Intracellular precursors and secretion of alkaline extracellular protease of Yarrowia lipolytica.

S Matoba, J Fukayama, R A Wing and D M Ogrydziak
Mol. Cell. Biol. **1988, 8(11):4904. DOI: 10.1128/MCB.8.11.4904.**

MOLECULAR AND CELLULAR BIOLOGY, Nov. 1988, p. 4904–4916
0270-7306/88/114904-13$02.00/0

Vol. 8, No. 11

Intracellular Precursors and Secretion of Alkaline Extracellular Protease of *Yarrowia lipolytica*

SAM MATOBA, JUNE FUKAYAMA,† ROD A. WING,‡ AND DAVID M. OGRYDZIAK*

Institute of Marine Resources, University of California, Davis, California 95616

Received 6 June 1988/Accepted 4 August 1988

Processing and secretion of the alkaline extracellular protease (AEP) from the yeast *Yarrowia lipolytica* was studied by pulse-chase and immunoprecipitation experiments. Over half of newly synthesized AEP was secreted by 6 min. Over 99% of AEP activity which was external to the cytoplasmic membrane was located in the supernatant medium. Polypeptides of 55, 52, 44, 36, and 32 kilodaltons (55K, 52K, 44K, 36K, and 32K polypeptides) were immunoprecipitated from [^{3}H]leucine-labeled cell extracts by rabbit antibodies raised against mature, secreted AEP (32K polypeptide). Experiments with tunicamycin and endoglycosidase H indicated that the 55K, 52K, and 44K polypeptides contained about 2 kilodaltons of N-linked oligosaccharide and that the 36K and 32K polypeptides contained none. Results of pulse-chase experiments did not fit a simple precursor-product relationship of 55K→52K→44K→36K→32K. In fact, maximum labeling intensity of the 52K polypeptide occurred later than for the 44K and 36K polypeptides. Secretion of polypeptides of 19 and 20 kilodaltons derived from the proregion of AEP indicated that one major processing pathway was 55K→52K→32K. The gene coding for AEP (*XPR2*) was cloned and sequenced. The sequence and the immunoprecipitation results suggest that AEP is originally synthesized with an additional preproI-proII-proIII amino-terminal region. Processing definitely involves cleavage(s) after pairs of basic amino acids and the addition of one N-linked oligosaccharide. Signal peptidase cleavage, dipeptidyl aminopeptidase cleavages, and at least one additional proteolytic cleavage may also be involved.

Yarrowia lipolytica (previously *Candida*, *Endomycopsis*, or *Saccharomycopsis lipolytica*) secretes high levels (1 to 2% of total cell protein) of an alkaline extracellular protease (AEP) (52). At high cell densities, over 1 g of AEP per liter has been reported (71). We have been using AEP production by *Y. lipolytica* as a model system for the study of protein processing and secretion. This system is amenable to biochemical and genetic analyses. The power of such a combined approach has been demonstrated in *Saccharomyces cerevisiae* in studies of protein translocation and secretion (11, 20, 62, 65), signal sequence functions (34), protein processing (1, 33, 45, 56, 73, 79), endocytosis (12), and vacuolar localization (30, 61, 69, 73).

Y. lipolytica and *S. cerevisiae* are not closely related (46, 68). *Y. lipolytica* exhibits dimorphic growth (59), grows on hydrocarbons (4), and differs in codon bias (18) and in organization of the rDNA genes (14, 24, 74). *Y. lipolytica* also secretes several enzymes in addition to AEP into the extracellular medium (9, 51, 54, 80). *Y. lipolytica* is not as well characterized as *S. cerevisiae*. However, it offers many of the same advantages for studies of protein processing and localization. Almost all the biochemical and genetic manipulations possible with *S. cerevisiae* are also possible with *Y. lipolytica*—tetrad analysis (50), transformation (17, 26), cloning of genes by complementation (19), gene replacement (J. DeZeeuw, personal communication), and expression of heterologous genes (25). Recently, a small ribonucleoprotein homologous to signal recognition particle has been isolated from *Y. lipolytica* and the putative signal recognition particle RNA gene has been cloned and sequenced (55).

Y. lipolytica has been used on a large scale industrially (35), and it has industrial potential as a host for secretion of foreign proteins. The AEP promoter and preproregions have been used to secrete prochymosin from *Y. lipolytica* (European Patent Office application 0220864, E. P. Bulletin May 1987/19). Because (i) large amounts of AEP are secreted, (ii) the AEP preproregion has industrial potential, and (iii) no analogous protein is secreted by *S. cerevisiae*, we decided to study AEP processing in detail.

AEP has been purified and partially characterized (52), the structural gene (*XPR2*) has been identified (66), and the *XPR2* gene from three different strains (including CX161-1B used in this study) has been cloned and sequenced (19; C. Gaillardin, personal communication). AEP is a serine protease with an alkaline pH optimum (52), and based on sequence similarities, it is similar to subtilisin (19).

In this paper, we report on studies of intracellular precursors of AEP detected by immunoprecipitation and on the cloning and sequencing of the gene coding for AEP (*XPR2*). These results are combined to make a model which suggests that AEP processing is quite complex. It involves several precursors, as many as six processing events, and possibly several alternative processing pathways. We also report on the kinetics of AEP secretion and on the location (cell wall versus extracellular medium) of secreted AEP.

MATERIALS AND METHODS

Strains and media. The haploid *Y. lipolytica* CX161-1B *adel* A (ATCC 32338) obtained from J. Bassell and R. Mortimer (University of California, Berkeley) is wild type for AEP production. *Y. lipolytica* PS3065 *adel xpr2-33* A (ATCC 46027) and PS6069 *adel xpr2-34* A (ATCC 46028), which contain mutations in the AEP structural gene, and DO613 *adel xpr6-13* A were isolated by UV mutagenesis of

* Corresponding author.
† Present address: Department of Molecular Biology, Princeton University, Princeton, NJ 08544.
‡ Present address: Plant Gene Expression Center, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA 94710.

CX161-1B (51, 66). Yeast cultures were maintained on YM medium (50) and grown in glycerol-Proteose Peptone (Difco Laboratories, Detroit, Mich.) (GPP) medium (52). Glycerol-casein (GC) medium is GPP with 0.4 g of casein per liter substituted for the Proteose Peptone.

Escherichia coli HB101, K802, and JM101 used for library construction, λ bacteriophage propagation, and DNA sequencing are described by Maniatis et al. (42). *E. coli* media LB and M9 were prepared as described by Miller (44). Recombinant M13 phage were grown on 2× TY medium (1.6% Difco tryptone, 1% yeast extract, 0.5% NaCl).

Materials. ρ-[*methyl*-^{3}H]toluenesulfonyl-L-arginine methyl ester ([^{3}H]TAME; specific activity, 1.3 Ci/mmol) and Protosol and Econofluor were obtained from Dupont, NEN Research Products (Boston, Mass). L-[4,5-^{3}H]leucine (40 to 60 or 120 to 190 Ci/mmol), L-[^{35}S]cysteine (>600 Ci/mmol), L-[^{35}S]methionine (>800 Ci/mmol), L-[G-^{3}H]tryptophan (13.4 Ci/mmol), 5′-[α-^{35}S]dATP (650 Ci/mmol), 5′-[γ-^{32}P] dATP (5,000 Ci/mmol), and PCS liquid scintillation cocktail were obtained from Amersham Corp. (Arlington Heights, Ill.). Phenylmethysulfonyl fluoride (PMSF) was from Calbiochem-Behring (La Jolla, Calif.), and dimethyl sulfoxide was from Mallinckrodt, Inc. (St. Louis, Mo.). Casein (Hammersten) was purchased from ICN Pharmaceuticals Inc. (Irvine, Calif.), and polypropylene glycol (2,000 molecular weight) was from Aldrich Chemical Co., Inc. (Milwaukee, Wis.). Protein A-Sepharose CL-4B, Triton X-100, and 2,5-diphenyloxazole were obtained from Sigma Chemical Co. (St. Louis, Mo.).

Restriction enzymes, T4 DNA ligase, and DNA polymerase I (Klenow fragment) were purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.), and polynucleotide kinase was from Amersham. The enzymes were used according to the specifications of the manufacturer.

[^{3}H]TAME hydrolysis. The assay based on the release of [^{3}H]methanol and its preferential extraction from the aqueous phase by toluene in the scintillation fluid is adapted from that of Roffman et al. (60). Cells were grown in GPP medium containing 50% more glycerol and Proteose Peptone. Samples were collected during exponential growth. Cell-free supernatant samples were obtained by centrifugation for 10 min at 10,000 × *g*. Cells were washed twice in 10 mM sodium azide and suspended at a cell density of 1,140 Klett units in assay buffer consisting of 46 mM Tris hydrochloride buffer (pH 8.0) containing 10 mM sodium azide. The substrate was prepared by diluting the original [^{3}H]TAME solution 100-fold in unlabeled 0.3 M TAME in 46 mM Tris hydrochloride buffer (pH 8.0). For supernatant samples, 5 μl of sample, 15 μl of assay buffer, 10 μl of substrate, and 10 ml of scintillation fluor (15.2 g of POPOP [1,4-bis(5-phenyloxazolyl)benzene] and 0.19 g of PPO per liter of toluene) were combined in a scintillation vial, shaken, and counted with a Beckman 7500 scintillation counter. Samples were incubated at room temperature and periodically shaken and counted. For cell samples, 50 μl of substrate and 100 μl of cell suspension were combined and incubated at room temperature with shaking. At time zero and after 1 h, 30-μl samples were removed and added to 10 ml of scintillation fluor and counted. In all cases, less than 10% of the substrate was hydrolyzed. Controls in which assay buffer was substituted for supernatant medium or cells were run in parallel.

Antiserum preparation. Purified AEP was prepared by ultrafiltration, gel filtration, and ion-exchange chromatography as described previously (52) and was inactivated with 2 mM PMSF. A sodium dodecyl sulfate (SDS)-treated sample containing 100 to 125 μg of AEP was emulsified with Freund complete adjuvant and injected subcutaneously into two or four places on the back of a New Zealand White rabbit. Booster injections of 50 or 100 μg in Freund incomplete adjuvant were done at 2- to 3-week intervals. Antibody titers were monitored by an enzyme-linked immunosorbent assay with the Vectastain ABC kit (Vector Laboratories, Burlingame, Calif.) with 2,2-azino-di-(3-ethyl-benzthiazoline-6-sulfonic acid) (Sigma) as the substrate. Preparation of a crude immunoglobulin fraction by ammonium sulfate precipitation (50% saturation) and Ouchterlony double-diffusion analysis were done as described previously (9).

MAb preparation. Mice were injected with 50 μg of purified AEP diluted 1:1 in Fruend complete adjuvant. The AEP had been inactivated by 1 mM PMSF and denatured in 0.5% SDS. Booster injections of 50 μg of PMSF-inactivated, SDS-denatured AEP in Freund incomplete adjuvant were given after 4, 11, and 16 weeks. Spleen cells from two hyperimmune mice were fused in 50% (wt/vol in saline) polyethylene glycol 1450 (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) with myeloma cell line $P_3x63Ag8$ (53). Cells were grown in RPMI medium containing hypoxanthine, thymidine, and aminopterin, and the supernatant medium was tested for monoclonal antibody (MAb) against AEP by an enzyme-linked immunosorbent assay. A strongly positive clone was subcloned three times and then frozen. Ascites fluid from mice injected with these hybridoma cells was precipitated with ammonium sulfate (50% saturation) to prepare a crude MAb fraction.

Preparation of labeled cell extracts and supernatant media. Cells were grown in GPP medium and suspended in GC medium at 23°C at a cell density of 1,000 Klett units (~4 mg [dry weight] of cells per ml) as described previously (10). After 30 min of incubation, L-[4,5-^{3}H]leucine was added. Typically, 250 μCi of [^{3}H]leucine was added for every 6 ml of suspended cells. Chase with unlabeled leucine was done, when appropriate, by the addition of a 3,000-fold excess of L-leucine. Protein synthesis was stopped by pouring 6 ml of the cell suspension into 2.3 g of crushed ice and 1.0 ml of 100 mM sodium azide containing 25 mM L-leucine. In later experiments, 2 mM PMSF was included in the ice, sodium azide, and L-leucine mixture. Total label incorporated was measured at several time points (to confirm that the chase had been effective) by scintillation counting of trichloroacetic acid (10%, vol/vol) precipitates from two 100-μl samples of the cell suspension.

After the ice had melted, the cell suspension was centrifuged at 8,000 × *g* for 10 min at 4°C. The supernatant sample was precipitated with 10% (vol/vol) trichloroacetic acid on ice for 2 to 3 h, dissolved in Laemmli gel-loading buffer (39) containing 2 mM PMSF, neutralized, and frozen at −20°C as described previously (10). PMSF (2 mM) was added to the cell pellet, the cells were disrupted with glass beads in a Braun homogenizer, and a clarified cell extract prepared as described previously except that 4% Triton X-100 was used and L-leucine was substituted for L-methionine (10).

Protein assays. The Lowry procedure as described by Herbert et al. (28) and the dye-binding procedure described by Bradford (8) were used for protein assays with bovine plasma albumin as a standard.

Immunoprecipitation. The immunoprecipitation procedures were essentially those described by Cheng and Ogrydziak (10). The amounts of antiserum and protein A-Sepharose CL-4B needed to be in excess were determined empirically. For every 6 ml of suspended cells, 30 μl of AEP antiserum was added to the clarified cell extract, and 25 mg (dry weight) of Sepharose beads was used to precipitate the

antigen-antibody complexes. The complex was washed as described previously (10), 300 μl of Laemmli gel-loading buffer (without glycerol and bromphenol blue) was added, and the complex was boiled for 4 min. The mixture was centrifuged, and the supernatant liquid (~250 μl) was removed and frozen at −20°C.

Immunocompetition. Competition experiments with unlabeled PMSF-inactivated purified AEP were done to examine the specificity of the immunoprecipitation reactions. AEP was adjusted to 1% SDS, incubated for 10 min at room temperature, adjusted to 2% Triton X-100, and incubated for 5 min at room temperature. The competition experiments were identical to the immunoprecipitation experiments except that 30 μl of antiserum was first incubated with 500 μg of the treated AEP for 1.5 h at room temperature and overnight at 4°C before it was added to the cell extract.

PAGE and fluorography. Proteins were analyzed by polyacrylamide gel electrophoresis (PAGE) in 10 to 15% linear gradient polyacrylamide slab gels (1.5 mm; 14 by 16 cm) in the presence of SDS by the system described by Laemmli (39). A 3% stacking gel was used, and gels were run at a constant current of 25 mA for 4 to 4.5 h or for a total of 110 mA-h for overnight runs. All samples were boiled for 5 min before being loaded on the gel. For immunoprecipitates, 12.5 μl of sample, equivalent to 0.3 ml of suspended cells, was applied per lane. Proteins precipitated with trichloroacetic acid were dissolved in 500 μl of Laemmli buffer, and 25 μl, equivalent to 0.3 ml of suspended cells, was applied per lane.

After PAGE, gels were stained for 15 to 20 min with Coomassie brilliant blue R250 (0.4% [wt/vol] Coomassie brilliant blue R250, 40% methanol, 5% acetic acid) and destained in 10% methanol–15% acetic acid. The gels were dried after treatment with dimethyl sulfoxide and the scintillator 2,5-diphenyloxazole and exposed to preflashed X-Omat AR X-ray film at −80°C (7). Fluorographs were scanned with a LKB Ultro-Scan laser densitometer.

Gel slices were rehydrated in 50 μl of water and solubilized by shaking overnight at 37°C in 10 ml of an Econofluor-Protosol mixture (95:5) in glass vials. Counting was done on several consecutive days until the numbers of counts stabilized.

Molecular weight estimation. Molecular weights were estimated by SDS-PAGE. Values were obtained from linear regression analysis of $\log_{10}$ molecular weight versus the $\log_{10}$ of the polyacrylamide concentration (%T) to which the protein migrated (27).

Electroelution. Labeled AEP-related polypeptides were resolved by SDS-PAGE and located by fluorography. The labeled bands were cut out of the gel and cut into small pieces. Electroelution was done with an electrophoretic concentrator at 4°C in a pH 8.3 buffer containing 14 g of glycine per liter, 3 g of Trizma base per liter, 2 mM PMSF, and 0.05% SDS. The dialysis membranes and chambers were first rinsed with 100 μl of a 2-mg/ml solution of bovine serum albumin (type IV; Sigma) to minimize losses of labeled material caused by adsorption.

Tunicamycin inhibition. For tunicamycin inhibition, cells were prepared and treated as in a normal immunoprecipitation experiment except that tunicamycin (Sigma) was included at a final concentration of 10 μg/ml in the GC medium.

Endo H. Endoglycosidase H (endo H; Miles Scientific, Div. Miles Laboratories, Inc., Naperville, Ill.) was used as recommended by the supplier. Radioactive bands were cut out of a gel and electroeluted. About 3,000 counts (5 to 10 μl) were diluted 10-fold in 50 mM sodium citrate buffer (pH 5.5) containing 5 mM sodium azide. PMSF (2 mM, final concentration) and 5.0 mU of endo H were added to the sample. The sample was incubated at 37°C for 12 h, an additional 5.0 mU of endo H was added, and the incubation continued for 12 h. Laemmli buffer (4×) was added, the mixture was boiled for 5 min, and the total volume was applied to a gel. Invertase (grade VII; Sigma) from *S. cerevisiae*, purified as described by Trimble and Maley (72), served as a positive control for endo H treatment.

Peptide mapping. Peptide mapping was done by the method of Cleveland et al. (15). Labeled AEP-related polypeptides isolated from cell extracts by immunoprecipitation were resolved by SDS-PAGE and located by fluorography. Gel slices cut from the first gel were inserted into the well of a 17.5% polyacrylamide gel and digested with *Staphylococcus aureus* V8 protease (Miles Scientific) in the stacking gel as described by Cleveland et al. (15). Concentrations and incubation conditions are listed in the figure legends. SDS-PAGE and fluorography were done as described above.

Immunoblotting. The procedures for immunoblotting were essentially as described by Cheng and Ogrydziak (9) with the following modifications. The sample used was an extract prepared from cells suspended at 1,000 Klett units for 30 min in GC medium as in an immunoprecipitation experiment. Material obtained from 1.0 ml of suspended cells was applied per lane of a 10 to 15% polyacrylamide gel. The blots were incubated first with a 1:50 dilution of mouse monoclonal AEP antiserum in 1% bovine serum albumin and then with a 1:125 dilution of horseradish peroxidase-conjugated rabbit anti-mouse Immunoglobulin G (IgG) (Accurate Chemical and Scientific Corp., Westbury, N.Y.) in 1% bovine serum albumin.

λ Charon 4 library construction and screening. *Y. lipolytica* CX161-1B DNA was isolated from protoplasts (67) essentially as described by Cryer et al. (16). The DNA was partially digested with *Hae*III and *Alu*I, and 15- to 20-kilobase fragments were recovered from a 10 to 40% sucrose gradient. The DNA was methylated with *Eco*RI methylase, and *Eco*RI linkers were added.

The λ Charon DNA was prepared as described by Zehnbauer and Blattner (82). The *Y. lipolytica* DNA was ligated to the λ Charon 4 arms and packaged in vitro, and its titer was determined. The library contained 65,000 phage, and 89% were recombinant. The library was amplified once in *E. coli* K802. The library was screened by the Benton-Davis in situ plaque hybridization procedure as modified by Maniatis et al. (42).

Oligonucleotide hybridization. Oligonucleotide probes designed to hybridize to the *XPR2* gene were kindly synthesized by B. Dujon using the solid-phase phosphotriester method (21). The probes, 51 and 44 bases in length, were synthesized based on the N-terminal amino acid sequence of mature AEP utilizing the codon bias of the *Y. lipolytica LEU2* gene (25). The probes were 5′-AAGCGAGCCAT TCAGACCACTCCCGTCACTCAGTGGGGCCTCTC-3′ and 3′-GTCACCCCTGAGAGAACATAGAGATGGTTCTT CCGGGTCTGACCGTTGATG-5′. The probes were purified on polyacrylamide-urea gels and further purified as described by Atkinson and Smith (2). The probe was then phosphorylated with [γ^{32}P]ATP and polynucleotide kinase and separated from unincorporated label on a Bio-Gel P60 column.

Hybridization conditions were established for the "guessmers" by probing Southern blots of restriction enzyme digests of *Y. lipolytica* genomic DNA under varied formamide concentrations (10 to 30%). The DNA was transferred to

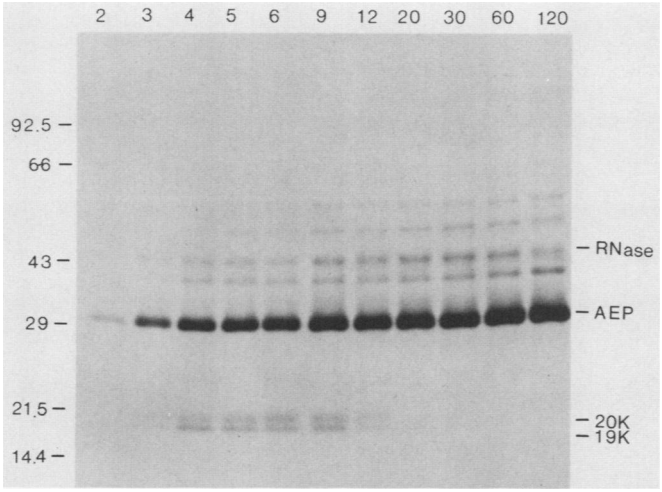

FIG. 1. Time course of secretion of AEP. *Y. lipolytica* cells were labeled for 40 s with 250 μCi of [^{3}H]leucine, and the radioactivity was chased with unlabeled leucine for the time periods (minutes) indicated in the figure. The culture supernatant medium was precipitated with trichloroacetic acid, and the precipitates were analyzed by SDS-PAGE on a 10 to 15% gel and fluorography. Left lane, Position and apparent molecular masses (in kilodaltons) of unlabeled molecular size standards (Bio-Rad) are indicated. Marker proteins were phosphorylase *b* (92,500), bovine serum albumin (66,000), ovalbumin (43,000), carbonic anhydrase (29,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400). Far right lane, Positions of RNase, mature AEP, and the 19K and 20K polypeptides are indicated. Total trichloroacetic acid-precipitable radioactive counts incorporated into 6 ml of cells and supernatant medium were 43,400,000. The equivalent of 40 μl of supernatant medium was applied per lane. Film exposure was for 4 days.

Zeta Probe nylon membrane (Bio-Rad Laboratories, Richmond, Calif.) by the rapid alkaline blotting technique described by Reed and Mann (58). The blots were hybridized in 4× SSPE–1% SDS–0.5% (wt/vol) evaporated milk–1 mg of heat-denatured sonicated salmon sperm DNA per ml–formamide for 12 h at 37°C. Prehybridization had been done for several hours in similar solutions lacking formamide. After hybridization, the filters were washed twice for 30 min at 23°C with 4× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and once for 30 min at 45°C with 0.2% SSC–0.1% sodium PP_i. The filters were blotted dry and autoradiographed with intensifying screens at −70°C.

DNA sequencing. DNA sequencing was done by the dideoxy method of Sanger et al. (63) with M13mp18 and M13mp19 (81). Both [α-^{35}S]dATP (6) and wedge gels were used. A total of 2,974 nucleotides, which includes the entire *XPR2* coding region, were sequenced from both strands.

RESULTS

AEP secretion. Pulse-chase experiments with [^{3}H]leucine were used to examine AEP secretion. Supernatant samples were taken at various times, and AEP was inhibited by the addition of PMSF. The samples were precipitated with trichloroacetic acid and examined by SDS-PAGE and fluorography (Fig. 1). The heavily labeled band is AEP, and AEP is clearly the major protein secreted under these conditions. The bands were cut out, and and their radioactivities were counted. At 9 min, 70% of the label was in the AEP band, at 20 min, 76% was in the AEP band, and at 60 min, 79% was in the AEP band. AEP and the larger polypeptides were relatively stable, and the counts in the bands increased slightly with time. The 19- and 20-kilodalton polypeptides (19K and 20K polypeptides) were quite unstable.

Western blots (immunoblots) of supernatant media from cultures handled in the same manner as in the pulse-chase experiments were prepared. When blots were probed with rabbit polyclonal antibody raised against purified extracellular RNase, one band of about 45,000 daltons reacted as expected (9) (data not shown). Therefore, the second band above AEP in Fig. 1 is most likely the secreted RNase. When the blots were probed with rabbit polyclonal antibody or mouse MAb raised against purified AEP, only one band corresponding to mature AEP was detected (data not shown). Considering the sensitivity of the Western blots for AEP-related proteins, the other labeled bands are probably not AEP precursors or degradation products of mature AEP.

A comparison of the counts incorporated into the AEP band with the total counts incorporated into the cells and supernatant medium yields the estimate that AEP synthesis accounts for 6.9% of total protein synthesis under these labeling conditions.

Based on these results and results of similar pulse-chase experiments, the transit time from addition of label to appearance of AEP in the supernatant medium was 190 to 215 s. These values were obtained by plotting the counts incorporated into the AEP bands versus time and extrapolating to zero counts.

The number of counts in the AEP band increased gradually from 19,000 at 30 min to 20,600 at 60 min. If the 60-min value is taken as 100% AEP secreted, then over 50% of labeled AEP had appeared in the medium by 6 min after the start of the pulse. Little or no labeled AEP was detectable in immunoprecipitates of cell extracts of samples taken at 30 min or later (data not shown). All these results indicate that AEP secretion is a rapid process.

Localization of secreted AEP. Since cell extracts were prepared from washed intact cells, labeled AEP which was secreted but was trapped in the cell wall should be detected in immunoprecipitates of cell extracts. The absence of labeled AEP in cell extracts 30 min after the start of the chase suggests (i) that AEP which has passed through the cytoplasmic membrane is not retained by the cell wall for a significant length of time and (ii) that most of the secreted AEP was in the supernatant medium and not cell associated.

The localization of secreted AEP was also examined by using the low-molecular-weight proteolytic substrate [^{3}H]TAME. The assumption was that this substrate would be able to penetrate the cell wall but not the cell membrane (13); therefore, cell-associated but not intracellular protease activity would be measured. The relative hydrolysis of [^{3}H]TAME by the cell-free supernatant medium and by washed cells indicated that 99% of the activity was in the cell-free supernatant (Table 1). Results with strain PS3065, which produces no AEP activity, indicated that 99% of [^{3}H]TAME hydrolysis by the supernatant medium and 77% by the washed cells was due to AEP activity.

Precursors of AEP. The purified AEP used to produce antibodies ran as a single band on SDS-polyacrylamide gels. In Ouchterlony double-diffusion analysis, the antiserum (one-, two-, and fourfold dilutions) formed a single precipitation band against the purified protease (1.3 μg per well) (data not shown).

In early studies, only the 36K and 32K polypeptides were detected by immunoprecipitation of [^{3}H]leucine-labeled cell extracts. Results from pulse-chase experiments, however, suggested that processing had not been adequately stopped

TABLE 1. Hydrolysis of [^{3}H]TAME by washed cells and cell-free supernatant medium from wild type and a strain containing an *xpr2* mutation

Strain	[^{3}H]TAME hydrolysis[a]		% cell-associated activity
	Supernatant fraction	Washed cells	
CX161-1B	1.33 ± 0.18	0.013 ± 0.002	1
PS3065	0.011 ± 0.018	0.003 ± 0.003	21

[a] Values given represent the average and standard deviation in micromoles of [^{3}H]TAME hydrolyzed per minute per 250 Klett units of cells for three independent experiments, each performed in duplicate.

after the samples were taken. The amount of the 36K putative precursor never increased from one time point to another, and it was always less than the amount of 32-kilodalton mature AEP. With improved procedures for stopping processing, an additional polypeptide (44K polypeptide) and a diffuse band (~55 kilodaltons) running under the heavy-chain IgG band were also detected. We assumed that the diffuse band was an artifact caused by trapping by the IgG heavy chain. Attempts to overcome this problem by not treating the immunoprecipitates with mercaptoethanol before SDS-PAGE or by affinity purifying the antibodies were largely unsuccessful.

A mouse MAb was prepared against AEP. The AEP MAb did not bind well to staphylococcal protein A used in our immunoprecipitation procedures. Direct binding of the MAb to Sepharose CL-4B was tried in immunoprecipitations, but the bands obtained were much less intense than for the rabbit polyclonal antibody. Cross-linking of an MAb–rabbit anti-mouse IgG–protein A-Sepharose CL-4B complex with dimethyl pimelimidate (64) gave somewhat better results, as did cross-linking the MAb with a protein A-Sepharose CL-4B complex. These procedures greatly reduced the amount of IgG heavy chain added per lane of the polyacrylamide gel, and yet a labeled band was still detected at 55 kilodaltons. This suggested that the band at 55 kilodaltons was real and not an artifact. This was confirmed by the detection of a 55K polypeptide in Western blots of cell extracts probed with the MAb (data not shown). Ultimately, interference caused by the heavy-chain IgG was overcome by increasing severalfold the amount of [^{3}H]leucine used, by increasing the exposure time for fluorography, and by decreasing severalfold the amount of sample applied per gel lane.

In recent immunoprecipitation experiments, the 55K, 52K, 44K, and 36K polypeptides, in addition to mature 32-kilodalton AEP, were detected (Fig. 2). The specificity of the immunoprecipitation was indicated by the ability of excess unlabeled AEP to block the immunoprecipitation of labeled polypeptides and by the inability of preimmune serum to immunoprecipitate any polypeptides (data not shown). An additional 29-kilodalton band was sometimes seen below the 32-kilodalton band. The amount of radioactivity in the 29-kilodalton band correlated better with that in the 55-kilodalton band than with that in the 32-kilodalton band, and the 29K polypeptide probably is a degradation product of the 55K precursor.

Results of peptide mapping suggest that the four polypeptides are related to AEP (Fig. 3) and that they therefore probably are precursors of AEP. In these samples, the 55K polypeptide probably was contaminated with 52K polypeptide and vice versa.

N-glycosylation of AEP-related polypeptides. The drug tunicamycin blocks the synthesis of N-linked glycosyl chains

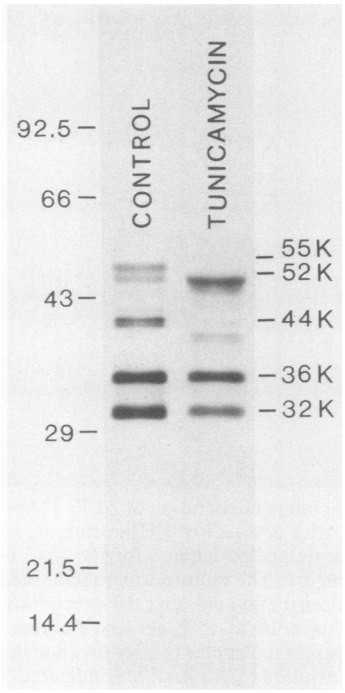

FIG. 2. N-linked glycosylation of AEP precursors. In the first lane, cells were labeled with 250 μCi of [^{3}H]leucine per 6 ml of resuspended cells for 90 s, and extracts were prepared immediately and then immunoprecipitated. The molecular size standards are the same as described in the legend to Fig. 1. In the second lane, the cells were treated for 30 min with 10 μg of tunicamycin per ml before labeling and labeled for 5 min with 250 μCi of [^{3}H]leucine per 6 ml of resuspended cells, and extracts were prepared and immunoprecipitated. The film was exposed for 12 days.

(41), and the enzyme endo H removes N-linked oligosaccharide chains from the peptide backbone (70). Immunoprecipitation of cell extracts of cells labeled in the presence of tunicamycin revealed that the 55K, 52K, and 44K polypeptides contained N-linked carbohydrate, while the 36K and 32K polypeptides did not (Fig. 2). The same result was obtained by endo H digestion of electroeluted, immunoprecipitated, AEP-related polypeptides (Fig. 4). The differences in apparent molecular mass for the treated and untreated polypeptides were about 2 kilodaltons, which corresponds to a single core oligosaccharide chain (40).

Precursor-product relationships. Several pulse-chase experiments were done in an attempt to establish precursor-product relationships among the immunoprecipitated polypeptides. If the largest precursor is converted to the next largest precursor and so on in order of decreasing size, then the peak amount of label in the larger precursors should occur earlier than for the smaller precursors. The 55K precursor appeared first, and it was the largest precursor detected (Fig. 5). With as little as a 30-s interval between samples, the amount of radioactivity in the 55K, 44K, and 36K precursors still peaked at the same time (1.5 min). Surprisingly, the peak for the 52K precursor was 1.0 min later at 2.5 min. The level of the labeled mature AEP 32K polypeptide dropped after 3.5 min when AEP began to be secreted from the cells. These results strongly suggest that processing does not occur by the single route 55K →52K→44K→36K→32K. It is also possible that the 36K and 44K polypeptides are artifacts resulting from proteolysis during sample preparation (see Discussion).

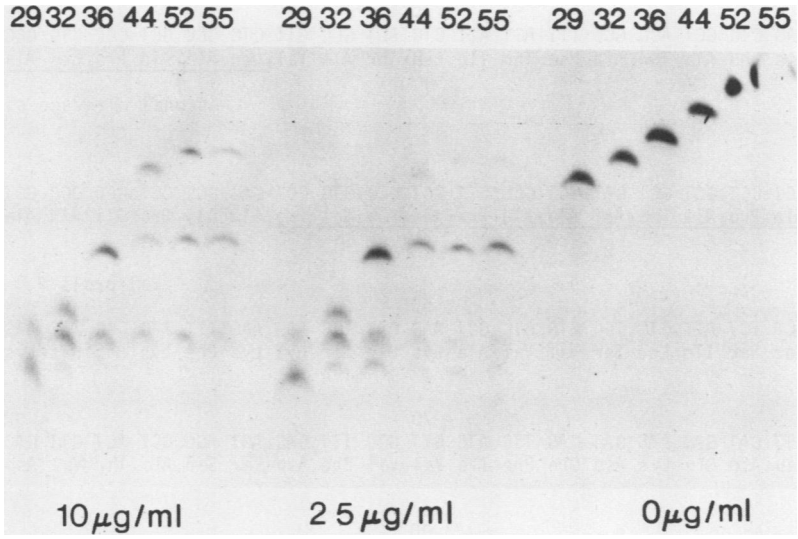

FIG. 3. Peptide mapping of labeled bands immunoprecipitated from cell extracts by anti-AEP rabbit polyclonal antibody. The bands were located by fluorography and cut out of the dried gel. An estimate of the radioactive counts in each band was obtained by scanning the fluorograph with a densitometer. The size and number of gel pieces loaded in each lane were adjusted in an attempt to equalize the number of radioactive counts in each lane. The estimated molecular mass of each sample (in kilodaltons) is indicated at the top of each lane. V8 protease (10 μl) at the concentrations indicated at the bottom of the figure was added to each lane. The gel was run at 20 mA for 40 min, and the power was then shut off for 30 min. The gel was then run until the dye front reached the top of the separating gel, and the power was shut off again for 30 min. The gel was then run at 20 mA for 6.5 h, dried, and fluorographed. The gel was exposed for 14 days.

Cloning of *XPR2* gene. The *XPR2* gene which codes for AEP (66) was cloned from strain CX161-1B and sequenced so that the positions and types of processing sites could be identified. The *XPR2* gene was cloned from a λ Charon 4 library of *Y. lipolytica* DNA by using two long oligonucleotide probes synthesized on the basis of the N-terminal amino acid sequence of mature AEP (52) and on the codon bias of the *Y. lipolytica LEU2* gene (25). Of 15,000 plaques, 6 hybridized to both probes, and one of these, λ33-1, was chosen for subsequent studies.

A small *Bgl*II-*Kpn*I fragment of λ33-1 which hybridized to the probes was sequenced. The translated DNA sequence contained the amino-terminal sequence of mature AEP (52), thus proving that the *XPR2* gene had been cloned (Fig. 6). Two sequence differences were found, but these were for amino acids at positions 14 and 17 of mature AEP for which the original amino acid sequence was not considered confirmed (52).

While this project was in progress, a patent application from Pfizer, Inc. (New York, N.Y.) containing the DNA

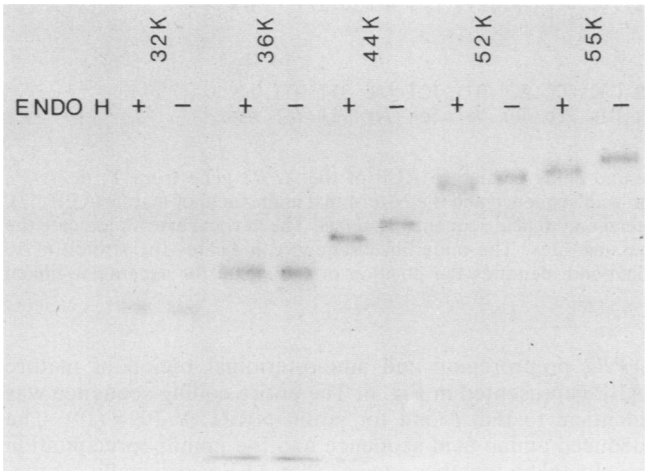

FIG. 4. N-linked glycosylation of AEP-related polypeptides. Bands obtained from the sample used in the control lane of Fig. 2 were electroeluted and digested with endo H as described in Materials and Methods. The estimated molecular mass (in kilodaltons) of each sample is indicated at the top of the gel. The samples were analyzed by SDS-PAGE and fluorography. Film exposure was for 12 days.

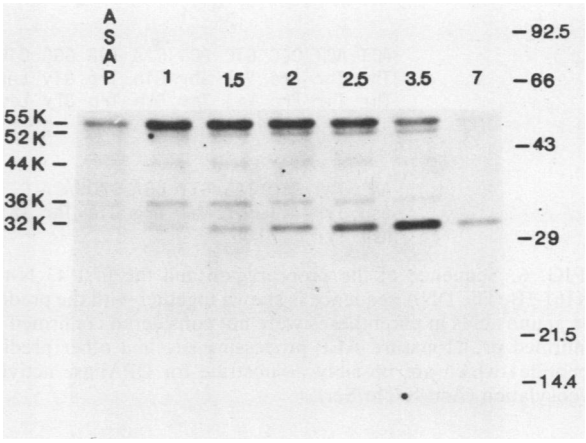

FIG. 5. Pulse-chase experiment. Cells were labeled for 45 s with 277 μCi of [^{3}H]leucine per 6 ml of resuspended cells, excess unlabeled leucine was added, one sample was taken as soon as possible (ASAP), and the remaining samples were obtained at the indicated times (in minutes) after the start of the chase. Cell extracts were immunoprecipitated, and the immunoprecipitates were analyzed by SDS-PAGE and fluorography. Total trichloroacetic acid-precipitable radioactive counts incorporated into 6 ml of cells and supernatant medium were 36,000,000 for the 1-min sample, 26,900,000 for the 2.5-min sample, and 28,800,000 for the 3.5-min sample. Film exposure was for 14 days. Molecular size standards (in kilodaltons) are shown on the right side of the figure.

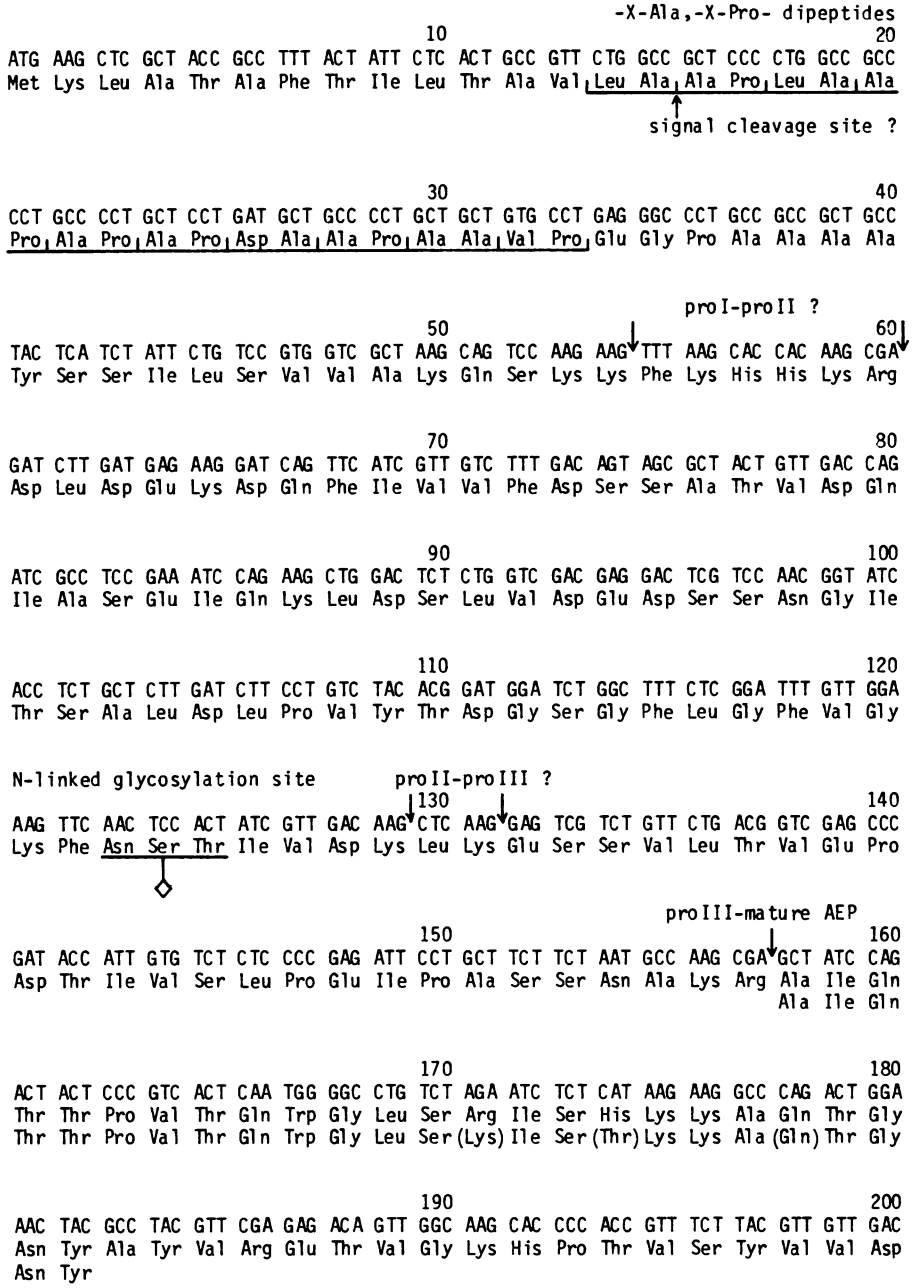

FIG. 6. Sequence of the preproregion and the first 43 N-terminal amino acids of mature AEP of the *XPR2* gene from *Y. lipolytica* CX161-1B. The DNA sequence is shown together with the predicted amino acid sequence and the N-terminal amino acid of mature AEP (52). The amino acids in parentheses were not considered confirmed in the original amino acid sequence analysis. The vertical arrows indicate the confirmed proIII-mature AEP processing site and other predicted processing sites. The underlined sequence indicates the stretch of 10 dipeptides which are possibly a substrate for DPAPase activity. The diamond identifies the position of the signal for asparagine-linked glycosylation (Asn-X-Thr/Ser).

sequence of the *XPR2* gene from *Y. lipolytica* NRRL Y-1094 appeared (European Patent Office application 0220864, E. P. Bulletin 1987/19). Crosses between CX161-1B and strains related to NRRL Y-1094 result in very poor spore viability (unpublished data), suggesting that the strains have substantial genetic differences. Because even minor changes in the amino acid sequence could cause significant problems in identifying processing sites, especially in planned radiosequencing experiments, the sequencing of the *XPR2* gene from strain CX161-1B was completed.

Predicted AEP-processing sites. The DNA sequence of the *XPR2* preproregion and amino-terminal region of mature AEP is presented in Fig. 6. The entire coding sequence was identical to that found for strain NRRL Y-1094 (19). The deduced amino acid sequence and the immunoprecipitation results taken together suggest that AEP is originally synthesized with an additional preproI-proII-proIII amino-terminal region.

A pair of basic amino acids (Lys and Arg at positions 156 and 157, respectively) directly precedes the N-terminal sequence of mature AEP. This indicates that the cleavage at the proIII-mature AEP processing site is probably done by a

TABLE 2. Comparisons of calculated molecular masses of AEP and AEP precursors as determined from the DNA sequence with their apparent molecular masses as determined by SDS-PAGE

Polypeptide designation	Putative processing event	Polypeptide N-terminal amino acid	Calculated molecular mass of polypeptide portion (daltons)	N-linked oligosaccharide	Estimated total molecular mass (daltons)[a]	Molecular mass on SDS-PAGE (kDa)[b]
55K	None; initiator Met	Met-1	46,903	+	48,900	54.8 ± 0.7 (14)
	Signal peptidase	Ala-16	45,357	+	46,800	
52K	DPAPase	Glu-34	43,807	+	45,800	52.2 ± 0.6 (12)
44K	*KEX2*-like at proI-proII site	Phe-55	41,720	+	43,700	44.1 ± 0.9 (14)
		Asp-61	40,886	+	42,900	
36K	Unknown proteolytic at proII-proIII site	Leu-130	33,487	−	33,500	36.4 ± 1.0 (16)
		Glu-132	33,246	−	33,200	
32K	*KEX2*-like at proIII-mature AEP site	Ala-158	30,524	−	30,500	31.9 ± 1.1 (16)

[a] Assuming 2 kilodaltons (kDa) for the oligosaccharide chain.
[b] Mean ± standard deviation (number of determinations).

processing endoprotease similar to the *KEX2* gene product of *S. cerevisiae* (32). There are two other pairs of basic amino acids (Lys and Lys at positions 53 and 54 and Lys and Arg at positions 59 and 60) in the prepro region, both of which are appropriately located to yield the 44K precursor after a *KEX2*-like cleavage at the proI-proII processing site.

The fact that the 55K, 52K, and 44K precursors contain N-linked carbohydrate but the 36K precursor and mature AEP do not indicates that the carbohydrate is attached in the proII region. The only Asn-X-Ser/Thr sequence, the recognition sequence for N-linked glycosylation (29), in the preproregion is in the proII sequence (Asn at position 123) as predicted from the immunoprecipitation data. The proII-proIII processing site must be downstream of this glycosylation site because the 36K precursor does not contain N-linked carbohydrate. There is an Asn-Phe-Ser sequence in mature AEP which does not seem to be glycosylated.

The predicted N-terminal sequence starting at methionine at position 1 contains many features of a secretory signal sequence (76): a basic N-terminal region, a central hydrophobic region, and a more polar C-terminal region. Based on the method of von Heijne (77) for predicting signal sequence cleavage sites, the most likely site is after alanine at position 15. Starting with leucine at position 14, there is a run of 20 amino acids which fit the pattern -X-Ala- or -X-Pro- characteristic of a substrate for a dipeptidyl aminopeptidase type IV (DPAPase) enzyme (31, 36, 37). The 52K precursor may result from DPAPase processing.

The molecular masses of the precursors which would result from processing at the sites predicted above were calculated based on the deduced amino acid sequence of AEP (Table 2). These values were compared with the apparent molecular masses of the immunoprecipitated polypeptides which were determined by SDS-PAGE (Table 2). In all cases, the values determined by SDS-PAGE were somewhat higher than the predicted values. The results are fairly consistent, especially for the 44K, 36K, and 32K polypeptides for which the differences ranged from 0.4 to 3.2 kilodaltons. The differences were significantly greater (5.4 to 8.2 kilodaltons) for the 55K and 52K precursors. Purified mature AEP may contain low levels of carbohydrate (52). It is not N linked, so it probably is O linked. This might account for some of the differences between calculated and experimentally determined molecular masses.

Secreted 19K and 20K polypeptides are derived from AEP preproregion. The possibility that the secreted 19K and 20K polypeptides (Fig. 1) were derived from the preproregion of AEP was considered. If this was true, then this would confirm that one pathway of AEP processing would be the direct formation of mature AEP from a large AEP precursor(s) without cleavage at the proI-proII and proII-proIII sites.

The 19K and 20K polypeptides are the sizes expected for the proregion of the 52-kilodalton AEP precursor (see Discussion). The estimated molecular masses were 19,400 ± 1,000 daltons for the 19K polypeptide (six determinations) and 20,300 ± 1,200 daltons for the 20K polypeptide (five determinations). The calculated molecular mass for the polypeptide portion of the proregion of AEP beginning at Glu-34 (see Discussion) was 13,283 daltons. Adding 2,000 for the N-linked carbohydrate, the estimated total molecular mass was 15,300 daltons, and again the calculated values are lower than the estimates from SDS-PAGE.

The addition of PMSF to the culture medium before labeling stabilized these polypeptides (Fig. 7A). Both polypeptides were relatively unstable in the absence of PMSF (Fig. 1), which suggests that a PMSF-sensitive protease(s) (perhaps AEP itself) normally degrades these polypeptides.

A pulse-chase experiment was done in the presence of PMSF, and the labeled polypeptides in the supernatant medium were resolved by SDS-PAGE and detected by fluorography. The bands corresponding to mature AEP and the 19K and 20K polypeptides were cut out, and their radioactivity was counted. Assuming these polypeptides contained the number of leucines predicted for the proregion of AEP, the number of molecules recovered was nearly 90% of that for mature AEP (data not shown).

Based on the immunoprecipitation and sequencing results, if the 19K and 20K polypeptides are derived from the AEP proregion, then several predictions can be made. As shown below, these are all confirmed.

First, the 19K and 20K polypeptides should contain about 2 kilodaltons of N-linked oligosaccharide. This was confirmed by endo H treatment of electroeluted samples (Fig. 7B). The presence of N-linked carbohydrate on these polypeptides and the fact that they did not react with polyclonal antibody prepared against mature AEP indicate that they are not derived from mature AEP.

Second, an AEP structural gene mutation which prevents secretion of mature AEP might be expected to prevent secretion of the AEP proregion. This was confirmed by a pulse-chase experiment with strain PS6069 containing the *xpr2-34* mutation (66). Neither mature AEP nor the 19K and 20K polypeptides were secreted by this strain (Fig. 7C, lanes 2 and 3).

Third, a mutation which affects AEP processing should affect production of the 19K and 20K polypeptides. This was confirmed by a pulse-chase experiment with strain DO613

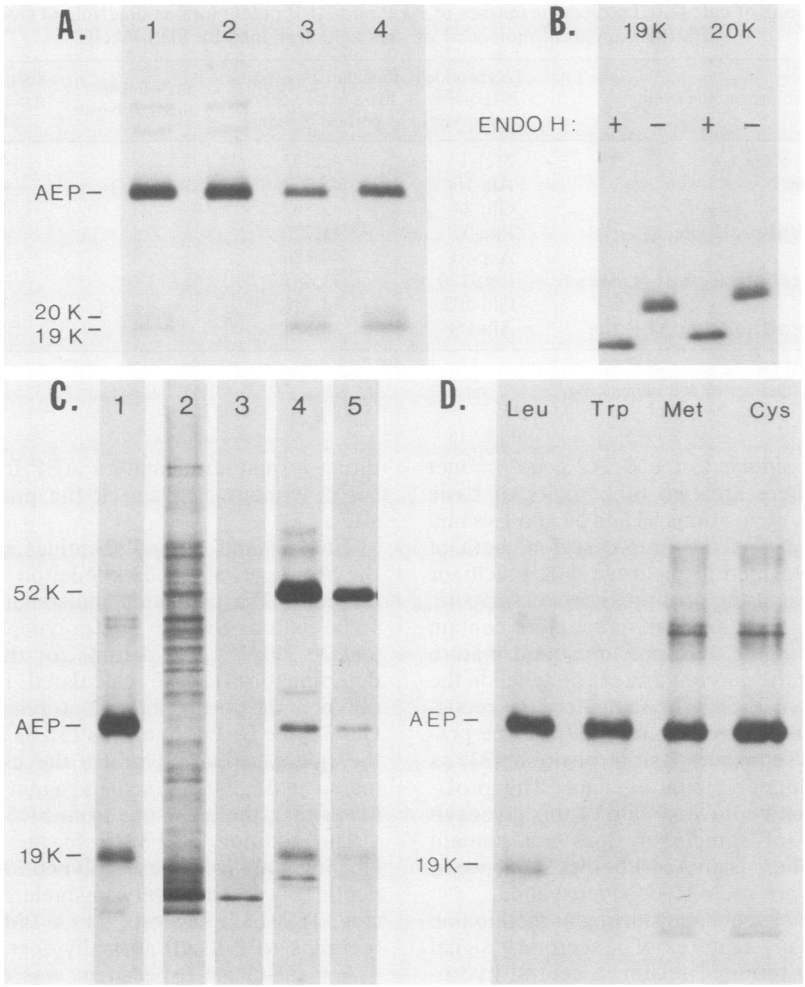

FIG. 7. The 19K and 20K polypeptides are derived from the AEP proregion. All the figures are fluorographs of pulse-chase supernatant samples (90-s labeling) resolved by SDS-PAGE on 10 to 15% gradient gels. (A) Stabilization of the 19K and 20K polypeptides by PMSF. Lanes 1 and 2, Samples from a standard [^{3}H]leucine pulse-chase experiment. Lanes 3 and 4, Samples from a pulse-chase experiment in which 2 mM PMSF was added 10 min before the start of labeling. Samples were taken 10 min (lanes 1 and 3) and 20 min (lanes 2 and 4) after the start of the chase. The equivalent of 40 μl of supernatant medium was applied per lane, and the film was exposed for 3 days. (B) The 19K and 20K polypeptides contain N-linked carbohydrate. Electroeluted samples were treated with endo H as described in Materials and Methods. A total of 5,000 counts were applied per lane, and the film was exposed for 3 days. (C) Mutations affecting production and processing of AEP affect appearance of the 19K and 20K polypeptides. Pulse-chase experiments were done in the presence (lanes 3 and 5) and absence (lanes 1, 2, and 4) of 2 mM PMSF (as in panel A). Lane 1, 9-min sample from wild-type CX161-1B. Lanes 2 and 3, 10-min samples from PS6069 containing the *xpr2-34* mutation. Lanes 4 and 5, 10-min samples from DO613 containing the *xpr6-13* mutation. The lanes are nonadjacent lanes from the same gel. The equivalent of 40 μl of supernatant medium was applied per lane, and the film was exposed for 4 days. (D) The 19K and 20K polypeptides do not contain tryptophan, methionine, or cysteine as predicted from the translated DNA sequence. Pulse-chase experiments were done by labeling the cells with [^{3}H]leucine, [^{3}H]tryptophan, [^{35}S]methionine, and [^{35}S]cysteine at 188, 375, 375, and 375 μCi per 6 ml of resuspended cells, respectively. Counts incorporated per 6 ml of culture were 31,280,000, 15,300,000, 132,050,000, and 13,600,000, respectively. Samples were taken 10 min after the start of the chase. The equivalent of 30, 75, 15, and 120 μl of supernatant medium was applied per lane for Leu, Trp, Met, and Cys, respectively. Based on the translated DNA sequence, mature AEP contains 18 leucines, 5 tryptophans, 1 methionine, and 3 cysteines. The film was exposed for 4 days.

containing the *xpr6-13* mutation. This strain secretes the 52-kilodalton AEP precursor and relatively little of the mature AEP and 19K and 20K polypeptides (Fig. 7C, lanes 4 and 5).

Finally, the AEP proregion contains no cysteine, methionine, or tryptophan. This was confirmed for the 19K and 20K polypeptides by pulse-chase experiments in which cells were labeled with [^{35}S]cysteine, [^{35}S]methionine, or [^{3}H]tryptophan. These amino acids are present in mature AEP, and in each case mature AEP was clearly visible on the fluorograph, but no labeled bands were detected where the 19K and 20K polypeptides would be expected (Fig. 7D). These amino acids are the least frequently found amino acids in *Y. lipolytica* proteins (23). But assuming that the 19-kilodalton band contains 17 kilodaltons of polypeptide (~155 amino acids) and using the frequencies of appearances for the three amino acids (23), the chance that an average *Y. lipolytica* 17-kilodalton protein contains none of these three amino acids is between 0.2 and 0.5%, depending on the amino acid composition used for the calculation.

Peptide mapping of the 52K, 32K, and 19K polypeptides was done in an attempt to demonstrate that the 19K poly-

peptide was from the preproregion. Some correlation could be made between the 52K and 19K polypeptides, but the overall fragment patterns obtained after individual treatment with several proteases happened not to confirm a conclusive relationship between the 19K and 52K polypeptides.

The above results strongly suggest that the 19K and 20K polypeptides are derived from the AEP proregion. In the presence of PMSF, direct conversion of 52K to 32K appears to be the major pathway for AEP processing, and 55K→ 52K→32K is definitely a major pathway in the absence of PMSF.

DISCUSSION

Secretion of AEP from *Y. lipolytica* is a rapid process. The transit time for AEP is similar to the transit times reported for *S. cerevisiae* proteins of less than 5 min for invertase (48, 62) and between 3 and 8 min for several surface proteins (49). The pulse-chase and AEP localization experiments both suggested that AEP secreted through the cell membrane is not retained to any great extent by the cell wall and appears rapidly in the extracellular medium.

Four polypeptides which are precursors of mature AEP were detected in *Y. lipolytica* cell extracts. Several results support our conclusions that these polypeptides are related to AEP. Their immunoprecipitation by antibody raised against AEP and immunocompetition by purified AEP indicate that they are antigenically related to AEP. The similarities of their peptide maps with that of mature AEP further demonstrates this relatedness. Results of the pulse-chase experiments do not fit simple precursor-product relationships, but do suggest that the 55K polypeptide is synthesized first and that the 52K, 44K, and 36K polypeptides are intermediates in the formation of 32-kilodalton mature AEP. Finally, AEP precursors of approximately these sizes can be predicted from the amino acid sequence deduced from the coding sequence of the *XPR2* structural gene.

The following model of AEP processing is suggested. AEP is originally synthesized as a preproI-proII-proIII AEP polypeptide. On the basis of the DNA sequence data, the molecule seems to have an N-terminal signal peptide. It cannot be determined from the immunoprecipitation data whether or not the 55K polypeptide contains the signal peptide. Removal of the signal peptide would be expected to occur cotranslationally in a eucaryote, and the full-length precursor would not be detected in vivo (57). The differences between the calculated and measured molecular masses of the 55K polypeptide (Table 2) are minimized if one assumes that the 55K polypeptide contains the signal peptide. The significantly greater differences in these values for the 52K and 55K polypeptide compared with the 32K, 36K, and 44K polypeptides may be due to the presence of the -X-Ala-, -X-Pro-dipeptide stretch. Removal of an Ala- and Pro-rich sequence from the bovine β-crystalline B_1 chains resulted in an unexpectedly large decrease in apparent M_r which was attributed to the inability of the Ala- and Pro-rich region to bind normal amounts of SDS (5). Another possibility is an as yet unknown posttranslational modification in the preproI region of AEP.

The first precursor seen (55K polypeptide) contains about 2 kilodaltons of N-linked carbohydrate. This is consistent with cotranslational addition of core oligosaccharide as the AEP precursor enters the lumen of the rough endoplasmic reticulum (22). Essentially nothing is known about glycosylation in *Y. lipolytica*, but if it is identical to glycoslation in *S. cerevisiae*, then it would be expected that three glucoses and a mannose would be removed before the 55K precursor was processed to the 52K and 44K precursors (38). Reliably detecting such a small difference (~0.65 kilodaltons) in the mass of the attached carbohydrates would be difficult for polypeptides of the size of the glycosylated AEP precursors. There is no evidence for the addition of high-mannose sugars to AEP as occurs for many *S. cerevisiae* glycoproteins (3).

The late appearance of the 52K precursor relative to the 44K and 36K precursors was difficult to explain until the DNA sequence of the preproregion was obtained. The presence of a stretch of 10 consecutive dipeptides with the sequence of -X-Ala- or -X-Pro- suggested that DPAPase activity was involved in AEP processing. N-terminal amino acid sequencing of the 52-kilodalton AEP precursor secreted by strains containing an *xpr6* mutation confirm that DPA Pase activity is involved in formation of the 52K precursor (Matoba and Ogrydziak, manuscript submitted).

Our working hypothesis is that the 55K polypeptide contains the signal peptide and one core N-linked oligosaccharide. As it moves through the secretory pathway, it first encounters the *KEX2*-like processing endoprotease(s) and the processing enzyme which cleaves at the proII-proIII site to yield the 36K precursor. A Lys-Arg pair of basic amino acids directly precedes the N-terminal amino acid sequence of mature AEP, confirming that this is a *KEX2*-like processing site (32). The cleavage which yields the 44K precursor is predicted to be *KEX2*-like and most likely is after Lys-Arg at positions 59 and 60. The nature of the proII-proIII cleavage is unknown except that it must occur downstream of the glycosylation site. Conversion of the 55K precursor to mature AEP occurs rather slowly in the early part of the secretory pathway. Later in the secretory pathway, the 55K precursor encounters the DPAPase activity which results in formation of the 52K precursor. If the signal peptide is not removed by signal peptidase activity, then another proteolytic cleavage (possibly after Val at position 13) would be required to allow DPAPase activity to initiate, because DPAPase requires a free N-terminal dipeptide as a substrate (37). Mature AEP is formed at a higher rate in this part of the secretory pathway, which suggests that the 52K polypeptide is a better substrate than the 55K polypeptide for processing to mature AEP or that conditions are more favorable for further processing in this part of the secretory pathway.

The secretion of the intact proregion of AEP confirms that one processing route is from the 52K precursor directly to the 32K polypeptide. The 52K polypeptide is favored as the direct precursor for several reasons. First, the 19K and 20K polypeptides are closer to the size expected after removal of 32-kilodalton mature AEP from a 52-kilodalton polypeptide than from a 55-kilodalton polypeptide. Second, in some preparations, two distinct bands (19 and 20 kilodaltons) are visible (Fig. 1), but in others only one band (19 kilodaltons) is visible and its upper portion is fuzzy (Fig. 7). The 52-kilodalton AEP precursor secreted by strains with an *xpr6* mutation has a similar fuzzy appearance (Matoba and Ogrydziak, manuscript submitted). We suspect that the 20-kilodalton band consists of polypeptides containing various extents of the -X-Ala-, -X-Pro- dipeptide stretch and that this stretch of 10 dipeptides is completely removed from the 19-kilodalton band. If the 19K polypeptide is derived from the 52-kilodalton AEP precursor, then the 20K polypeptide cannot be derived directly from the 55-kilodalton AEP precursor because the difference in estimated molecular mass between the 19K and 20K polypeptides (0.9 kilodaltons) is much less than for the 52- and 55-kilodalton AEP precursors (2.5 kilodaltons). Also, the fact that the 19K and

20K polypeptides were not labeled with [^{35}S]methionine indicates that they are not derived directly from the original translation product containing the Met-1 initiator methionine.

The presence of the AEP-related 44K and 36K polypeptides and the results of the pulse-chase experiments suggest that there are multiple pathways for AEP processing. An alternative possibility is that there is only one pathway, 55K→52K→32K, and that the AEP-related 44K and 36K polypeptides are artifacts resulting from proteolysis of the 55- or 52-kilodalton AEP precursor or both during sample preparation. In this case, one might assume that the amounts of labeled 44K and 36K precursors in the sample would be proportional to the amount of labeled 55K (or 55K plus 52K) precursor. This would explain why the 55K, 44K, and 36K precursors all reached maximum labeling intensity at the same time (1.5 min) in the pulse-chase experiment (Fig. 5). However, based on the radioactive counts in the bands, the ratios of labeled 44K or 36K precursor to larger precursors were not constant. For the 1- to 2.5-min samples, the counts in the 36- and 44-kilodalton bands were severalfold higher than background, and therefore these results would be expected to be more meaningful. The ratios of 36K polypeptide/(52K polypeptide + 55K polypeptide) and 44K polypeptide/(52K polypeptide + 55K polypeptide) were both highest at 1.5 min (100%). The values ranged from 82 to 91% for the 36K polypeptide and 81 to 91% for the 44K polypeptide for the 1-, 2-, and 2.5-min samples. This suggests, but by no means proves, that the 36K and 44K polypeptides are precursors formed in vivo.

Why is AEP processing so complicated? We believe that one reason is to protect the cell from premature activation of AEP. There are many examples of proteases which are activated by removal of proregions, including extracellular proteases of *Bacillus* species (75, 78), vacuolar proteases of *S. cerevisiae* (1, 45, 74, 79), and zymogens of higher eucaryotes such as trypsinogen (47). Although AEP is an alkaline protease, it still has activity at neutral pH and below (52). The 52-kilodalton AEP precursor secreted by strains containing an *xpr6* mutation is inactive (51; unpublished data), but it is not known whether the 44K and 36K precursors are active. The pH of the endoplasmic reticulum is thought to be near neutrality, and the secretory pathway is known to become progressively more acidic as it goes along (43). The low rate of conversion of 55K polypeptide to 32K polypeptide until the 52K intermediate is formed may delay AEP processing and activation until the secretory pathway is acidic enough so that AEP will not cause damage.

We are presently testing our model of AEP processing by amino acid and radiosequencing of the N terminals of the AEP precursors. We plan to test our hypotheses on the functions of the various processing steps and on the authenticity of the 44K and 36K precursors by isolating mutations in the AEP processing enzymes and processing sites.

ACKNOWLEDGMENTS

We thank S.-C. Cheng for performing the [^{3}H]TAME experiments and J. Montero for help with antiserum preparation. We thank E. Bandman for advice and for making facilities available for the preparation of MAbs. We are extremely grateful to D. Fraser and D. Zdanis for help with MAb preparation. We thank L. Davidow for helpful discussion, P. Carpenter for preparing the manuscript, and members of the Institute of Marine Resources laboratory for critical reading of the manuscript.

The research was supported by grants PCM8011510 and PCM8307663 from the National Science Foundation, by a grant from Pfizer, Inc., supporting R.W., and by Agricultural Experiment Station fund CA-D-FST-3590-H.

LITERATURE CITED

1. **Ammerer, G., G. P. Hunter, J. H. Rothman, G. A. Saari, L. A. Valls, and T. H. Stevens.** 1986. *PEP4* gene of *Saccharomyces cerevisiae* encodes proteinase A, a vacuolar enzyme required for processing of vacuolar precursors. Mol. Cell. Biol. **6**:2490–2499.
2. **Atkinson, T., and M. Smith.** 1984. Solid phase synthesis of oligodeoxyribonucleotides by the phosphate triester method, p. 33–81. *In* M. J. Gait (ed.), Oligonucleotide synthesis: a practical approach. IRL Press, Washington, D.C.
3. **Ballou, C. E.** 1982. Yeast cell wall and cell surface, p. 335–360. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
4. **Bassel, J. B., and R. K. Mortimer.** 1982. Genetic and biochemical studies of N-alkane non-utilizing mutants of *Saccharomycopsis lipolytica*. Curr. Genet. **5**:77–88.
5. **Berbers, G. A. M., W. A. Hockman, H. Bloemendel, W. W. deJong, T. Kleinschmidt, and G. Braunitzer.** 1983. Proline- and alanine-rich N-terminal extension of the basic bovine β-crystalline B_1 chains. FEBS Lett. **161**:225–229.
6. **Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
7. **Bonner, W. M.** 1983. Use of fluorography for sensitive isotope detection in polyacrylamide gel electrophoresis and related techniques. Methods Enzymol. **96**:215–222.
8. **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.
9. **Cheng, S.-C., and D. M. Ogrydziak.** 1986. Extracellular RNase produced by *Yarrowia lipolytica*. J. Bacteriol. **168**:581–589.
10. **Cheng, S.-C., and D. M. Ogrydziak.** 1987. Processing and secretion of the *Yarrowia lipolytica* RNase. J. Bacteriol. **169**: 1433–1440.
11. **Chirico, W. J., M. G. Waters, and G. Blobel.** 1988. 70K heat shock related proteins stimulate protein translocation into microsomes. Nature (London) **332**:805–810.
12. **Chvatchko, Y., I. Howald, and H. Riezman.** 1986. Two yeast mutants defective in endocytosis are defective in pheromone response. Cell **46**:355–364.
13. **Ciejek, E., and J. Thorner.** 1979. Recovery of *S. cerevisiae a* cells from GI arrest by α factor pheromone requires endopeptidase action. Cell **18**:623–635.
14. **Clare, J. J., L. S. Davidow, D. C. J. Gardner, and S. G. Oliver.** 1986. Cloning characterizations of the ribosomal RNA genes of the dimorphic yeast, *Yarrowia lipolytica*. Curr. Genet. **10**:449–452.
15. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1102–1106.
16. **Cryer, D. R., R. Eccleshall, and J. Marmur.** 1975. Isolation of yeast DNA. Methods Cell Biol. **12**:39–44.
17. **Davidow, L. S., D. Apostolakos, M. M. O'Donnell, A. R. Proctor, D. M. Ogrydziak, R. A. Wing, I. Stasko, and J. R. DeZeeuw.** 1985. Integrative transformation for the yeast *Yarrowia lipolytica*. Curr. Genet. **10**:39–48.
18. **Davidow, L. S., F. S. Kaczmarek, J. R. DeZeeuw, S. W. Conlon, M. R. Lauth, D. A. Pereira, and A. E. Franke.** 1986. The *Yarrowia lipolytica LEU2* gene. Curr. Genet. **11**:377–383.
19. **Davidow, L. S., M. M. O'Donnell, F. S. Kaczmarek, D. A. Pereira, J. R. DeZeeuw, and A. E. Franke.** 1987. Cloning and sequencing of the alkaline extracellular protease gene of *Yarrowia lipolytica*. J. Bacteriol. **169**:4621–4629.
20. **Deshaies, R. J., B. D. Koch, M. Werner-Washburne, E. A. Craig, and R. Schekman.** 1988. A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature (London) **332**:800–805.

21. **Efimov, V. A., A. A. Burykova, S. V. Reverdatto, O. G. Chakmakcheva, and Y. A. Ouchinnikov.** 1983. Rapid synthesis of long-chain deoxyribonucleotides by the *N*-methylimidazolide phosphotriester method. Nucleic Acids Res. **11**:8369–8387.
22. **Esmon, B., P. Novick, and R. Schekman.** 1981. Compartmentalized assembly of oligosaccharide chains on exported glycoproteins in yeast. Cell **25**:451–460.
23. **Evans, G. H.** 1968. Industrial production of single-cell protein from hydrocarbons, p. 243–254. *In* R. I. Mateles and S. R. Tannenbaum (ed.), Single-cell protein. MIT Press, Cambridge, Mass.
24. **Fournier, P., C. Gaillardin, M.-A. Persuy, J. Klootwijk, and H. van Heerikhuizen.** 1986. Heterogeneity in the ribosomal family of the yeast *Yarrowia lipolytica*: genomic organization and segregation studies. Gene **42**:273–282.
25. **Gaillardin, C., and A.-M. Ribet.** 1986. *LEU2* directed expression of β-galactosidase activity and phleomycin resistance in *Yarrowia lipolytica*. Curr. Genet. **11**:369–375.
26. **Gaillardin, C., A.-M. Ribet, and H. Heslot.** 1985. Integrative transformation of the yeast *Yarrowia lipolytica*. Curr. Genet. **10**:49–58.
27. **Hames, B. D.** 1981. An introduction to polyacrylamide gel electrophoresis, p. 1–91. *In* B. D. Hames and D. Rickwood (ed.), Gel electrophoresis of proteins: a practical approach. IRL Press, Washington, D.C.
28. **Herbert, D., P. J. Phipps, and E. R. Strange.** 1971. Chemical analysis of microbial cells. Methods Microbiol. **5B**:209–344.
29. **Hubbard, S. C., and R. J. Ivatt.** 1981. Synthesis and processing of asparagine-linked oligosaccharides. Annu. Rev. Biochem. **50**: 555–583.
30. **Johnson, L. M., V. A. Bankaitis, and S. D. Emr.** 1987. Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease. Cell **48**:875–885.
31. **Julius, D., L. Blair, A. Brake, G. Sprague, and J. Thorner.** 1983. Yeast α factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. Cell **32**:839–852.
32. **Julius, D., A. Brake, L. Blair, R. Kinisawa, and J. Thorner.** 1984. Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-α-factor. Cell **37**:1075–1089.
33. **Julius, D., R. Schekman, and J. Thorner.** 1984. Glycosylation and processing of prepro-α-factor through the yeast secretory pathway. Cell **36**:309–318.
34. **Kaiser, D. A., D. Preuss, P. Grisafi, and D. Botstein.** 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. Science **235**:312–317.
35. **Kapoor, K. K., K. Chaudhary, and P. Tauro.** 1982. Citric acid, p. 709–747. *In* G. Reed (ed.), Prescott and Dunn's industrial microbiology, 4th ed. AVI Publishing Co., Inc., Westport, Conn.
36. **Kreil, G.** 1985. Late reactions in the processing of peptide precursors: stepwise cleavage of dipeptides and formation of terminal amides, p. 49–52. *In* R. B. Freedman and H. C. Hawkins (ed.), The enzymlogy of post-translational modification of proteins, vol. 2. Academic Press, Inc. (London), Ltd., London.
37. **Kreil, G., L. Haiml, and G. Suchanek.** 1980. Stepwise cleavage of the pro part of promelittin by dipeptidylpeptidase IV. Eur. J. Biochem. **111**:49–58.
38. **Kukuruzinska, M. A., M. L. E. Bergh, and B. J. Jackson.** 1987. Protein glycosylation in yeast. Annu. Rev. Biochem. **56**:915–944.
39. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
40. **Lehle, L.** 1980. Biosynthesis of the core region of yeast mannoproteins. Eur. J. Biochem. **109**:589–601.
41. **Mahoney, W. C., and D. Duskin.** 1979. Biological activities of the two major components of tunicamycin. J. Biol. Chem. **254**: 6572–6576.
42. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
43. **Mellman, I., R. Fuchs, and A. Helenius.** 1986. Acidification of the endocytic and exocytic pathways. Annu. Rev. Biochem. **55**: 663–700.
44. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
45. **Moehle, C. M., R. Tizard, S. K. Lemmon, J. Smart, and E. W. Jones.** 1987. Protease B of the lysosomelike vacuole of the yeast *Saccharomyces cerevisiae* is homologous to the subtilisin family of serine proteases. Mol. Cell. Biol. **7**:4390–4399.
46. **Nakase, T., and K. Komagata.** 1968. Taxonomic significance of base composition of yeast DNA. J. Gen. Appl. Microbiol. **14**: 345–357.
47. **Neurath, H.** 1975. Limited proteolysis and zymogen activation, p. 51–64. *In* E. Reich, D. B. Rifkin, and E. Shaw (ed.), Proteases and biological control. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
48. **Novick, P., S. Ferro, and R. Schekman.** 1981. Order of events in the yeast secretory pathway. Cell **25**:461–469.
49. **Novick, P., and R. Schekman.** 1983. Export of major cell surface proteins is blocked in yeast secretory mutants. J. Cell Biol. **96**: 541–547.
50. **Ogrydziak, D., J. Bassel, R. Contopoulou, and R. Mortimer.** 1978. Development of genetic techniques and the genetic map of the yeast *Saccharomycopsis lipolytica*. Mol. Gen. Genet. **163**: 229–239.
51. **Ogrydziak, D. M., S.-C. Cheng, and S. J. Scharf.** 1982. Characterization of *Saccharomycopsis lipolytica* mutants producing lowered levels of alkaline extracellular protease. J. Gen. Microbiol. **128**:2271–2280.
52. **Ogrydziak, D. M., and S. J. Scharf.** 1982. Alkaline extracellular protease produced by *Saccharomycopsis lipolytica* CX161-1B. J. Gen. Microbiol. **128**:1225–1234.
53. **Oi, V. T., and L. A. Herzenberg.** 1980. Immunoglobulin-producing hybrid cell lines, p. 351–371. *In* B. B. Mishell and S. M. Shugi (ed.), Selected methods in cellular immunology. W. H. Freeman and Co., San Francisco.
54. **Ota, Y., K. Goni, S. Kato, T. Sugiura, and Y. Minoda.** 1982. Purification and some properties of cell-bound lipase from *Saccharomycopsis lipolytica*. Agric. Biol. Chem. **46**:2885–2893.
55. **Poritz, M. A., V. Siegel, W. Hansen, and P. Walter.** 1988. Small ribonucleoproteins in *Schizosaccharomyces pombe* and *Yarrowia lipolytica* homologous to signal recognition particle. Proc. Natl. Acad. Sci. USA **85**:4315–4319.
56. **Powers, S., S. Michaelis, D. Brock, S. Santa Anna-A., J. Field, I. Herskowitz, and M. Wigler.** 1986. *RAM*, a gene of yeast required for a functional modification of *RAS* proteins and for production of mating pheromone a-factor. Cell **47**:413–422.
57. **Rapoport, T. A., and M. Wiedmann.** 1985. Application of the signal hypothesis to the incorporation of integral membrane proteins. Curr. Top. Membr. Transp. **24**:1–63.
58. **Reed, K. C., and D. A. Mann.** 1985. Rapid transfer from agarose gels to nylon membranes. Nucleic Acids Res. **13**:7207–7221.
59. **Rodriguez, C., and A. Dominguez.** 1984. The growth characteristics of *Saccharomycopsis lipolytica*: morphology and induction of mycelium formation. Can. J. Microbiol. **30**:605–612.
60. **Roffman, S., V. Sanocka, and W. Troll.** 1970. Sensitive proteolytic enzyme assay using differential solubilities of radioactive substrates and products in biphasic systems. Anal. Biochem. **36**:11–17.
61. **Rothman, J. H., and T. H. Stevens.** 1986. Protein sorting in yeast: mutants defective in vacuole biogenesis mislocalize vacuolar proteins into the late secretory pathway. Cell **47**:1041–1051.
62. **Salminen, A., and P. J. Novick.** 1987. A *ras*-like protein is required for a post-golgi event in yeast secretion. Cell **49**:527–538.
63. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
64. **Scheider, C., R. A. Newman, D. R. Sutherland, V. Asser, and M. F. Greaves.** 1982. A one-step purification of membrane proteins using a high efficiency immunomatrix. J. Biol. Chem.

257:10766–10769.
65. **Schekman, R.** 1985. Protein localization and membrane traffic in yeast. Annu. Rev. Cell Biol. **1**:115–143.
66. **Simms, P. C., and D. M. Ogrydziak.** 1981. Structural gene for the alkaline extracellular protease of *Saccharomycopsis lipolytica*. J. Bacteriol. **145**:404–409.
67. **Stahl, U.** 1978. Zygote formation and recombination between like mating types in the yeast *Saccharomycopsis lipolytica* by protoplast fusion. Mol. Gen. Genet. **160**:111–113.
68. **Stenderup, A., and A. Leth Bak.** 1968. Deoxyribonucleic acid base composition of some species within the genus *Candida*. J. Gen. Microbiol. **52**:231–236.
69. **Stevens, T., B. Esmon, and R. Schekman.** 1982. Early stages in the yeast secretory pathway are required for transport of carboxypeptidase Y to the vacuole. Cell **30**:439–448.
70. **Tarentino, A. L., T. H. Plummer, Jr., and F. Maley.** 1974. The release of intact oligosaccharides from specific glycoproteins by endo-β-*N*-acetylglucosaminidase H. J. Biol. Chem. **249**:818–824.
71. **Tobe, S., T. Takami, S. Ikeda, and K. Mitsugi.** 1976. Production and some enzymatic properties of alkaline proteinase of *Candida lipolytica*. Agric. Biol. Chem. **40**:1087–1092.
72. **Trimble, R. B., and F. Maley.** 1977. Subunit structure of external invertase from *Saccharomyces cerevisiae*. J. Biol. Chem. **252**:4409–4412.
73. **Valls, L. A., C. P. Hunter, J. H. Rothman, and T. H. Stevens.** 1987. Protein sorting in yeast: the localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide. Cell **48**:887–897.
74. **van Heerikhuizen, H., A. Ykema, J. Klootwijk, C. Gaillardin, C. Ballas, and P. Fournier.** 1985. Heterogeneity in the ribosomal RNA genes of the yeast *Yarrowia lipolytica*: cloning and analysis of two size classes of repeats. Gene **39**:213–222.
75. **Vasantha, N., L. D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula.** 1984. Genes for alkaline protease and neutral protease from *Bacillus amyloliquefaciens* contain a large open reading frame between the regions coding for signal sequence and mature protein. J. Bacteriol. **159**:811–819.
76. **von Heijne, G.** 1985. Signal sequences. The limits of variation. J. Mol. Biol. **184**:99–105.
77. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
78. **Wells, J. A., E. Ferrari, D. J. Henner, D. A. Estell, and E. Y. Chen.** 1983. Cloning, sequencing and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*. Nucleic Acids Res. **11**:7911–7925.
79. **Woolford, C. A., L. B. Daniels, F. J. Park, E. W. Jones, J. N. Van Arsdell, and M. A. Innis.** 1986. The *PEP4* gene encodes an aspartyl protease implicated in the postranslational regulation of *Saccharomyces cerevisiae* vacuolar hydrolases. Mol. Cell. Biol. **6**:2500–2510.
80. **Yamada, T., and D. M. Ogrydziak.** 1983. Extracellular acid proteases produced by *Saccharomycopsis lipolytica*. J. Bacteriol. **154**:23–31.
81. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp and pUC vectors. Gene **33**:103–119.
82. **Zehnbauer, B. A., and F. R. Blattner.** 1982. Construction and screening of recombinant DNA libraries with Charon vector phages, p. 249–278. *In* J. K. Setlow and A. Hollaender (ed.), Genetic engineering: principles and methods, vol. 4. Plenum Publishing Corp., New York.