MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 121–128
0270-7306/87/010121-08$02.00/0

Vol. 7, No. 1

Signal Peptide Specificity in Posttranslational Processing of the Plant Protein Phaseolin in *Saccharomyces cerevisiae*†

JANE HARRIS CRAMER,[1]* KRISTI LEA,[1] MICHAEL D. SCHABER,[2] AND RICHARD A. KRAMER[2]

Agrigenetics Advanced Science Co., Madison, Wisconsin 53716,[1] and Department of Molecular Genetics, Roche Research Center, Hoffmann-La Roche Inc., Nutley, New Jersey 07110[2]

Received 7 July 1986/Accepted 23 September 1986

We linked the cDNA coding region for the bean storage protein phaseolin to the promoter and regulatory region of the *Saccharomyces cerevisiae* repressible acid phosphatase gene (*PHO5*) in multicopy expression plasmids. Yeast transformants containing these plasmids expressed phaseolin at levels up to 3% of the total soluble cellular protein. Phaseolin polypeptides in *S. cerevisiae* were glycosylated, and their molecular weights suggested that the signal peptide had been processed. We also constructed a series of plasmids in which the phaseolin signal-peptide-coding region was either removed or replaced with increasing amounts of the amino-terminal coding region for acid phosphatase. Phaseolin polypeptides with no signal peptide were not posttranslationally modified in *S. cerevisiae*. Partial or complete substitution of the phaseolin signal peptide with that from acid phosphatase dramatically inhibited both signal peptide processing and glycosylation, suggesting that some specific feature of the phaseolin signal amino acid sequence was required for these modifications to occur. Larger hybrid proteins that included approximately one-half of the acid phosphatase sequence linked to the amino terminus of the mature phaseolin polypeptide did undergo proteolytic processing and glycosylation. However, these polypeptides were cleaved at several sites that are not normally used in the unaltered acid phosphatase protein.

A number of foreign genes have been expressed in the yeast *Saccharomyces cerevisiae*. One of the major advantages of this sytem for the expression of heterologous eucaryotic proteins is the presence of well-characterized pathways for the posttranslational alteration and transport of proteins (26), offering the potential for correct modification and for intra- or extracellular targeting of foreign-gene products. Several plant proteins have been shown to undergo signal peptide cleavage, glycosylation, or both when synthesized in *S. cerevisiae* (10, 11, 24). However, the nature of the plant sequences required for correct protein modification in the yeast system has not been examined in detail.

We have shown that phaseolin, the major seed storage protein from French bean plants (*Phaseolus vulgaris* cv. Tendergreen), is expressed in *S. cerevisiae* (10). The term phaseolin applies to a group of closely related peptides that are heterogeneous in molecular mass and isoelectric charge. On one-dimensional sodium dodecyl sulfate (SDS)-polyacrylamide gels, the proteins separate into a 45.5-kilodalton (kDa) nonglycosylated form and two glycosylated components of 48 and 51 kDa. These polypeptides are encoded by six to eight closely related genes (30). DNA sequence data (13, 27, 28) have indicated that the genes can be classified into two groups, α and β, each containing a limited number of base substitutions and deletions. β Phaseolin contains 421 amino acids, including the initial methionine, and α phaseolin is 14 amino acids larger. Each type of protein contains two N-glycosylation recognition sites and a hydrophobic amino-terminal signal sequence (see Fig. 2). N-terminal analysis of phaseolin polypeptides purified from bean cotyledons has shown that heterogeneous processing of the signal peptide occurs. Threonine, serine, and, to a minor extent, leucine are found at the amino-terminal ends of the mature polypeptides (H. E. Paaren, R. J. Blagrove, A. S. Inglis, J. L. Slightom, and T. C. Hall, Phytochemistry, in press; see Fig. 2). In *S. cerevisiae*, both α and β polypeptides are glycosylated, and the molecular weights of the products indicate that the phaseolin signal peptides have been removed (10). In earlier studies, cDNA copies of the genes were expressed after being spliced into the normal genomic phaseolin 5′- and 3′-flanking sequences, and the level of protein observed was quite low, only 0.01 to 0.03% of the total yeast protein. To enhance the level of protein expression and to examine the features of the phaseolin polypeptides which allow their posttranslational modification in *S. cerevisiae*, we linked the coding region of a β-type phaseolin to the promoter and regulatory region of the *S. cerevisiae* repressible acid phosphatase gene (*PHO5*) in multicopy expression plasmids (18, 25).

Vectors containing the *PHO5* promoter provided several desirable characteristics for our studies. Transcription of the *PHO5* gene is repressed when inorganic phosphate is present in the medium but is induced to high levels in the absence of phosphate (1, 5, 17); therefore, foreign genes linked to the promoter can be expressed or not, depending on the culture conditions of the transformants. Yeast acid phosphatase is a secreted glycoprotein found in the periplasmic space (21). Sequence analysis of the *PHO5* gene and the mature protein has revealed a 17-amino-acid signal sequence (2, 31), which is apparently required for processing and secretion of the protein (12), and 11 recognition sites for asparagine-linked glycosylation (2). Thus, acid phosphatase undergoes posttranslational processing similar to that of phaseolin. In this paper we examine the effect of yeast or plant amino acid sequences in determining the nature of these events when hybrid proteins are synthesized.

MATERIALS AND METHODS

Recombinant-plasmid construction. Plasmids used in forming *PHO5*-phaseolin expression plasmids are diagramed in

* Corresponding author.
† Agrigenetics Advanced Science Co. publication no. 62.

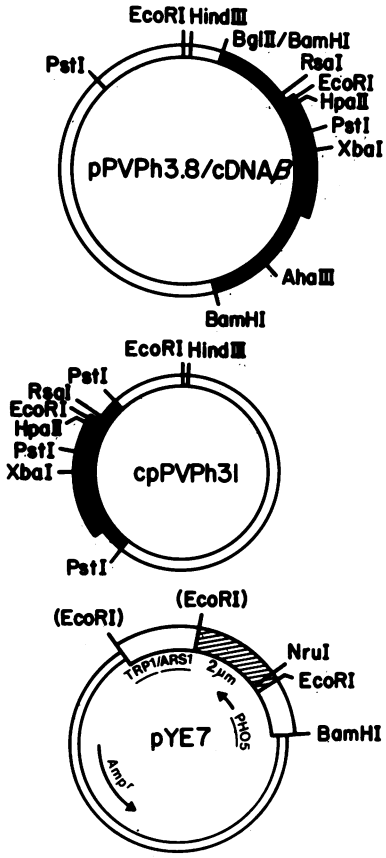

FIG. 1. Parental plasmids used in the construction of *PHO5*-phaseolin expression vectors. pPvPh3.8/cDNAβ and cpPvPh31 contain the phaseolin DNA used in expression plasmid constructions, and pYE7 is the yeast vector. The positions of relevant restriction endonuclease sites are noted. Only those *Rsa*I, *Hpa*II, and *Aha*III sites used in construction of the fusion plasmids are shown; these sites are not unique in the phaseolin DNA. Narrow open bars are pBR322; shaded wide and narrow bars are phaseolin-coding and -noncoding regions, respectively; open wide bars are yeast sequences; hatched regions represent yeast 2μm DNA.

Fig. 1, and details of the plasmid constructions are described in Results. Plasmid cpPvPh31 (13) is pBR322 containing a complete β-phaseolin cDNA sequence. Plasmid pPvPh3.8/cDNAβ (10) contains the cDNA coding region from cpPvPh31 spliced into the flanking sequences from an identical genomic β-phaseolin gene. The *S. cerevisiae* expression vector pYE7 (25) is composed of part of pBR322, the yeast 1.4-kilobase *Eco*RI fragment carrying *TRP1/ARS1*, the yeast 2μm plasmid 2.2-kilobase *Eco*RI fragment containing the plasmid replication origin, and a *PHO5* promoter fragment that extends from a *Bam*HI site 550 base pairs (bp) upstream to a point 9 bp upstream of the acid phosphatase ATG (18). *Eco*RI restriction sites flanking the *TRP1/ARS1* fragment were removed to facilitate cloning at the *Eco*RI site adjacent to the *PHO5* promoter. The *Bam*HI-*Rsa*I, *Bam*HI-*Hae*III, and *Bam*HI-*Aha*III *PHO5* promoter-regulatory restriction fragments used in some constructions were from the plasmid pAP20 (31). Restriction enzymes, T4 ligase, and *Escherichia coli* DNA polymerase (Klenow fragment) were used as recommended by their suppliers.

Yeast strains and culture conditions. *S. cerevisiae* W301-18A (*MATα ade2-1 trp1 can1-100 his3-11* or *his3-15 leu2-3* or *leu2-112 ura3-1*) (23) and 20B-12 (*MATα trp1 pep4-3*) (16) were used as yeast host strains in experiments designed to study the induction of phaseolin under control of the *PHO5* promoter. *S. cerevisiae* 29A21 (*MATα pho80 trp1 ade2 leu2*) (18) carries a mutation in the *PHO80* gene, a repressor of acid phosphatase synthesis, and expresses acid phosphatase constitutively. Yeast transformations were performed as previously described (14). High-phosphate yeast nitrogen base plus Casamino Acids medium, supplemented with uracil and adenine (18), and phosphate-free UMD medium (25) both lack tryptophan and allow for selection of yeast transformants containing pYE7 or its derivatives.

RNA isolation and analysis. Transformed yeast cells were grown in high-phosphate selective medium to a cell density of approximately 1.2×10^7 cells per ml (optical density at 600 nm, 2), harvested, washed with sterile H_2O, and suspended in high-phosphate or phosphate-free medium at an optical density at 600 nm of 0.5. After a 6-h incubation, total cellular RNA was extracted from the cultures (4), fractionated in 1.4% agarose gels after denaturation with glyoxal (22), transferred to nitrocellulose (32), and hybridized to a ^{32}P-labeled restriction fragment encompassing the phaseolin cDNA coding region.

Protein extraction. *S. cerevisiae* W301-18A and 20B-12 transformants were grown in high-phosphate or phosphate-free medium as described above for RNA preparations. *S. cerevisiae* 29A21 transformants were grown in high-phosphate medium to the stationary phase. Tunicamycin treatment was at 5 μg/ml as previously described (10). For strains W301-18A and 20B-12, the drug was added at the time of cell suspension in phosphate-free medium. Cells were harvested and disrupted by vortex mixing with glass beads in a buffer containing 10 mM NaH_2PO_4 (pH 7.5), 1 mM phenylmethylsulfonyl fluoride, and 0.1% Triton X-100. Extracts were clarified by centrifugation, and the total soluble protein was quantitated by the Bradford protein assay (6).

Immunoblot analysis and phaseolin quantitation. Protein extracts were fractionated on 13% SDS-polyacrylamide gels (acrylamide/bisacrylamide ratio, 30:0.15) (19) and then transferred to nitrocellulose by electroelution (33). Filters were incubated with polyclonal rabbit phaseolin antiserum in the presence of TBS (20 mM Tris hydrochloride [pH 7.5], 500 mM NaCl) plus 1% gelatin, washed extensively with TBS, and then incubated with ^{125}I-labeled *Staphylococcus aureus* protein A (New England Nuclear Corp.) in TBS plus 1% gelatin. After final washing with TBS to remove unbound radioactivity, antigen-antibody complexes were visualized by autoradiography (7).

Glycosylated polypeptides were identified by their ability to bind to concanavalin A (ConA)-Sepharose 4B (Sigma Chemical Co.). Yeast protein extracts were incubated with ConA-Sepharose at 20°C for 1.5 h in 20 mM Tris hydrochloride (pH 7.5)–500 mM NaCl–1 mM $MgCl_2$–1% Triton X-100, and the supernatant was reserved. The ConA-Sepharose fraction was washed exhaustively with 20 mM Tris hydrochloride (pH 7.5)–500 mM NaCl–0.05% SDS, suspended in 1% SDS, and boiled to elute glycoproteins bound to the ConA. This fraction and the reserved supernatant were both diluted with 10 mM NaH_2PO_4 (pH 7.5)–0.1% Triton X-100, and the proteins were concentrated by precipitation with 80% acetone before electrophoresis.

Quantitation of phaseolin in soluble yeast protein extracts was performed by dot-immunobinding assays. Cellular protein extracts were diluted in phosphate-buffered saline (6 mM KH_2PO_4 [pH 6.8], 125 mM NaCl) to yield concentrations of 1 to 5 μg of protein per 100 μl, and the samples were bound to nitrocellulose filters by gentle suction through a

Minifold apparatus (Schleicher & Schuell, Inc.). Purified phaseolin protein for standards was quantitated by the Bradford assay. Duplicate standard samples over a range of 1 to 20 ng were prepared in 100 μl of phosphate-buffered saline containing 5 μg of bovine serum albumin and were included in the samples bound to each filter. Filters were incubated with antiphaseolin antibody and ^{125}I-labeled *S. aureus* protein A as described above. Sections containing the radioactive dots were cut out and assayed in a scintillation counter. Determinations of the amounts of phaseolin in cellular extracts were based on the standard curve, which was linear over the range examined.

RESULTS

Construction of *PHO5*-phaseolin expression plasmids. In previous studies of phaseolin expression in *S. cerevisiae* (10), we replaced the protein-coding region of a β-phaseolin genomic restriction fragment with an identical cDNA copy of the gene to produce a sequence equivalent to that in the bean genome except for the absence of intervening sequences. The plasmid containing this genomic-cDNA sequence, pPvPh3.8/cDNAβ, and a β-phaseolin cDNA-containing plasmid, cpPvPh31 (Fig. 1), were used as the sources of phaseolin DNA in constructing the *PHO5*-phaseolin expression plasmids. The yeast expression vector pYE7 is a multicopy extrachromosomal plasmid that contains the *PHO5* promoter and regulatory signals, the *TRP1* gene for selection and maintenance in auxotrophic host cells, and the yeast 2μm plasmid origin of replication (25). Features of the *PHO5* and phaseolin genes relevant to the plasmid constructions are illustrated in Fig. 2.

We linked the phaseolin protein-coding sequence to the *PHO5* promoter-regulatory region either alone or with increasing amounts of the amino-terminal coding sequence for the acid phosphatase enzyme, as shown in Fig. 3. The plasmids in the pYE7Ph series contained phaseolin DNA from pPvPh3.8/cDNAβ which extended through the coding region to an *Aha*III site in 3′ nontranscribed DNA. This site was ligated to the unique *Nru*I site in the 2μm DNA region of pYE7. The pYE8Ph series contained less phaseolin 3′-flanking DNA. For these plasmids, the *Pst*I site adjacent to the cDNA poly(A) sequence in cpPVPh31 was converted to an *Eco*RI site via the insertion of a synthetic linker so that it could be ligated directly into the *Eco*RI site of pYE7.

To make pYE7Ph, the *Eco*RI 5′ extension in pYE7 was filled in with the Klenow fragment and ligated to the *Rsa*I site in the 5′ noncoding region of phaseolin DNA from pPVPh3.8/cDNAβ. Thus, pYE7Ph (Fig. 3) contained the intact *PHO5* promoter, 57 bp of phaseolin DNA 5′ to the translation initiation codon, and the entire phaseolin polypeptide-coding sequence. The short phaseolin 5′-flanking region does not include the TATA and CCAAT sequences thought to be required for transcription of this gene in bean tissue (28). The remaining *PHO5*-phaseolin plasmids had no phaseolin 5′-flanking sequences. pYE72Ph and pYE82Ph contained the *Bam*HI-AhaIII *PHO5* restriction fragment, which included the sequence for the first two acid phosphatase amino acids, fused to the sequence for amino acid 28 of phaseolin. For this construction, the *Hpa*II site in the phaseolin DNA was filled in with the Klenow fragment to create a blunt end. These two plasmids encoded a protein which contained no signal peptide and was equivalent to the mature phaseolin polypeptide except for the first two N-terminal amino acids. pYE83Ph included the sequence for 15 amino acids of acid phosphatase on a *Bam*HI-*Hae*III *PHO5* fragment connected to the sequence for amino acid 22 of phaseolin. The codons for amino acids 22 through 28 were added by the insertion of a synthetic oligonucleotide linker extending from the *PHO5 Hae*III site to the phaseolin *Hpa*II site. Of these codons, five had been altered (see Fig. 3) to conform to more-favored usage in *S. cerevisiae* (3). The protein encoded by pYE83Ph had a hybrid signal peptide and the signal peptide cleavage site normally found in phaseolin. In pYE71Ph and pYE81Ph, the sequence for the first 18 amino acids of acid phosphatase on a *Bam*HI-*Rsa*I restriction fragment was fused to that of amino acid 28 of phaseolin through the filled-in *Hpa*II site. These constructions removed the entire coding sequence for the signal peptide of phaseolin and replaced it with that for the signal peptide plus the signal peptidase cleavage site of acid phosphatase. pYE81Ph was modified to form pYE85Ph by inserting an oligonucleotide linker with the coding sequences for the amino acids Thr-Ser-Leu between the codons for the glycine of acid phosphatase and the arginine of phaseolin, thereby regenerating the *PHO5* sequence for the first 19 amino acids, the acid phosphatase signal peptide cleavage site, and the second two phaseolin signal peptidase cleavage sites. The codons for threonine, serine, and leucine were altered as for pYE83Ph to favor usage in *S. cerevisiae*. The final plasmid, pYE84Ph, included the coding region for approximately half of acid phosphatase and extended to an *Sca*I site after amino acid 257. This sequence included five of the N-glycosylation recognition sites as well as the signal peptide and signal peptidase cleavage site of acid phosphatase and encoded a polypeptide of approximately 75 kDa.

Transcription of the phaseolin gene from the *PHO5* promoter. The plasmids described in Fig. 3 were transformed into *S. cerevisiae* and certain ones were examined to determine the nature of the phaseolin transcripts produced under culture conditions which did or did not induce acid phosphatase expression (Fig. 4). Two major polyadenylated (J. H. Cramer, unpublished data) RNA species were evident in yeast transformants containing the different plasmids when the cells were grown in the absence of phosphate (see Fig. 4, minus [−] lanes). If RNA synthesis starts at the sites normally used for *PHO5* transcription initiation (34 and 40 bp upstream from the translation initiation codon [31]), the presence of two transcript size classes might result from alternative processing of the mRNAs or from the use of different transcription termination sites. pYE7Ph transformants (Fig. 4A and B) contained phaseolin transcripts of approximately 1,500 and 1,800 bases, while the two species transcribed from pYE71Ph (Fig. 4C) and pYE72Ph (Fig. 4E) appeared to be slightly smaller, as expected, because the protein-coding sequences were shorter and the phaseolin 5′-flanking sequences were absent. In the 3′ deletion derivatives, pYE81Ph and pYE82Ph, approximately 540 bp of 3′-flanking phaseolin DNA was removed. Transformants containing either of these plasmids also had two phaseolin transcripts (Fig. 4D and F). The smaller, 1,500-base component may or may not have been equivalent to the smaller transcripts in the corresponding parental plasmids. The larger transcript, about 2,100 bases, must have arisen from transcription that extended into the 2μm plasmid sequence, judging from its size.

Transcription of the *PHO5*-phaseolin fusion genes was clearly regulated by the level of inorganic phosphate in the culture medium. For pYE7Ph, which included 57 bp of phaseolin genomic DNA between the *PHO5* promoter and

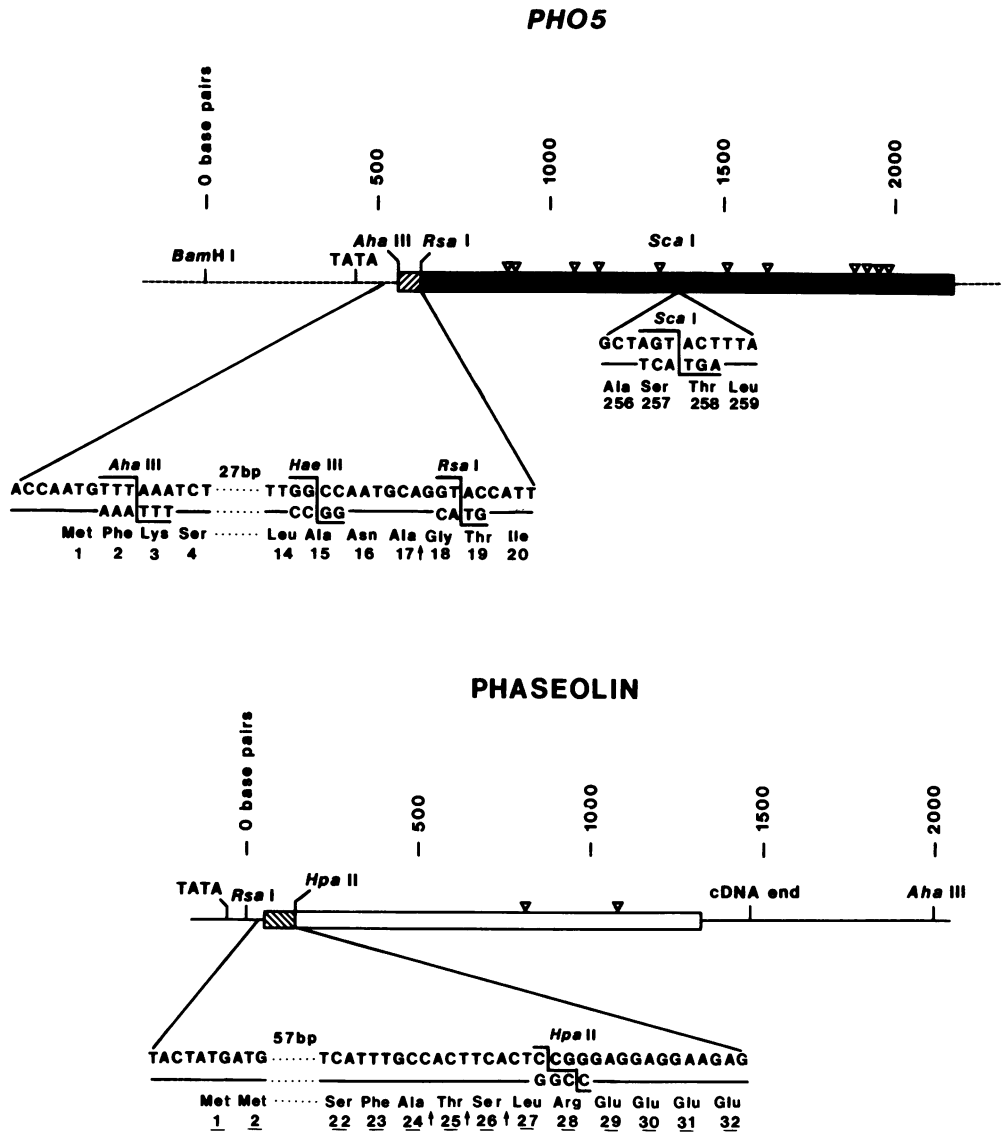

FIG. 2. Organization of the *PHO5* and phaseolin genes. Features relevant to construction of the phaseolin expression plasmids are illustrated. Solid and open bars denote protein-coding sequences of the *PHO5* and phaseolin genes, respectively. Signal peptide regions are indicated by diagonal lines. Horizontal lines represent flanking regions, with *PHO5* regions dashed and phaseolin regions solid. N-glycosylation recognition sites (▽) and the positions of signal peptide cleavage (↑) are indicated for both proteins. Amino acids are numbered from the initiator methionine. The numbers of the phaseolin amino acids are underlined.

the translation initiation codon, the degree of repression in high-phosphate concentrations was not complete and varied among yeast strains (compare Fig. 4A and B; Cramer, unpublished results). Phaseolin transcription in pYE71Ph and pYE72Ph as well as their 3′ deletion derivatives was tightly regulated in both *S. cerevisiae* 20B-12 and W301-18A. Transcription of pYE83Ph and pYE84Ph was also regulated by the phosphate concentration in these strains (data not shown).

We estimated the quantity of phaseolin mRNA produced from the different plasmid constructions under fully induced conditions by comparing transformant total cellular RNAs with serial dilutions of purified phaseolin mRNA on nitrocellulose blots (data not shown). The amount of phaseolin transcript in each case was approximately 2 to 3% of the total yeast mRNA. Notably, the shift in use of transcription termination-polyadenylation sites between plasmids pYE71Ph and pYE81Ph and between plasmids pYE72Ph and pYE82Ph did not markedly affect the amount of phaseolin mRNA in these transformants.

Phaseolin polypeptide synthesis in yeast transformants. Soluble protein extracts were prepared from yeast cultures containing the various recombinant plasmids, and the phaseolin polypeptides were characterized by immunodetection techniques (Fig. 5, Table 1). Phaseolin polypeptides synthesized in transformants containing pYE7Ph (Fig. 5A) were indistinguishable from the three protein species (49, 47, and 44.5 kDa) observed when the protein is expressed from bean genomic sequences in the vector YEp13 (10), although the level of expression from the *PHO5* promoter was much higher. The 49- and 47-kDa polypeptides bound to ConA and were not synthesized in the presence of tunicamycin, an antibiotic that blocks the synthesis of N-linked oligosaccharides (20). The residual amounts of the 47- and 49-kDa

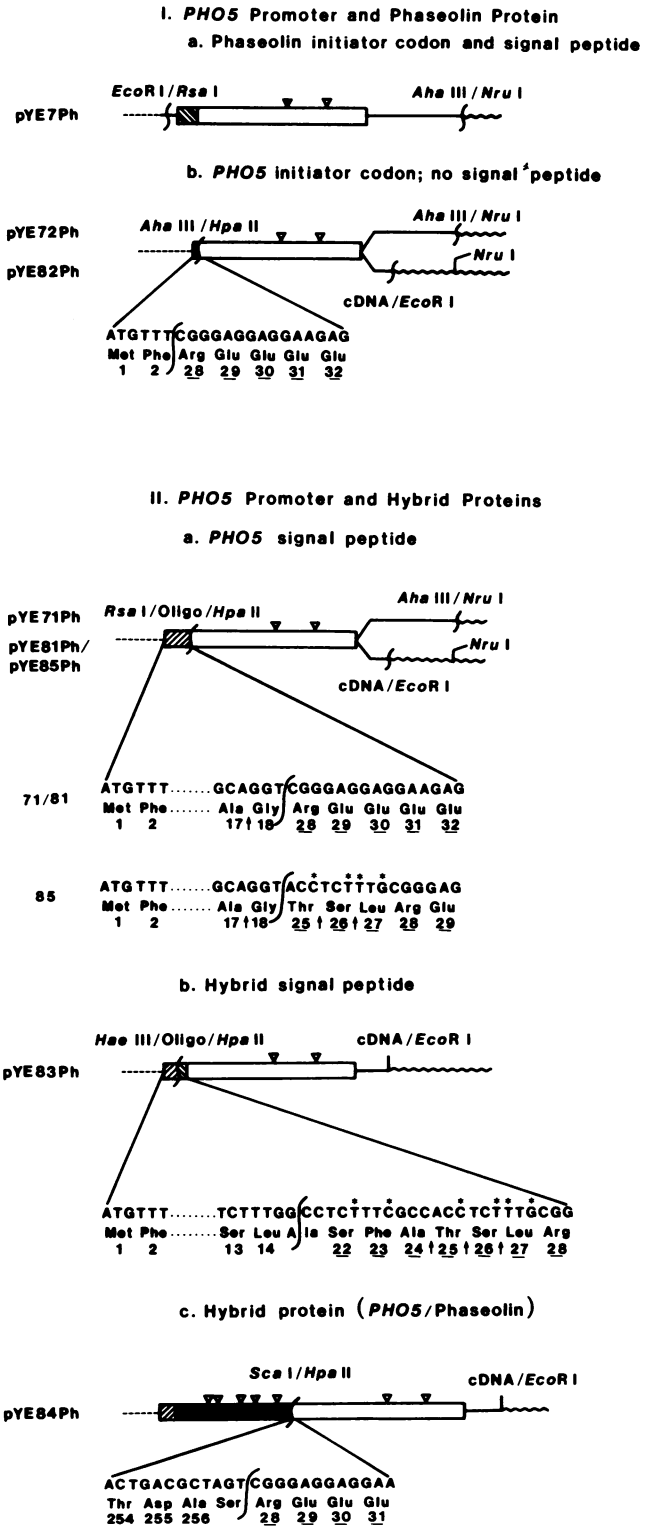

FIG. 3. *PHO5*-phaseolin gene fusions. Relevant details of the gene fusions are shown, and details of plasmid constructions are given in the text. Symbols are the same as described in the legend to Fig. 2 with the addition of wavy lines to denote yeast 2μm DNA and asterisks to indicate those nucleotides which were changed to conform to preferred codons in *S. cerevisiae*.

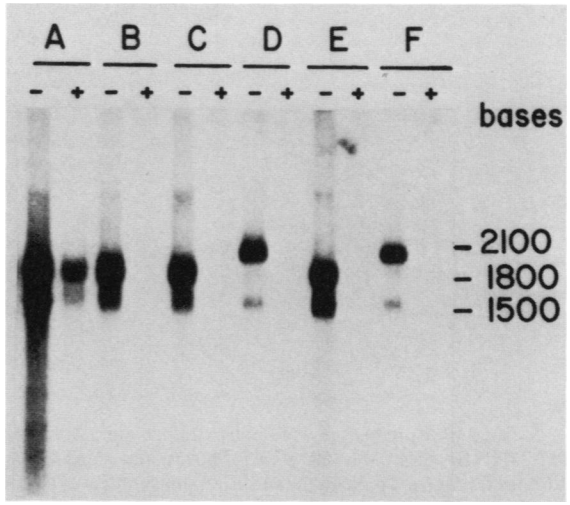

FIG. 4. Transcription of the phaseolin gene in *PHO5* expression plasmids. Total cell RNA was extracted from yeast transformants grown in the presence (+) or absence (−) of inorganic phosphate. Phaseolin transcripts were fractionated and analyzed as described in Materials and Methods. Although equal amounts of total RNA, as determined by A_{260}, were used in all lanes, the amounts of poly(A)$^+$ RNA were not necessarily equivalent in all the samples. Lanes: A, *S. cerevisiae* W301-18A(pYE7Ph); B, *S. cerevisiae* 20B-12(pYE7Ph); C, *S. cerevisiae* 20B-12(pYE71Ph); D, *S. cerevisiae* 20B-12(pYE81Ph); E, *S. cerevisiae* 20B-12(pYE72Ph); F, *S. cerevisiae* 20B-12(pYE82Ph).

components seen in the tunicamycin experiments were presumably polypeptides glycosylated before the addition of the drug. The nature of the low-molecular-size polypeptide in this and other samples from tunicamycin-grown cultures is unknown. The 49- and 47-kDa components represented increasing degrees of glycosylation of the unglycosylated 44.5-kDa polypeptide, which was the size expected for phaseolin after removal of the 24- to 26-amino-acid signal peptide. Molecular sizes of phaseolin polypeptides synthesized in *S. cerevisiae* were not equivalent to those observed for native phaseolin, probably because the extent of glycosylation differs in *S. cerevisiae* and bean plants, and also because these polypeptides represent proteins synthesized from a single β gene rather than the entire phaseolin gene family.

Manipulation of the amino terminus of the phaseolin polypeptide dramatically altered the pattern of posttranslational processing in *S. cerevisiae*. Removal of the signal peptide altogether (pYE72Ph [Fig. 5B]) resulted in the synthesis of a single, unglycosylated peptide close to the predicted molecular size (see Table 1). Synthesis of phaseolin proteins with a hybrid signal peptide, encoded by pYE83Ph, or a complete acid phosphatase signal peptide (pYE71Ph, pYE85Ph) severely impaired processing. For pYE83Ph, only a single, unglycosylated component was synthesized, and its size suggested that the signal peptide had not been cleaved (Fig. 5C). The polypeptide patterns in transformants containing pYE71Ph or pYE85Ph were identical, and the results only for pYE71Ph are shown (Fig. 5D). A limited amount of processing appeared to occur when the complete acid phosphatase signal peptide was present in polypeptides encoded by these plasmids. A 44-kDa unglycosylated component had the molecular size predicted for the protein minus the signal peptide. The 45.5-kDa component contained glycosylated and unglycosylated polypeptides, possibly a mixture of the

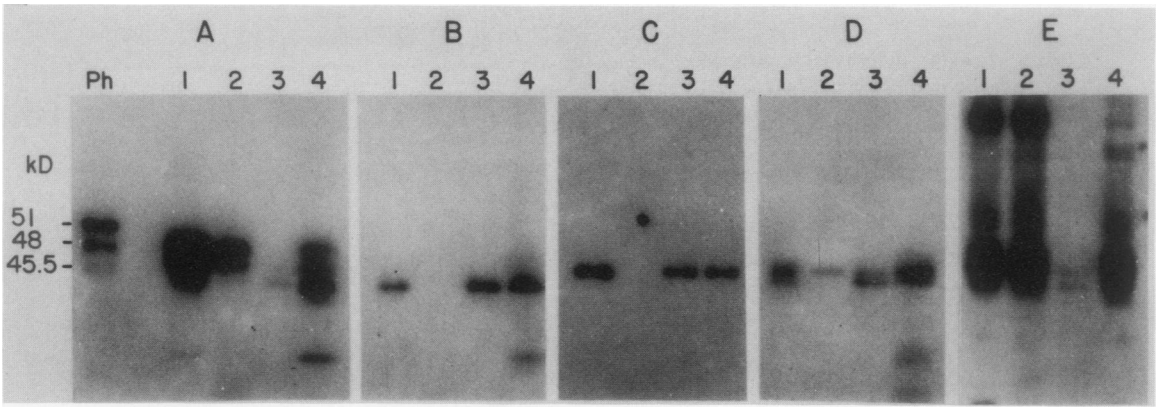

FIG. 5. Acid phosphatase-phaseolin hybrid proteins. Soluble protein extracts were prepared from yeast transformants containing pYE7Ph (A), pYE72Ph (B), pYE83Ph (C), pYE71Ph (D), and pYE84Ph (E). Phaseolin polypeptides were fractionated and characterized as described in Materials and Methods. Each panel shows phaseolin polypeptides as follows: lane 1, in total soluble cellular protein; lane 2, bound to ConA; lane 3, not bound to ConA; and lane 4, from cells grown in tunicamycin. Lane Ph contains purified bean phaseolin polypeptides as molecular size markers. Phosphorylase B (92.5 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), and soybean trypsin inhibitor (21.5 kDa) were also included in the gels and used as molecular size markers (not shown).

entire protein with its signal peptide and the glycosylated protein after signal peptide cleavage.

The pattern of phaseolin polypeptides in transformants containing pYE84Ph (Fig. 5E), which encoded a much-larger fusion protein, was more complex. A 76-kDa protein and several polypeptides of 54 to 58 kDa, all of which were glycosylated, were evident. In addition, a large proportion of the phaseolin migrated at the same positions as those of the three polypeptides synthesized from the intact phaseolin gene in pYE7Ph. Their composition differed, however, in that both the 44.5- and 47-kDa components in pYE84Ph transformants were mixtures of glycosylated and small amounts of unglycosylated polypeptides. Analysis of the tunicamycin sample was complicated by the presence of glycosylated polypeptides synthesized before the addition of the drug. However, distinct new polypeptides were visible at 69 and 71 kDa. Their glycosylated counterparts were probably not seen as discrete species in the absence of tunicamycin because of various degrees of glycosylation at the sites in the acid phosphatase portion of the polypeptide. A 42-kDa polypeptide may represent the precursor to the glycosylated component of the 44.5-kDa band.

We also quantitated the total amount of phaseolin-immunoreactive material in soluble protein extracts from transformants containing the different plasmids (Table 1). Values were determined for cultures grown under fully induced conditions for *PHO5* expression and were similar for *S. cerevisiae* W301-18A, 20B-12, and 29A-21, the latter being a strain constitutive for *PHO5* expression. Percentages in Table 1 represent at least three independent determinations and are presented as ranges. Changes in the pattern of posttranslational processing did not appear to alter, in a consistent and significant way, the level of total phaseolin expression.

TABLE 1. Characteristics of acid phosphatase-phaseolin hybrid polypeptides

Plasmid(s)	Polypeptide molecular sizes (kDa)[a]	Glycosylation	% Total yeast protein[b]
pYE7Ph	49	+	1.3–2.5
	47	+	
	44.5	−	
pYE72Ph, pYE82Ph	44	−	2.6–3.4
pYE83Ph	46	−	2.3–2.7
pYE71Ph, pYE81Ph, pYE85Ph	45.5	+ and −	1.5–3.3
	44	−	
pYE84Ph	76	+	1.6–1.8
	71, 69	−	
	54–56	+	
	49	+	
	47	+	
	44.5	+ and −	

[a] Polypeptide molecular sizes were determined from SDS-polyacrylamide gels.

[b] Protein quantity represents the sum of individual phaseolin polypeptide components and was determined by the immunodot method described in Materials and Methods.

DISCUSSION

In developing bean cotyledons, phaseolin polypeptides undergo several processing steps while en route from their site of synthesis on the rough endoplasmic reticulum to their deposition in membrane-delimited storage vacuoles termed protein bodies. Cleavage of the hydrophobic signal peptide and core glycosylation occur cotranslationally, while further alterations in the carbohydrate moiety take place during subsequent steps of the transfer process (9). We showed previously that phaseolin undergoes similar types of post-translational processing when expressed at low levels in yeast cells (10). When the level of phaseolin expression was enhanced by placing the gene under the control of the yeast *PHO5* promoter, similar results were obtained, indicating that the cellular capacity to process this foreign protein was not saturated at least at levels up to 3% of the total soluble protein.

There was a high degree of specificity in the requirement for the phaseolin signal peptide sequence to direct posttranslational modification of this protein in *S. cerevisiae*. As expected, removal of the signal sequence prevented processing, presumably because the protein could not bind to a

signal recognition particle and subsequently be transported across the endoplasmic reticulum to the location of the signal peptidase and glycosylating enzymes. Furthermore, the phaseolin signal peptide could not be effectively replaced by that from yeast acid phosphatase, an enzyme that normally travels through the secretory pathway. The failure to process hybrid proteins was apparently not the result of overloading the normal acid phosphatase processing pathway. Wild-type enzyme is correctly processed at expression levels as high as 15 times normal (12).

Polypeptides encoded by pYE71Ph (and pYE81Ph) underwent limited signal peptide cleavage and only a single glycosylation event. Although these proteins had the complete signal peptide and cleavage recognition site of acid phosphatase, the environment of the site had been altered. Other researchers (12) have also found that a change in the amino acid sequence of the mature protein adjacent to the signal peptide recognition site of acid phosphatase reduces the efficiency of cleavage. However, in that case glycosylation of the protein appeared to be unaffected. Although the acid phosphatase cleavage site is normally surrounded by hydrophobic amino acids, the hybrid protein encoded by pYE71Ph brought the charged amino acid sequence Arg-Glu-Glu-Glu-Glu from phaseolin into close proximity. Moving these amino acids to their normal distance from the cleavage site in phaseolin by insertion of three uncharged amino acids in the protein encoded by pYE85Ph did not improve the efficiency of posttranslational processing. Creation of a phaseolin polypeptide with a hybrid signal sequence and a phaseolin signal peptide cleavage site resulted in loss of posttranslational processing altogether.

We observed a different type of result when the products synthesized from the pYE84Ph construction were examined. The protein encoded by this plasmid included over half of the acid phosphatase amino acid sequence as well as the sequence for the mature phaseolin polypeptide. This hybrid protein was glycosylated in yeast cells and underwent a number of proteolytic processing steps, finally yielding species that appeared very similar to the polypeptides synthesized from the normal phaseolin gene. When such a large portion of the mature acid phosphatase protein was present in association with the phaseolin mature polypeptide, the phosphatase signal peptide was apparently effective in directing the protein to the site of the processing and glycosylating enzymes. In addition, some feature of the hybrid protein appeared to allow proteolytic cleavage at numerous sites in the acid phosphatase amino acid sequence that were not used in the wild-type form of the enzyme.

Because we detected the pYE84Ph proteins with antibody to phaseolin, we did not follow the fate of the acid phosphatase portion during the processing steps. Although we cannot rule out some processing at the C terminus, we assume that proteolytic processing occurs predominantly at the amino-terminal ends of the hybrid proteins. The glycosylation pattern of the polypeptides was consistent with this hypothesis. The molecular size of the glycosylated 76-kDa polypeptide and its discrete nature suggest that it underwent only limited glycosylation, presumably during translation. The 71- and 69-kDa polypeptides were observed only as unglycosylated components in the presence of tunicamycin, possibly because various degrees of glycosylation at the five acid phosphatase and two phaseolin recognition sites in the absence of the drug caused them to migrate as smears on polyacrylamide gels. If the acid phosphatase glycosylation sites were removed by N-terminal processing, there would be less possibility for molecular size variation as was observed for the 56-kDa and smaller polypeptides, which could be seen as discrete bands.

It is not clear whether the failure of the acid phosphatase signal peptide to direct posttranslation processing in the smaller hybrid proteins examined here was a function of the acid phosphatase sequence itself, yeast signal sequences as a whole, or some special feature of the phaseolin polypeptide. Although the acid phosphatase signal has not been used previously for this purpose, the signal from a different yeast secretory protein, invertase, directs the processing and secretion of several different animal polypeptides despite considerable differences among the amino acid sequences of the mature proteins (8, 29). An analysis of the local average hydrophobicity (15) of the amino acids in the signal sequences of acid phosphatase and phaseolin does reveal some differences. Signal sequences from a variety of sources include a charged amino acid residue in the first five amino acids (34). Phaseolin contains two arginine residues in this region, creating a short hydrophilic sequence, whereas acid phosphatase has a single lysine and the N terminus of the signal is hydrophobic in nature. It may be that the relationship between the phaseolin signal and the mature protein is designed to require this specific region. Clearly, further analysis is necessary to characterize the specific requirements for processing in this system and to develop a more general understanding of the requirements for posttranslational modification of hybrid proteins as a whole.

LITERATURE CITED

1. **Anderson, N., G. P. Thill, and R. A. Kramer.** 1983. RNA and homology mapping of two DNA fragments with repressible acid phosphatase genes from *Saccharomyces cerevisiae*. Mol. Cell. Biol. **3**:562–569.
2. **Arima, K., T. Oshima, I. Kubota, N. Nakamura, T. Mizunaga, and A. Toh-e.** 1983. The nucleotide sequence of the yeast *PHO5* gene: a putative precursor of repressible acid phosphatase contains a signal peptide. Nucleic Acids Res. **11**:1657–1672.
3. **Bennetzen, J. L., and B. D. Hall.** 1982. Codon selection in yeast. J. Biol. Chem. **257**:3026–3031.
4. **Bostian, K. A., J. E. Hopper, D. T. Rogers, and D. J. Tipper.** 1980. Translational analysis of the killer-associated virus-like particle dsRNA genome of *S. cerevisiae*: M-dsRNA encodes toxin. Cell **19**:403–414.
5. **Bostian, K. A., J. M. Lemire, L. E. Cannon, and H. O. Halvorson.** 1980. *In vitro* synthesis of repressible yeast acid phosphatase: identification of multiple mRNAs and products. Proc. Natl. Acad. Sci. USA **77**:4504–4508.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. **72**:248–254.
7. **Burnette, W. N.** 1981. "Western blotting." Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112**:195–203.
8. **Chang, C. N., M. Matteucci, L. J. Perry, J. J. Wulf, C. Y. Chen, and R. A. Hitzeman.** 1986. *Saccharomyces cerevisiae* secretes and correctly processes human interferon hybrid proteins containing yeast invertase signal peptides. Mol. Cell. Biol. **6**:1812–1819.
9. **Chrispeels, M. J.** 1984. Biosynthesis, processing and transport of storage proteins and lectins in cotyledons of developing legume seeds. Philos. Trans. R. Soc. London Ser. B **304**:309–322.
10. **Cramer, J. H., K. Lea, and J. L. Slightom.** 1985. Expression of phaseolin cDNA genes in yeast under control of natural plant DNA sequences. Proc. Natl. Acad. Sci. USA **82**:334–338.
11. **Edens, L., I. Bom, A. M. Ledeboer, J. Maat, M. Y. Toonen, C. Visser, and C. T. Verrips.** 1984. Synthesis and processing of the plant protein thaumatin in yeast. Cell **37**:629–633.

12. **Haguenauer-Tsapis, R., and A. Hinnen.** 1984. A deletion that includes the signal peptidase cleavage site impairs processing, glycosylation, and secretion of cell surface yeast acid phosphatase. Mol. Cell. Biol. **4**:2668–2675.
13. **Hall, T. C., J. L. Slightom, D. R. Ersland, M. G. Murray, L. M. Hoffman, M. J. Adang, J. W. S. Brown, Y. Ma, J. A. Matthews, J. H. Cramer, R. F. Barker, D. W. Sutton, and J. D. Kemp.** 1983. Phaseolin: nucleotide sequence explains molecular weight and charge heterogeneity of a small multigene family and also assists vector construction for gene expression in alien tissue, p. 123–142. *In* O. Ciferri and L. S. Dure (ed.), Structure and function of plant genomes. Plenum Publishing Corp., New York.
14. **Hinnen, A., J. B. Hicks, and G. R. Fink.** 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA **75**:1929–1933.
15. **Hopp, T. H., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3824–3828.
16. **Jones, E.** 1976. Proteinase mutants of *Saccharomyces cerevisiae*. Genetics **85**:23–30.
17. **Kramer, R. A., and N. Anderson.** 1980. Isolation of yeast genes with mRNA levels controlled by phosphate concentration. Proc. Natl. Acad. Sci. USA **77**:6541–6545.
18. **Kramer, R. A., T. M. DeChiara, M. D. Schaber, and S. Hilleker.** 1984. Regulated expression of a human interferon gene in yeast: control by phosphate concentration or temperature. Proc. Natl. Acad. Sci. USA **81**:367–370.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Lehle, L., and W. Tanner.** 1976. The specific site of tunicamycin inhibition in the formation of dolichol bound N-acetyl glucosamine derivatives. FEBS Lett. **71**:167–170.
21. **Linnemans, W. A. M., P. Boer, P. F. Elbers.** 1977. Localization of acid phosphatase in *Saccharomyces cerevisiae*—a clue to cell wall formation. J. Bacteriol. **131**:638–644.
22. **McMaster, G. K., and G. G. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. USA **74**:4835–4838.
23. **Rothstein, R.** 1983. One-step gene disruption in yeast. Methods Enzymol. **101**:202–211.
24. **Rothstein, S. J., C. M. Lazaras, W. E. Smith, D. C. Baulcombe, and A. A. Gatenby.** 1984. Secretion of a wheat α-amylase expressed in yeast. Nature (London) **308**:662–665.
25. **Schaber, M. D., T. M. DeChiara, and R. A. Kramer.** 1986. Yeast vectors for production of interferon. Methods Enzymol. **119**:416–423.
26. **Schekman, R.** 1982. The secretory pathway in yeast. Trends Biochem. Sci. **7**:243–246.
27. **Slightom, J. L., R. F. Drong, R. C. Klassy, and L. M. Hoffman.** 1985. Nucleotide sequences from phaseolin cDNA clones: the major storage proteins from *Phaseolus vulgaris* are encoded by two unique gene families. Nucleic Acids Res. **13**:6483–6498.
28. **Slightom, J. L., S. M. Sun, and T. C. Hall.** 1983. Complete nucleotide sequence of a French bean storage protein gene: phaseolin. Proc. Natl. Acad. Sci. USA **80**:1897–1901.
29. **Smith, R. A., M. J. Duncan, and D. T. Moir.** 1985. Heterologous protein secretion from yeast. Science **229**:1219–1224.
30. **Talbot, D. R., M. J. Adang, J. L. Slightom, and T. C. Hall.** 1984. Size and organization of a multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. Mol. Gen. Genet. **198**:42–49
31. **Thill, G. P., R. A. Kramer, K. J. Turner, and K. A. Bostian.** 1983. Comparative analysis of the 5′-end regions of two repressible acid phosphatase genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **3**:570–579.
32. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
33. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
34. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.