determinants in PHA and in the yeast proteinase A (Klionsky et al., 1988).

Multiple Targeting Determinants in PHA?

The QRD^{41} to EGN^{41} mutation, which caused high levels of secretion in PHA, was also introduced into the longest fusion, PHA240, where the effect was not as great: inver-

tase secretion increased from 7% to 28%. Introduction of the EGN mutation into full-length PHA did not give dramatic effects on the level of secretion when the mutant gene was expressed in yeast. Similarly, the Gln^{39} to Pro change, which caused increased secretion in PHA63, did not appear to affect targeting when the change was tested in full-length PHA. Therefore, it appears that although the domain in PHA63 is sufficient for vacuolar sorting, it is not strictly required for the sorting of PHA in yeast. A second, perhaps independent, signal may be present C-terminal to the signal found in these experiments.

One property of the PHA mutants that we were unable to test is their stability in periplasm. Because PHA-L has neither methionine nor cysteine, we were unable to conduct satisfactory pulse-chase experiments to test the stability of PHA itself or of the mutants. Pulse-labeling with ^{14}C-amino acids did not result in a high enough level of radioactivity in PHA.

METHODS

Reagents

Restriction enzymes were obtained from Bethesda Research Laboratories or from Boehringer Mannheim and were used as recommended by the manufacturers. Exonuclease III, T4 DNA ligase, and DNA polymerase (Klenow) were from Bethesda Research Laboratories. Mung bean nuclease was obtained from Promega-Biotec. Antimycin A was from Calbiochem. Ficoll 400 was from Pharmacia LKB Biotechnology Inc. RepliPlate colony transfer pads were obtained from FMC BioProducts. The Sequenase DNA sequencing kit was from United States Biochemical Corporation. 4-Chloro-1-naphthol was from Bio-Rad. The following reagents were obtained from Sigma: lyticase, *N*-benzoyl-L-tyrosine *p*-nitroanilide, cytochrome *c*, glucose oxidase, NADPH, triphenyltetrazolium chloride, horseradish peroxidase, *p*-nitrophenyl-α-D-mannoside, *p*-nitrophenyl-α-D-glucopyranoside, and bromcresol purple.

Strains and Media

Escherichia coli strains JA221 (*recA1*, *leuB-6*, *trpAE5*, *hsd* R^-M^+, *lacY*) and HB101 [*pro*, *leu*, *thi*, *lacY*, *hsd20*, *endA*, *recA*, *rpsL20* (*str*^{r}), *ara*-14, *galK2*, *xyl-5*, *mtl-1*, *supE44*] were used for plasmid maintenance. The F^+, *dut*$^-$, *ung*$^-$ *E. coli* strain RZ1032, used to obtain uracil containing bacteriophage M13 DNA, was originally obtained by our laboratory from W. Wickner. *Saccharomyces cerevisiae* strain 1403-1A (*a*, *suc*, *gal3*, *gal4*, *trp1*, *ura3*) was obtained from the Yeast Genetic Stock Center, Berkeley, CA. *S. cerevisiae* strain 20B12 (*a*, *pep4-3*, *trp1*) (Jones, 1977) was used as a wild-type *SUC* strain and was a kind gift from D. Brow.

E. coli medium was as described in Maniatis et al. (1982). Yeast-rich YEPD medium, synthetic complete medium lacking uracil, and bromcresol purple indicator plates were prepared as described (Sherman et al., 1986). Minimal fructose plates were prepared as described (Klionsky et al., 1988). Medium lacking phosphate was prepared as described (Rubin, 1975). This method selectively precipitates inorganic phosphate as $MgNH_4PO_4$ and has little effect on organic phosphates. Ten milliliters of 1 M $MgSO_4$ and 10 mL of concentrated ammonium hydroxide were added to 1 L of the appropriate medium. After 0.5 hr at room temperature, the $MgNH_4PO_4$ was removed by filtering over Whatmann 1 filter paper. The medium was adjusted to pH 5.8 with HCl and sterilized. The low-phosphate medium was estimated to contain 10^{-4} M inorganic phosphate, i.e., approximately 10 μg/mL. This low-phosphate medium induced the acid phosphatase (*PHO5*) promoter to the same extent as synthetic minimal medium lacking phosphate (Bostian et al., 1980). When it was necessary to increase the inorganic phosphate concentration, KH_2PO_4 was added to the medium.

Construction of Fusion Plasmids

Standard recombinant DNA protocols (Maniatis et al., 1982) were used unless otherwise noted. Plasmids pSEYC306, pCYI-20, pCYI-50, P4I-23, and P41-137 were kind gifts from S. Emr. Plasmid pSEYC306 (Johnson et al., 1987) is a yeast/*E. coli* shuttle vector that contains a promoterless, truncated invertase gene, beginning at the codon for amino acid 3 (threonine) of mature invertase, placed behind a multiple cloning site. pCYI-20 and pCYI-50 (Johnson et al., 1987) are pSEYC306 derivatives containing portions of the *PRC1* gene coding for yeast carboxypeptidase Y that lead to secretion and vacuolar location of invertase activity, respectively. Similarly, P4I-23 and P4I-137 (Klionsky et al., 1988) are derivatives of pSEYC306 that contain sufficient information from the *PEP4* gene (Ammerer et al., 1986) coding for yeast proteinase A to lead to secretion and vacuolar location of the invertase activity.

Plasmid pYE7/PHA-L (Tague and Chrispeels, 1987) was used as a starting point for construction of the PHA-invertase fusions. A BamHI fragment containing the inorganic phosphate-repressible acid phosphatase (*PHO5*) promoter and the complete coding region of PHA-L was cloned into the plasmid pBS^- (StrataGene; Figure 1). This plasmid was linearized with XbaI (5′ overhang, ExoIII sensitive) and PstI (3′ overhang, ExoIII resistant) and treated with ExoIII as recommended by the StrataGene manual. Treatment at room temperature resulted in a deletion rate of about 50 bp/min. A time course of digestion created a nested set of deletions up to the 5′ end of the PHA coding sequence. Each timepoint results in fairly discrete endpoints of deletion as determined by gel electrophoresis. The DNA from each timepoint was made blunt ended with mung bean nuclease, ligated with T4 DNA ligase, and used to transform *E. coli* to ampicillin resistance.

Approximately 100 ampicillin-resistant colonies from each timepoint were pooled and used to prepare plasmid DNA. These deletion plasmids were digested with BamHI and HindIII, and the *PHO5* promoter/PHA-L deletion fragments were isolated from agarose gels. These fragments were directionally ligated into pSEYC306, which had been digested with the same two restriction enzymes, and the resulting plasmids were used to transform *E. coli*. Again, approximately 100 ampicillin-resistant colonies were pooled. Plasmid DNA from these pools was used to transform *S. cerevisiae* strain 1403-1A to uracil prototrophy by the LiCl method (Ito et al., 1983). A similar protocol has been used to generate Bal31-deleted invertase gene fusions to bacterial β-galactosidase (Emr et al., 1984).

Isolation of Fusion Plasmids and Sequencing

To rescue fusion plasmids of interest from yeast, total yeast DNA was prepared (Sherman et al., 1986). The total yeast DNA was used to transform *E. coli* to ampicillin resistance; routinely 5 to 10 ampicillin-resistant colonies were obtained from 100 ng of yeast DNA. Plasmid DNA from *E. coli* was purified on CsCl gradients and sequenced using the double-stranded DNA sequencing protocol described in the Sequenase manual (United States Biochemical Corporation). The sequencing primer (5′-CCAGCCCTTGTTGGGTGTGA-3′) covered bases 98 to 117 of the invertase gene (numbering system in Taussig and Carlson, 1983) and was complementary to the RNA-like strand of the gene. The primer is approximately 40 bases 3′ to the HindIII site of pSEYC306 used to join the PHA deletions to invertase. The primer was prepared by Louis Schmidt at the Center for Molecular Genetics, UCSD.

Construction of Fusion Plasmid Derivatives

Derivatives of fusion plasmids were constructed as follows. For construction of PHA63Δ23-33, the EcoRV site, which cuts in-frame between codons D^{22} and I^{23}, was joined to a filled-in FokI site, which is in-frame at D^{40}. The approximately 740-bp BamHI-HindIII fragment of PHA63 was subcloned to pBS$^-$ cut with the same enzymes. This plasmid (pBS$^-$/PHA63) was restricted with FokI (a Type-III restriction enzyme); the DNA fragment containing the COOH-terminal portion of PHA, the HindIII site, and a portion of the cloning vector was isolated. After filling in the FokI overhang with Klenow, this fragment was cleaved with HindIII, yielding an approximately 70-bp fragment containing the coding sequences for PHA amino acids 41 to 63 and a larger fragment consisting of pBS$^-$ sequences. This mixture was ligated to the pBS$^-$/PHA63 that had been cut with EcoRV and HindIII. After transformation and selection, plasmids were screened to obtain the correct pBS$^-$/PHA63Δ24-40 derivative. The BamHI-HindIII fragment from this plasmid was ligated to pSEYC306 cut with the same enzymes to generate PHA63Δ24-40.

Other derivatives of fusion plasmids were generated by site-directed oligonucleotide mutagenesis according to Kunkel (1985). The BamHI-HindIII fragment from PHA63 or from PHA240 was subcloned into the replicative form of M13mp18 and uracil-containing single-stranded DNA isolated from bacteriophages grown on the *dup*$^-$,*ung*$^-$ host RZ1032. The single-stranded DNA had the PHA coding sequences in the antisense orientation. The synthetic oligonucleotide 5′-GCAAACTCAAGCAACACCAACCTTATCC (prepared by Gary Sumnicht, San Diego State University) was used to delete the PHA coding sequences for amino acids 23 to 33, the oligonucleotide 5′-CGCGATGCCAGTACTAATCTTAATGGC was used to delete sequences coding for amino acids 44 to 52, and oligonucleotide 5′-CCTTATCCTCGAGGCCAATGCCTCCG was used to change the sequence QRD[41] to EGN. The oligonucleotides were used to prime second-strand synthesis of the uracil-containing single-stranded DNA by *E. coli* polymerase I Klenow fragment. Transformation of the reaction product into a *dut*$^+$,*ung*$^+$ host resulted in efficient production of mutant derivatives, which were screened for the loss of the EcoRV site, the addition of a ScaI site, or the addition of an XhoI site, respectively. All three oligonucleotides were used to mutagenize PHA63; only the QRD to EGN mutation was introduced into PHA240. After confirmation of the changes by DNA sequencing, the BamHI-HindIII fragment of the mutagenized plasmids was ligated back into pSEYC306 for analysis.

Site-directed mutagenesis of the region of PHA63 from Leu^{39} to Ser^{43} was performed with a degenerate mutagenic primer to obtain single residue changes at six different codon positions in the PHA-invertase hybrid. The BamHI-HindIII fragment of PHA63 (mentioned previously) was ligated into the phasmid vector pMa5-8 (P. Stanssens et al., Plant Genetic Systems, Gent, Belgium) for mutagenesis by the gapped duplex DNA method (Kramer et al., 1984). The mutagenic primer was a 30-mer composed of the sequence 5′-GAGACGGAGGCATCGCGTTGGAGGATAAGG with degeneracies occurring at the 12 underlined positions. Primer synthesis was carried out so that substitutions at the degenerate positions would occur at a rate of 8.1% (2.7% for each of the three substituting nucleotides). With this amount of degeneracy, 38% of the oligonucleotides were predicted to contain single base substitutions and 36% were predicted to contain no substitutions. Clones with single residue changes were isolated directly by DNA sequencing or by screening transformed yeast clones for external invertase activity. In the latter case, a library of *E. coli* clones was generated by reintroducing the mutagenized BamHI-HindIII fragments into pSEYC306 as a pool of DNA prepared from more than 5000 independent clones. Plasmid DNA from this library was transformed into *S. cerevisia* strain 1403-1A, and 250 of the resulting clones were assayed for external invertase by the liquid assay described below. Selected yeast transformants were rescued in *E. coli* and analyzed by DNA sequencing.

Assays

For screening individual yeast colonies for invertase activity, the filter plate assay was used as described (Klionsky et al., 1988) with the modification that 1 mg/mL 4-chloro-1-naphthol was used in place of *o*-dianisidine as the horseradish peroxidase (HRP) substrate. Invertase activity gels using triphenyltetrazolium chloride were performed as described in Gabriel and Wang (1969) with the modifications of Carlson et al. (1981).

Quantitative invertase assays using a coupled glucose oxidase-HRP reaction were performed as described in Goldstein and Lampen (1975) with some modifications. Yeast cultures were centrifuged and the cell pellet was washed twice and resuspended in 0.1 M $NaCOOCH_3$. The resuspended cell culture was split in two. One portion was assayed in 100 μL of 0.1 M $NaCOOCH_3$, 0.1 M sucrose to detect external invertase. The second portion was vortexed in the presence of glass beads before the assay to detect total invertase. After addition of 100 μL of phosphate buffer, pH 7, and boiling for 3 min, glucose produced in these assays was detected by the coupled glucose oxidase/HRP procedure (Goldstein and Lampen, 1975) using 1 mg/mL 4-chloro-1-naphthol as the HRP color reagent. The reaction was stopped by the addition of 0.5 mL of MeOH and 1.5 mL of 6 N HCl and the absorbance read at 550 nm.

Spheroplasts of yeast were prepared using lyticase as described in Scott and Schekman (1980) at a concentration of 50 units/OD_{600} yeast cells in 50 mM Tris-HCl, pH 7.5, 0.6 M sorbitol, 10 mM β-mercaptoethanol, and 2 mM NaN_3. Spheroplasts were lysed and vacuoles were purified by buoyant density centrifugation on a stepped Ficoll gradient as described previously (Tague and Chrispeels, 1987) with the modification that fractions were pooled

to yield a vacuole fraction, the load volume ("cytoplasmic" fraction), and a pellet. These fractions and the original spheroplast lysate were analyzed for invertase and enzyme marker activities (α-mannosidase, α-glucosidase, CPY, NADPH-cytochrome *c* reductase, total protein) as previously described (Tague and Chrispeels, 1987).

ACKNOWLEDGMENTS

We would like to thank D. Brow for providing the yeast strain 20B12, Patrick Stanssens for the mutagenesis vectors, and Scott Emr for the kind gift of the plasmids used in this work. We would also like to thank Scott Emr, Thomas Stevens, Russell Doolittle, and Jay Carberry for helpful discussions; Gary Sumnicht and Louise Schmidt for preparation of the oligonculeotides; and Bobbie Bradford for expert technical assistance. This work was supported by grants from the Department of Energy (DE-FG03-86ER13497) and from the National Science Foundation (DMB-88 17211) and from the University of California Biotechnology Research and Training program.

Received April 9, 1990.

REFERENCES

Ammerer, G., Hunter, C.P., Rothman, J.H., Saari, G.C., Valls, L.A., and Stevens, T.H. (1986). *PEP4* gene of *Saccharomyces cerevisiae* encodes proteinase A, a vacuolar enzyme required for processing of vacuolar precursors. Mol. Cell. Biol. **6,** 2490–2499.

Argos, P., Narayana, S.V.L., and Nielsen, N.C. (1985). Structural similarity between legumin and vacilin storage proteins from legumes. EMBO J. **4,** 1111–1117.

Baumgartner, B., Tokuyasu, K.T., and Chrispeels, M.J. (1978). Localization of vicilin peptidohydrolase in the cotyledons of mung bean seedlings by immunofluorescence microscopy. J. Cell Biol. **79,** 10–19.

Beachy, R.N., Chen, Z.-L., Bray, E., Lawton, M., Komeda, Y., Nakamura, I., Naito, S., Anderson, E., Dube, P., Rogers, S.G., and Fraley, R.T. (1986). Expression of subunits of β-conglycinin in transgenic tobacco and petunia plants. In Molecular Biology of Seed Storage Proteins and Lectins, L.M. Shannon and M.J. Chrispeels, eds (Rockville, MD: American Society of Plant Physiologists), pp. 193–201.

Becker, J.W., Reeke, G.N., Jr., Wang, J.L., Cunningham, B.A., and Edelman, G.M. (1975). The covalent and three-dimensional structure of concanavalin A. J. Biol. Chem. **250,** 1513–1524.

Beckers, C.J.M., Block, M.R., Glick, B.S., Rothman, J.E., and Balch, W.E. (1989). Vesicular transport between the endoplasmic reticulum and the Golgi stack requires the NEM-sensitive fusion protein. Nature **339,** 397–398.

Boller, T., and Kende, H. (1979). Hydrolytic enzymes in the central vacuole of plant cells. Plant Physiol. **63,** 1123–1132.

Bollini, R., Ceriotti, A., Daminati, M.G., and Vitale, A. (1985). Glycosylation is not needed for the intracellular transport of phytohemagglutinin in developing *Phaseolus vulgaris* cotyledons and for the maintenance of its biological activities. Physiol. Plant **65,** 15–22.

Bostian, K.A., Lemire, J.M., Cannon, L.E., and Halvorson, H.O. (1980). In vitro synthesis of repressible yeast acid phosphatase: Identification of multiple mRNAs and products. Proc. Natl. Acad. Sci. USA **77,** 4504–4508.

Bradshaw, R.A. (1989). Protein translocation and turnover in eukaryotic cells. Trends Biochem. Sci. **14,** 276–279.

Carlson, M., and Botstein, D. (1982). Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. Cell **28,** 145–154.

Carlson, M., Osmond, B.C., and Botstein, D. (1981). Mutants of yeast defective in sucrose utilization. Genetics **98,** 25–40.

Chrispeels, M.J. (1983a). The Golgi apparatus mediates the transport of phytohemagglutinin to the protein bodies in bean cotyledons. Planta **158,** 140–151.

Chrispeels, M.J. (1983b). Incorporation of fucose into the carbohydrate moiety of phytohemagglutinin in developing *Phaseolus vulgaris* cotyledons. Planta **157,** 454–461.

Chrispeels, M.J. (1985). The role of the Golgi apparatus in the transport and post-translational modification of vacuolar (protein body) proteins. Oxford Surv. Plant Mol. Cell Biol. **2,** 43–68.

Chrispeels, M.J., and Bollini, R. (1982). Characteristics of membrane-bound lectin in developing *Phaseolus vulgaris* cotyledons. Plant Physiol. **70,** 1425–1428.

della-Cioppa, G., Kishore, G.M., Beachy, R.N., and Fraley, R.T. (1987). Protein trafficking in plant cells. Plant Physiol. **84,** 965–968.

Diaz, R., Mayorga, L.S., Weidman, P.J., Rothman, J.E., and Stahl, P.D. (1989). Vesicle fusion following receptor-mediated endocytosis requires a protein active in Golgi transport. Nature **339,** 398–400.

Dorel, C., Voelker, T.A., Herman, E.M., and Chrispeels, M.J. (1989). Transport of proteins to the plant vacuole is not by bulk flow through the secretory system, and requires positive sorting information. J. Cell Biol. **108,** 327–337.

Dunphy, W.G., Pfeffer, S.R., Clary, D.O., Wattenberg, B.W., Glick, B.S., and Rothman, J.E. (1986). Yeast and mammals utilize similar cytosolic components to drive protein transport through the Golgi complex. Proc. Natl. Acad. Sci. USA **83,** 1622–1626.

Einspahr, H., Parks, E.H., Suguna, K., Subramanian, E., and Suddath, F.L. (1986). The crystal structure of pea lectin at 3.0-Å resolution. J. Biol. Chem. **261,** 16518–16527.

Emr, S.D., Schauer, I., Hansen, W., Esmon, P., and Schekman, R. (1984). Invertase β-galactosidase hybrid proteins fail to be transported from the endoplasmic reticulum in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4,** 2347–2355.

Emr, S.D., Vassarotti, A., Garrett, J., Geller, B.L., Takeda, M., and Douglas, M.G. (1986). The amino terminus of the yeast F1-ATPase β-subunit precursor functions as a mitochondrial import signal. J. Cell Biol. **102,** 523–533.

Etzler, M.E. (1985). Plant lectins: Molecular and biological aspects. Annu. Rev. Plant Physiol. **36,** 209–234.

Gabel, C.A., Goldberg, D.E., and Kornfeld, S. (1983). Identification and characterization of cells deficient in the mannose 6-phosphate receptor: Evidence for an alternative pathway for

lysosomal enzyme targeting. Proc. Natl. Acad. Sci. USA **80,** 775–779.

Gabriel, O., and Wang, S. (1969). Determination of enzymatic activity in polyacrylamide gels. I. Enzymes catalyzing the conversion of nonreducing substrates to reducing products. Anal. Biochem. **27,** 545–554.

Goldstein, A., and Lampen, J.O. (1975). β-D-Fructofuranoside fructohydrolase from yeast. Methods. Enzymol. **42,** 504–511.

Hardman, K.D., and Goldstein, I.J. (1977). The Structure and Activity of Concanavalin A. (New York: Plenum Press).

Harris, N., and Chrispeels, M.J. (1975). Histochemical and biochemical observations on storage protein metabolism and protein body autolysis in cotyledons of germinating mung beans. Plant Physiol. **56,** 292–299.

Hoffman, L.M., and Donaldson, D.D. (1985). Characterization of two *Phaseolus vulgaris* phytohemagglutinin genes closely linked on the the chromosome. EMBO J. **4,** 883–889.

Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153,** 163–168.

Johnson, L.M., Bankaitis, V.A., and Emr, S.D. (1987). Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease. Cell **48,** 875–885.

Jones, E.W. (1977). Proteinase mutants of *Saccharomyces cerevisiae*. Genetics **85,** 23–33.

Jones, R.L., and Robinson, D.G. (1989). Protein secretion in plants. New Phytol. **111,** 567–597.

Kelly, R.B. (1985). Pathways of protein secretion in eukaryotes. Science **230,** 25–32.

Klausner, R.D. (1989). Sorting and traffic in the central vacuolar system. Cell **57,** 703–706.

Klionsky, D.J., Banta, L.M., and Emr, S.D. (1988). Intracellular sorting and processing of a yeast vacuolar hydrolase: Proteinase A propeptide contains vacuolar targeting information. Mol. Cell. Biol. **8,** 2105–2116.

Kornfeld, S. (1987). Trafficking of lysosomal enzymes. FASEB J. **1,** 462–468.

Kornfeld, S., and Mellman, I. (1989). The biogenesis of lysosomes. Annu. Rev. Cell Biol. **5,** 483–525.

Kramer, W., Drutsa, V., Jansen, H., Kramer, B., Pflugfelder, M., and Fritz, H.J. (1984). The gapped duplex DNA approach to oligonucleotide-directed mutation construction. Nucl. Acids Res. **12,** 9441–9456.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82,** 488–492.

Lamb, F.I., Roberts, L.M., and Lord, J.M. (1985). Nucleotide sequence of cloned cDNA coding for preproricin. Eur. J. Biochem. **148,** 265–270.

Lang, L., Reitman, M., Tang, J., Roberts, R.M., and Kornfeld, S. (1984). Lysosomal enzyme phosphorylation: Recognition of a protein-dependent determinant allows specific phosphorylation of oligosaccharides present on lysosomal enzymes. J. Biol. Chem. **259,** 14663–14671.

Lin, W., and Wittenbach, V.A. (1981). Subcellular localization of proteases in wheat and corn mesophyll protoplasts. Plant Physiol. **67,** 969–972.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Moore, H.H., and Kelly, R.B. (1986). Re-routing of a secretory protein by fusion with human growth hormone sequences. Nature **321,** 443–446.

Pâquet, M.R., Pfeffer, S.R., Burczak, J.D., Glick, B.S., and Rothman, J.E. (1986). Components responsible for transport between successive Golgi cisternae are highly conserved in evolution. J. Biol. Chem. **261,** 4367–4370.

Pelham, H.R.B., Hardwick, K.G., and Lewis, M.J. (1988). Sorting of soluble ER proteins in yeast. EMBO J. **7,** 1757–1762.

Perlman, D., Halvorson, H.O., and Cannon, L.E. (1982). Presecretory and cytoplasmic invertase polypeptides encoded by distinct mRNAs derived from the same structural gene differ by a signal sequence. Proc. Natl. Acad. Sci. USA **79,** 781–785.

Pfeffer, S.R., and Rothman, J.E. (1987). Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. Annu. Rev. Biochem. **56,** 829–852.

Reeke, G.N., Jr., and Becker, J.W. (1986). Three-dimensional structure of favin: Saccharide binding—Cyclic permutation in leguminous lectins. Science **234,** 1108–1111.

Rose, J.K., and Doms, R.W. (1988). Regulation of protein export from the endoplasmic reticulum. Annu. Rev. Cell Biol. **4,** 257–288.

Rothman, J.H., Howald, I., and Stevens, T.H. (1989a). Characterization of genes required for protein sorting and vacuolar function in the yeast *Saccharomyces cerevisiae*. EMBO J. **8,** 2057–2065.

Rothman, J.H., Yamashiro, C.T., Kane, P.M., and Stevens, T.H. (1989b). Protein targeting to the yeast vacuole. Trends Biochem. Sci. **14,** 347–350.

Rubin, G.M. (1975). Preparation of RNA and ribosomes from yeast. In Methods in Cell Biology, Volume XII: Yeast Cells, D.M. Prescott, ed (New York: Academic Press), pp. 45–64.

Schekman, R. (1985). Protein localization and membrane traffic in yeast. Annu. Rev. Cell Biol. **1,** 115–143.

Scott, J.H., and Schekman, R. (1980). Lyticase: Endoglucanase and protease activities that act together in yeast cell lysis. J. Bacteriol. **142,** 414–423.

Segev, N., Mulholland, J., and Botstein, D. (1988). The yeast GTP-binding YPT1 protein and a mammalian counterpart are associated with the secretion machinery. Cell **52,** 915–924.

Sherman, G., Fink, R., and Hicks, J.B. (1986). Methods in Yeast Genetics. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Slightom, J.L., Sun, S.M., and Hall, T.C. (1983). Complete nucleotide sequence of a French bean storage protein gene: Phaseolin. Proc. Natl. Acad. Sci. USA **80,** 1897–1901.

Stevens, T.H., Rothman, J.H., Payne, G.S., and Schekman, R. (1986). Gene dosage-dependent secretion of yeast vacuolar carboxypeptidase Y. J. Cell Biol. **102,** 1551–1557.

Strosberg, A.D., Buffard, D., Lauwereys, M., and Foriers, A. (1986). Legume lectins: A large family of homologous proteins. In The Lectins: Properties, Functions, and Applications in Biology and Medicine, I.E. Liener, N. Sharon, and I.J. Goldstein, eds (New York: Academic Press), pp. 249–264.

Sturm, A., and Chrispeels, M.J. (1986). The high mannose oligosaccharide of phytohemagglutinin is attached to asparagine 12 and the modified oligosaccharide to asparagine 60. Plant Physiol. **81,** 320–322.

Tague, B.W., and Chrispeels, M.J. (1987). The plant vacuolar protein, phytohemagglutinin, is transported to the vacuoles of transgenic yeast. J. Cell Biol. **105,** 1971–1979.

Tammi, M., Ballou, L., Taylor, A., and Ballou, C.E. (1987). Effect of glycosylation of yeast invertase oligomer stability. J. Biol. Chem. **262,** 4395–4401.

Taussig, R., and Carlson, M. (1983). Nucleotide sequence of the yeast *SUC2* gene for invertase. Nucl. Acids Res. **11,** 1943–1954.

Thomas, G., Thorne, B.A., Thomas, L., Allen, R.G., Hruby, D.E., Fuller, R., and Thorner, J. (1988). Yeast KEX2 endopeptidase correctly cleaves a neuroendocrine prohormone in mammalian cells. Science **241,** 226–230.

Toh-E, A., Ueda, Y., Kakimoto, S., and Oshima, Y. (1973). Isolation and characterization of acid phosphatase mutants in *Saccharomyces cerevisiae*. J. Bacteriol. **113,** 727–738.

Valls, L.A., Hunter, C.P., Rothman, J.H., and Stevens, T.H. (1987). Protein sorting in yeast: The localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide. Cell **48,** 887–897.

Vitale, A., Ceriotti, A., Bollini, R., and Chrispeels, M.J. (1984). Biosynthesis and processing of phytohemagglutinin in developing bean cotyledons. Eur. J. Biochem. **141,** 97–104.

Voelker, T.A., Herman, E.M., and Chrispeels, M.J. (1989). In vitro mutated phytohemagglutinin genes expressed in tobacco seeds: Role of glycans in protein targeting and stability. Plant Cell **1,** 95–104.

von Figura, K., and Hasilik, A. (1986). Lysosomal enzymes and their receptors. Annu. Rev. Biochem. **55,** 167–193.

Waheed, A., Gottschalk, S., Hille, A., Krentler, C., Pohlmann, R., Braulke, T., Hauser, H., Geuze, H., and von Figura, K. (1988). Human lysosomal acid phosphatase is transported as a transmembrane protein to lysosomes in transfected baby hamster kidney cells. EMBO J. **7,** 2351–2358.

Walter, P., and Lingappa, V.R. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. Annu. Rev. Cell Biol. **2,** 499–516.

Wieland, F.T., Gleason, M.L., Serafini, T.A., and Rothman, J.E. (1987). The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell **50,** 289–300.

Williams, R.S., Trumbly, R.J., MacColl, R., Trimble, R.B., and Maley, F. (1985). Comparative properties of amplified external and internal invertase from yeast *SUC2* gene. J. Biol. Chem. **260,** 13334–13341.

Wilson, D.W., Wilcox, C.A., Flynn, G.C., Chen, E., Kuang, W.-J., Henzel, W.J., Block, M.R., Ullrich, A., and Rothman, J.E. (1989). A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. Nature **339,** 355–359.

Yarwood, A., Richardson, M., Sousa-Cavada, B., and Rougé, P. (1985). The complete amino acid sequences of the β1- and β2-subunits of the isolectins LoL1 and LoL2 from seeds of *Lathyrus ochrus* (L.) DC. FEBS Lett. **184,** 104–109.