



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵: C12N 15/81, 9/04, C12P 7/10, C12N 9/32, C12Q 1/68	A1	(11) International Publication Number: WO 94/21802 (43) International Publication Date: 29 September 1994 (29.09.94)
(21) International Application Number: PCT/US94/03213 (22) International Filing Date: 24 March 1994 (24.03.94) (30) Priority Data: 037,617 25 March 1993 (25.03.93) US 037,618 25 March 1993 (25.03.93) US (71) Applicant: BIOSOURCE GENETICS CORPORATION [US/US]; 3333 Vaca Valley Parkway, Vacaville, CA 95688 (US). (72) Inventors: KUMAGAI, Monto, H.; 1330 Brown Drive, Davis, CA 95616 (US). SVERLOW, Genadie, G.; 249 Wrentham Drive, Vacaville, CA 95688 (US). (74) Agents: HALLUIN, Albert, P. et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).		(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KG, KR, KZ, LK, LV, MD, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, UA, UZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: <i>PICHA PASTORIS</i> ALCOHOL OXIDASE ZZA1 and ZZA2		
(57) Abstract		
Two novel genomic clones, ZZA1 and ZZA2, were isolated which encode for <i>Pichia pastoris</i> alcohol oxidase isozymes. The 5' non-coding region of ZZA1 contains common structural features involved in the transcription and translation of eukaryotic genes. Comparison of the nucleotide sequences of the ZZA1 and AOX1 51 noncoding regions showed that they are 66 % similar to each other. The rice α-amylase gene OS103 was placed under the transcriptional control of the ZZA1 promoter. The nucleotide sequences of ZZA1 and other methanol-regulated promoters were analyzed. A highly conserved sequence (TTGNNNGCTTCCAANNNTGGT) (SEQ ID NO:2) was found in the 51 flanking region. A yeast strain containing the ZZA1-OS103 fusion and secreting biologically active α-amylase into the culture media while converting starch to ethanol was produced. The ZZA1 and ZZA2 regulatory sequences may be used to control the expression of other heterologous proteins in multiple yeast species. Methods of purifying proteins that are regulators of alcohol oxidase expression (referred to as AOER proteins) and methods of isolating these proteins are also provided. The invention also provides for the efficient conversion of starch into ethanol.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

PICHIA PASTORIS ALCOHOL OXIDASE ZZA1 AND ZZA2

Related Applications

5 This patent application is a continuation-in-part of U.S patent application 08/037,618 filed March 25, 1993 and a continuation-in-part of U.S patent application 08/037,617 filed March 25, 1993.

10 **1. Field of the Invention**

 This invention relates to the field of recombinant DNA technology. More specifically, the invention relates to isolated proteins and nucleic acid sequences from methylotrophic yeast cells and the
15 production of ethanol using α -amylase.

2. Background

 There has been an intense interest in the structure and function of the regulatory regions
20 controlling the expression of genes in the pathway for methanol utilization. The methylotrophic yeasts are found in four genera that can be divided into two groups the ascosporogenous Hansenula and Pichia, and the asporogenous Candida and Torulopsis. The first
25 enzyme in the methanol-utilization pathway is alcohol oxidase. It catalyzes the oxidation of methanol to formaldehyde. During this reaction there is a simultaneous reduction of oxygen to hydrogen peroxide. When glucose grown yeast cells are transferred to
30 methanol containing media as a sole carbon source, the peroxisome, a subcellular organelle, begins to swell and proliferate. Large amounts of alcohol oxidase are sequestered into the peroxisomes. The
 compartmentalization of alcohol oxidase in the
35 peroxisomes protects the cytosol from hydrogen

peroxide. The high concentration of this enzyme compensates for its low affinity for oxygen. In methanol-grown Pichia pastoris cells, alcohol oxidase constitutes up to 30% of the total soluble protein.

5 Alcohol oxidases have numerous commercial applications. These uses include the measurement of alcohol levels in various biological and nonbiological fluids, and the conversion of alcohol precursor molecules into aldehydes, e.g., for use in artificial
10 flavor production.

 Several genes encoding alcohol oxidase from Pichia pastoris (AOX1, AOX2), Candida boidinii S2 (AOD1), and methanol oxidase from Hansenula polymorpha (MOX1) have been isolated and characterized. The
15 regulation of the synthesis of alcohol oxidase is primarily controlled at the level of transcription. During methanol induction, the rapid de novo synthesis of the enzyme is accompanied by a dramatic increase in alcohol oxidase mRNA. Previous studies by Ellis et
20 al., Mol. Cell. Biol., 5:1111-1121 (1985) of the AOX1 promoter indicates that it is strongly repressed by ethanol and glucose. Based on northern hybridizations, Pichia pastoris cells grown in the presence of ethanol did not synthesize alcohol oxidase
25 specific poly(A)⁺ RNA. The synthesis of alcohol oxidase is tightly catabolite repressed. In cells grown in glucose containing media, alcohol oxidase mRNA is not detectable. In order to study the regulation of the AOX1 gene Tschopp et al., Nucl.
30 Acids Res., 15:3859-3876 (1987) took the AOX1 promoter and fused it to the E. coli lacZ gene. Saccharomyces cerevisiae strains harboring an AOX1-lacZ fusion produced only small amounts of active enzyme when grown in glucose or ethanol. See Stroman et al., U.S.
35 Pat. No. 4,855,231.

The methylotrophic yeasts Hansenula polymorpha and Pichia pastoris have been used as hosts for heterologous gene expression. Cregg et al., Bio/Technology, 5:479-485 (1987); Tschopp et al., Bio/Technology, 5:1305-1308 (1987). Over twenty proteins of potential commercial value have been produced using methanol regulated promoters. These heterologous proteins can accumulate to high levels in the cytoplasm (tumor necrosis factor was expressed at 8 g/liter, Skeekrishna et al., Biochemistry, 28:4117-4125, 1989) or can be secreted into the media (Saccharomyces cerevisiae invertase was secreted at 2.5 g/liter, Tschopp et al., Bio/Technology, 5:1305-1308, 1987). Stable transformants have been obtained by integrating the alcohol oxidase promoter expression cassette into the yeast chromosomes. Increases in the level of expression have been obtained by increasing the number of integrated copies of the heterologous gene expression cassette (tetanus toxin fragment C was expressed at 12 g/liter, Clare et al., Bio/Technology, 9:455-460).

Methylotrophic yeasts have been used in the production of heterologous proteins. Increases in gene dosage, cell density, and promoter strength have resulted in high-level expression of valuable proteins. Several genes encoding alcohol oxidase (Pichia pastoris AOX1, AOX2, and Candida Boidinii S2 AOD1), methanol oxidase (Hansenula polymorpha MOX), and dihydroxyacetone synthase (Hansenula polymorpha DHAS) have been isolated and characterized, see, for example, Ellis et al., Mol. Cell. Biol, 5:1111-1121 (1985) (AOX1); Koutz et al., Yeast, 5:167-177 (1989) (AOX2); Sakai et al., Gene, 114:67-73 (1992) (AOD1); Ledebouer et al., Nucl. Acids Res., 13:3063-3082 (1985) (MOX1). The synthesis of these enzymes is tightly

controlled by methanol induction and glucose catabolite repression. When Pichia pastoris is grown on methanol as a sole carbon source, alcohol oxidase (a peroxisomal packaged enzyme) constitutes up to 30%
5 of the total soluble protein.

Although high-level expression of heterologous proteins has been achieved using some methylotrophic yeast promoters, very little is known about the molecular mechanism involved in methanol induction.
10 Previous studies comparing the methanol regulated promoters did not reveal any significant regions of homology. Koutz et al., Yeast, 5:167-177 (1989); Sakai et al., Gene, 114:67-73 (1992); Ledebor et al., Nucl. Acids Res., 13:3063-3082 (1985). A possible
15 mechanism for the methanol induction of the methanol regulated gene is by means of a positive effector molecule that activates transcription by binding at specific DNA sequences, e.g., gene regulatory proteins. Thus, it is of interest to identify
20 nucleotide sequences that are conserved among methanol regulated genes so as to identify molecules involved in their expression and regulation, and to confer similar forms of regulation on heterologous genes.

Given the recognized utility of methylotrophic
25 yeast, as well as other yeast, in the expression of heterologous proteins, it is of interest to provide promoters and other regulatory nucleotide sequences for the controlled and/or high level expression of heterologous proteins in yeast. It is also of
30 interest to identify new methanol regulated genes so as to provide for the increased production of the proteins encoded by these methanol regulated genes and to provide for nucleotide sequences involved in the expression of methanol regulated genes.

The breakdown of starch is an important process in the brewing, baking, and sweetener industries. The degradation of the starchy cereal grain endosperm is initiated by α -amylase. In rice (Oryza sativa),
5 α -amylase is a monomeric Ca^{2+} -requiring metalloprotein which catalyzes an $\alpha(1-4)$ endoglycolytic cleavage of amylose and amylopectin. Since the yeast Saccharomyces cerevisiae which is commonly used in industry lacks the genes encoding for α -amylase, it is
10 unable to directly ferment starch. During the commercial production of alcoholic beverages the starch is initially pretreated with an exogenous source of α -amylase.

Although the expression of a wheat α -amylase in
15 yeast has been reported previously (Rothstein et al., Gene, 55: 353-356, 1987), the secretion of this protein across the cell membrane was minimal. Because the wheat α -amylase gene used in their study did not encode an enzyme containing an N-glycosylation site,
20 the low level of secretion may be due to the lack of N-glycosylation. The secretion of barley α -amylase in yeast had higher levels of expression (Sogaard et al., Gene, 94: 173-179, 1990). The highest yield was 2-3 $\mu\text{g/ml}$ of barley α -amylase AMY1 under the control of
25 the 3'-phosphoglycerate kinase promoter. It is interesting to note that the rice α -amylase is the only cereal amylase known to be glycosylated. Kumagai, et al., Gene, 94: 209-216 (1990) expressed α -amylase OS103 (O'Neill et al., Mol. Gen. Genet.,
30 221: 235-244, 1990) in a laboratory strain of Saccharomyces cerevisiae under the transcriptional control of the enolase promoter. In that study, during the biosynthesis of rice α -amylase by yeast, the plant signal peptide was removed, the protein was
35 N-glycosylated, and the active, enzyme was secreted

into the culture media. Since the laboratory strain of Saccharomyces cerevisiae (LL20) used in that study was deficient in the enzyme maltase the recombinant strain LL20 [pEno/1031 is unable to produce
5 significant amounts of ethanol. The reported yield of secreted enzyme was 1.8 µg/ml of rice α-amylase OS103.

Other amylolytic genes have been expressed in Saccharomyces cerevisiae cDNAs encoding mouse (Thomsen, Carlsberg Res. Commun., 48: 545-555, 1983;
10 Filho et al., Bio/Technology, 4: 311-315, 1986) and human salivary gland α-amylase (Sato et al., Gene, 50: 245-247, 1986), and Aspergillus awamori (Innis et al., Science, 228: 21-26, 1985) are some of the previous examples. Inlow et al., Biotechnology and
15 Bioengineering, 32: 227-234, (1988) were able to convert Maltrin 150 (Grain Processing Corporation, Mucatin, IA) to ethanol. Maltrin 150 is a soluble starch which is enzymatically hydrolyzed by α-amylase.

Alcohol oxidases have numerous commercial
20 applications. These uses include the measurement of alcohol levels in various biological and non-biological fluids, and the conversion of alcohol precursor molecules into aldehydes, e.g., for use in artificial flavor production.

25 Previous studies by Ellis et al., Mol. Cell. Biol., 5: 1111-1121 (1985) of the AOX1 promoter indicates that it is strongly repressed by ethanol and glucose. Based on Northern hybridizations, Pichia pastoris cells grown in the presence of ethanol did
30 not synthesize alcohol oxidase specific poly(A)⁺ RNA. In order to study the regulation of the AOX1 gene Tschopp et al., Nucl. Acids Res., 15: 3859-3876 (1987) took the AOX1 promoter and fused it to the E. coli lacZ gene. Saccharomyces cerevisiae strains harboring
35 an AOX1-lacZ fusion produced only small amounts of

active enzyme when grown in glucose or ethanol. See Stroman et al., U.S. Pat. No. 4,855,231. This result indicates that the AOX1 promoter would not be useful in the production of heterologous proteins during ethanol production in Saccharomyces cerevisiae.

The subject invention provides for alcohol oxidase promoters that can direct high level expression of rice α -amylase in the brewer's yeast Saccharomyces cerevisiae and in other yeast species.

10 The ZZA1 regulatory region contains several nucleotide sequences involved in promoter activity. These sequences include a TATAA box, located 45 bp upstream of the putative transcription initiation site. In order to find additional regulatory sites, the

15 nucleotide sequence of ZZA1 promoter was compared to AOX1, AOX2, MOX1, AOD1, and DHAS genes. The highly conserved regions between these genes may be involved in binding methanol specific transacting factors. A region in the core consensus sequence of ZZA1

20 (GCTTCCA, position -123) is identical to the sequence which is adjacent the S. cerevisiae regulatory activating protein (RAPL) binding site, as noted in Buchanan et al., Mol. Cell. Biol., 8:210225, Brindle et al., Mol. Cell. Biol., 10:4872-4885. Deletion of

25 the terminal G of the GCTTCCA sequence causes a loss in ENO1 gene expression. The CTTCC motif may also serve as a binding site for the trans-acting factor GCR1. This regulatory protein is required for high-level expression of several glycolytic genes in S.

30 cerevisiae Huie, et al. Mol. Cell. Biol., 12:26900-2700 (1992).

3. Summary of the Invention

An aspect of the invention is to provide isolated

35 nucleotide sequences encoding the alcohol oxidases

ZZA1 and ZZA2, and to provide the purified alcohol oxidases themselves. Another aspect of the subject invention is to provide for nucleotide sequences comprising the promoters and/or other portions of the ZZA1 and ZZA2 regulatory regions. Yet another aspect of the subject invention is to provide genetic constructions for the expression of heterologous genes using nucleotide sequences derived from the regulatory regions of ZZA1 and ZZA2. Another aspect of the invention is to provide host cells for the expression of heterologous proteins, in which the expression of the heterologous proteins is driven by promoters from the ZZA1 and ZZA2 regulatory regions. It is of particular interest to provide for S. cerevisiae cells expressing the heterologous protein rice α -amylase (gene QS 103) under the control of the ZZA1 promoter so as to be able to be grown in culture that can convert starch to ethanol.

Another aspect of the invention is to provide the nucleotide sequence (TTG-N3-GCTTCCAA-N5-TGGT) (SEQ ID NO:2) and the use of the nucleotide sequence. This nucleotide sequence is shown to be conserved and was found in the 5' flanking regions of alcohol oxidase, methanol oxidase, and dihydroxyacetone synthase genes in Pichia pastoris, Hansenula polymorpha, and Candida Boidinii S2.

It is also an aspect of the subject invention to provide for methods of isolating genes encoding proteins that modulate the expression of genes encoding proteins that modulate the expression of genes under the regulatory control of alcohol oxidase gene regulatory sequences, i.e., AOER (alcohol oxidase expression regulator) genes. Another aspect of the invention is to provide isolated AOER genes and purified AOER proteins.

One aspect of the subject invention is to provide methods for the conversion of starch into ethanol by culturing α -amylase secreting yeast cells in strach containing media.

5

4. Brief Description of the Drawings

FIG. 1 is a map of the plasmid YEP 352.

FIG. 2 is a plasmid map of the E. coli/ yeast shuttle vector 352AO. This clone contains the 5' regulatory region of the ZZA1 gene and a portion of the protein ZZA1 structural gene.

FIG. 3 is a map of the plasmid KS+ZZA2T.

FIG. 4 is a map of the plasmid KS- A07'8'. This clone contains the 5' regulatory region of the ZZA1 gene, the protein ZZA1 structural gene, and the 3' noncoding region of the ZZA1 gene.

FIG. 5 is a sequence comparison between the 3' noncoding region of the ZZA1, AOX1, and AOX2 genes.

FIG. 6 is a map of the plasmid SK- 7'8'.

FIG. 7 is a map of the plasmid KS+/AO.

FIG. 8 is a map of the plasmid KS-/AO.

FIG. 9 is a partial nucleotide sequence (SEQ ID NO:1) of alcohol oxidase genomic clone ZZA1. The amino acid sequence of the ZZA1 encoded Pichia pastoris alcohol oxidase was determined from the nucleotide sequence of ZZA1. The restriction endonuclease sites are in bold and the putative transcription start point (+1) is marked with an asterisk. The putative TATA box is underlined.

FIG. 10 is the nucleotide sequence comparisons of Pichia pastoris ZZA1 and AOX1 alcohol oxidase 5'-flanking regions. The nucleotide sequences are aligned to maximize sequence similarity. The transcription start point (+1) is marked with an

35

asterisk. Gaps, as indicated by (.), have been inserted to maximize sequence similarity.

FIG. 11 is the sequence alignment of the TTCCAA and pyrimidine boxes of 6 methanol regulated genes from methylotrophic yeasts. The distance from the transcriptional start site to the first T of the consensus sequence is indicated as a negative number. P = Pichia pastoris, H = Hansenula polymorpha, C = Candida boidinii S2.

10 FIG. 12 is a map of the plasmid 352H AO.

FIG. 13 is a map of the α -amylase expression vector, pBGC103. This plasmid contains the Pichia pastoris ZZA1 promoter, the rice α -amylase CDNA pOS103, Saccharomyces cerevisiae 2 μ m ori, 15 Saccharomyces cerevisiae URA3, and part of the pUC18 plasmid.

FIG. 14 is a map of the plasmid pBGC103 BstXIL.

FIG. 15 is a map of the plasmid pBGC104.

FIG. 16 is a map of the plasmid 351 HAO103.

20 FIG. 17 is a western blot detecting the expression and secretion of α -amylase from yeast cells grown in 5mM CaCl₂, YEP plus various carbon sources (methanol, ethanol, glucose, or no carbon source). Samples were removed from the culture 0, 2, 4, and 6, 25 hours after inoculation.

FIG. 18 is the nucleotide sequence comparisons of Pichia pastoris ZZA2, ZZA1, AOX1, and AOX2 alcohol oxidase 3'-terminal regions. The nucleotide sequences are aligned to maximize sequence similarity. 30 The transcription start point (+1) is marked with an asterisk. Gaps, as indicated by (.), have been inserted to maximize sequence similarity.

5. Detailed Description of the Invention

5.1 Definitions

The terms "ZZA1" and "ZZA2" as used herein refers to two alcohol oxidases. The genes encoding ZZA1 and ZZA2 are referred to by the terms "ZZA1" and "ZZA2", respectively. ZZA1 has a 5' non-coding region nucleotide sequence and N-terminal amino acid sequence as indicated in Figure 9 (SEQ ID NO:1). The ZZA1 is at least partially encoded on plasmid pBGC104 (Figure 15). ZZA2 may be identified as the gene at least partially encoded by the nucleotide sequence that is the insert in plasmid KS+ZZA2T, i.e., the amplification product of a PCR reaction performed on genomic DNA from P. pastoris, (obtained from Dr. H. Phaff, University of California, Davis (University ID # 72-1033)) using 24-mer oligonucleotide 5' TCG ACC CAG GTT TCA TGA ACG ATG 3' (SEQ ID NO:4) and the 31-mer oligonucleotide (5' TCC TGC AGC AAC CAA TGA GGA GAA TGA CAA C 3' (SEQ ID NO:5) as primers .

The term "gene", as used herein, refers not only to the nucleotide sequence encoding a specific protein, but also to any adjacent 5' and 3' noncoding nucleotide sequence involved in the regulation of expression of the protein encoded by the gene of interest. These non-coding sequences include terminator sequences, promoter sequences, upstream activator sequences, regulatory protein binding sequences, and the like. These non-coding sequence gene regions may be readily identified by comparison with previously identified eukaryotic non-coding sequence gene regions. Furthermore, the person of average skill in the art of molecular biology is able to identify the nucleotide sequences forming the non-coding regions of a gene using well-known techniques

such as a site-directed mutagenesis, sequential deletion, promoter probe vectors, and the like.

The term "regulatory region" or "regulatory sequence" as used herein in reference to a specific gene refers to the non-coding nucleotide sequences within that gene that are necessary or sufficient to provide for the regulated expression of the coding region of a gene. Thus the term regulatory region includes promoter sequences, regulatory protein binding sites, upstream activator sequences, and the like. Specific nucleotides within a regulatory region may serve multiple functions. For example, a specific nucleotide may be part of a promoter and participate in the binding of a transcriptional activator protein.

The term "coding region" as used herein refers to that portion of a gene which codes for a protein. The term "non-coding region" refers to that portion of a gene that is not a coding region.

The term "alcohol oxidase", as used herein refers to an enzyme able to catalyze the reaction $RCH_2OH + O_2 = RCHO + H_2O_2$ where R is hydrogen or a lower alkyl, generally selected from the group H-, CH_3 -, CH_3CH_2 -, and $CH_3(CH_2)_2$ -.

The term "operably linked" refers to a juxtaposition of components, particularly nucleotide sequences, such that the normal function of the components can be performed. Thus, a coding sequence "operably linked" to regulatory sequences refers to a configuration of nucleotide sequences wherein the coding sequences can be expressed under the regulatory control i.e., transcriptional and/or translational control, of the regulatory sequences.

By "purified" it is meant, when referring to a peptide or nucleotide sequence, that the indicated molecule is present in the substantial absence of

other biological macromolecular, e.g., polypeptides, polynucleic acids, and the like of the same type. The term "purified" as used herein preferably means at least 95% by weight, more preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above.

The term "isolated" as used herein refers to a polypeptide, polynucleotide molecules separated not only from other peptides, DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule but also from other macromolecules and preferably refers to a macromolecule found in the presence of (if anything) only a solvent, buffer, ion or other component normally present in a solution of the same. "Isolated" and "purified" do not encompass either natural materials in their native state or natural materials that have been separated into components (e.g., in an acrylamide gel) but not obtained either as pure substances or as solutions.

In representations of nucleotide sequences provided herein, the following conventions are used. N refers to a nucleotide sequence location that may be of any nucleotide base. Y refers to a nucleotide base that is a pyrimidine.

Salts of any of the macromolecules described herein, i.e., proteins and polynucleotide sequences, will naturally occur when such molecules are present in (or isolated from) aqueous solutions of various pHs. All salts of peptides and other macromolecules having the indicated biological activity are considered to be within the scope of the present

invention. Examples include alkali, alkaline earth, and other metal salts of carboxylic acid residues, acid addition salts (e.g., HCl) of amino residues, and zwitterions formed by reactions between carboxylic acid and amino residues within the same molecule.

The degree of similarity between the nucleic acid sequences of two polynucleotides may be measured by determining whether the two polynucleotide sequences can hybridize to each other under a given set of conditions so as to form hybrid heterodimers between the two sequences. These hybridization conditions may be varied so that the hybridization interaction between the two polynucleotide sequences occurs at a certain number of degrees centigrade below the melting temperature, T_m , of the duplex polynucleotide molecule used as the hybridization probe. T_m is defined as the temperature at which half the duplex molecules have dissociated into their constituent single strands. The T_m for a given double stranded polynucleotide may be determined empirically or by reference to well-known formulas that take into account hybridization condition factors that influence T_m (including the types of nucleic acids used), e.g. for DNA-DNA hybridization probes longer than 50 nucleotides, $T_m = 81.5^\circ\text{C} + 16.6 \log M + 41 (\text{mole fraction G} + \text{C}) - 500/L - 0.62 (\% \text{ formamide})$ (M is the molar monovalent cation concentration), Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Volume 152, Academic Press, San Diego, CA (1987). The stringency of a hybridization may be defined as degrees centigrade below the T_m at nucleotide sequence for comparison. The degree of stringency of hybridization is said to decrease as hybridization takes place at a temperature increasingly below the T_m of the hybridization probe. Maximum stringency typically

occurs at about $T_m - 5^\circ\text{C}$, i.e., at a temperature 5°C below the T_m of the hybridization probe. "High stringency,, hybridization is said to take place at a temperature of about $5^\circ - 10^\circ\text{C}$ below T_m . "Intermediate stringency" hybridization is said to take place at a temperature of about $10^\circ - 20^\circ\text{C}$ below T_m . "Low stringency hybridization", i.e., maximum hybridization, is said to take place at a temperature of about $20^\circ - 25^\circ\text{C}$ below T_m . Two polynucleic acid sequences that can hybridize to one another under high stringency conditions can also hybridize to one another under low stringency conditions. Two polynucleic acid sequences that can hybridize to one another under low stringency conditions cannot necessarily hybridize to one another under low stringency conditions.

5. Detailed Description of the Invention

The present invention provides for isolated nucleotide sequences encoding alcohol oxidase proteins ZZA1 and ZZA2 and provides for the isolated alcohol oxidases themselves. The present invention also provides for the non-coding nucleotide sequences that are part of the ZZA1 and ZZA2 genes and various genetic constructions employing these non-coding nucleotide sequences to direct the expression of heterologous proteins. Also provided for are methods of expressing heterologous proteins using the ZZA1 and ZZA2 regulatory sequences and cells for the expression of the heterologous proteins. The present invention also provides for methods of isolating genes encoding proteins that modulate the expression of genes under the regulatory control of alcohol oxidase gene regulatory sequences (AOER genes), AOER genes themselves, and proteins encoded by AOER proteins.

Specifically described herein is the discovery of two previously unknown alcohol oxidases present in a naturally occurring P. pastoris strain (obtained from Dr. H. Phaff, University of California, Davis

5 (University ID # 72-1033)), ZZA1 and ZZA2, and the isolation of the genes encoding these newly discovered alcohol oxidases. The 5' non-coding region sequence of the ZZA1 gene is extensively analyzed and shown to possess regions of significant homology to the 5' non-

10 coding regions of other methanol regulated genes isolated from methylotrophic yeast. The 5'-non-coding homology provided herein reveals a hitherto undescribed conserved sequence (TTGNNNGCTTCCANNNNTGGT) (SEQ ID NO:2) present in the

15 51 regulatory region of methanol regulated genes present in methylotrophic yeast. The ZZA1 regulatory regions containing the conserved sequence are demonstrated to confer catabolite repression on the expression of heterologous proteins. The ZZA1

20 regulatory regions containing the conserved sequence confers heterologous gene expression in the presence of ethanol; i.e., this promoter is not ethanol repressed. Adding additional copies of the conserved sequence TTCAA to the 5' non-coding region of a

25 heterologous gene is demonstrated to increase the expression levels of that gene.

The subject invention provides for two novel alcohol oxidases, ZZA1 and ZZA2. ZZA1 and ZZA2 can be seen to be distinct from the previously identified P.

30 pastoris alcohol oxidases AOX1 and AOX2 because the primers used to isolate ZZA1 and ZZA2 would not be expected to amplify AOX1 and AOX2. Additionally, nucleotide sequence comparisons of ZZA1 and ZZA2 with AOX1 and AOX2 reveal that the genes (and consequently

35 the proteins) are distinct from one another. The non-

coding regions of the AOX1 and AOX2 genes have a different nucleotide sequence from corresponding regions of the ZZA1 gene. Additionally, partial sequence analysis of the predicted N-terminus of ZZA1 indicates that the protein is distinct from AOX1 and AOX2. The ZZA2 gene was obtained by performing PCR amplification on P. pastoris DNA using the primers (5' TCG ACC CAG GTT TCA TGA ACG ATG 3') (SEQ ID NO:4) and (5'TCC TGC AGC AAC CAA TGA GGA GAA TGA CAA C 3') (SEQ ID NO:5). The 910 bp fragment obtained from the PCR amplification was not expected based on the available sequence information of the AOX1, AOX2, or ZZA1 genes.

Although the complete nucleotide sequences of ZZA1 and ZZA2 are not explicitly provided herein, sufficient information is provided so as to enable the person of average skill in the art of molecular biology to obtain the complete nucleotide sequences by the use of routine molecular biology techniques. Similarly, the complete amino acid sequences of ZZA1 and ZZA2 are not explicitly provided; however, the information provided herein enables a person of average skill in the art to obtain the complete nucleotide sequence without undue experimentation. In other words ZZA1 and ZZA2 are described in sufficient detail, i.e. the PCR primers used for isolation and the nucleotide sequences of Figures 5, 9, and 10, to obtain the complete nucleotide sequence (and predicted encoded amino acid sequence) by applying routine molecular biology techniques, such as those described in Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook et al., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), e.g, the nucleotide sequence of cloned fragments may be completed by routine sequencing, and incompletely cloned sequences may be

used as hybridization probes to obtain the remainder of the sequence.

Although ZZA1 and ZZA2 possess unique primary amino acid sequences, the subject invention also provides for numerous ZZA1 and ZZA2 "variants." These "variants" are proteins that have alcohol oxidase activity and possess primary amino acid sequences similar (but not identical), i.e., homologous, to the primary amino of ZZA1 and/or ZZA2. These ZZA1 and ZZA2 "variants" may have one or more amino acid substitutions. Preferably these substitutions are the result of the substitution of one amino acid with another amino acid that has a similar chemical structure, i.e., a conservative substitution. Examples of conservative substitution are the replacement of a leucine with an isoleucine or a valine, an aspartate with a glutamate, and a threonine with a serine. Furthermore, the ZZA1 and ZZA2 variants provided for include proteins with the primary amino acid sequence of ZZA1 and ZZA2, but possessing various minor amino acid deletions and/or insertions, typically in the range of about 1-5 amino acids, as well as one or more amino acid substitutions. Other ZZA1 and ZZA2 variants provided for include chimeric proteins produced by adding heterologous signal sequences.

By providing for the amino acid sequence of hitherto unknown alcohol oxidases, the subject invention greatly enhances the level of guidance available to the person of average skill in the art of molecular biology attempting to make biologically active variants of alcohol oxidases, including ZZA1 and ZZA2. The level of guidance is enhanced because knowledge of regions of homology between proteins of similar biological activity indicate where amino acid

sequence changes can be made without abolishing biological activity, i.e., at the most variable regions.

In addition to providing for the nucleotide
5 sequences of genes ZZA1 and ZZA2. the subject
invention also provides for numerous nucleotide
sequences bearing homology to nucleotide sequences of
ZZA1 and ZZA2. Sequences of interest bearing homology
to the nucleotide sequences encoding ZZA1 and ZZA2
10 include nucleotide sequences encoding ZZA1 and ZZA2
variants.

It will be appreciated by those skilled in the
art that as a result of the degeneracy of the genetic
code, a multitude of ZZA1 and ZZA2 encoding nucleotide
15 sequences, some bearing minimal nucleotide sequence
homology to the nucleotide sequences of ZZA1 and ZZA2,
may be produced. The subject invention has
specifically contemplated each and every possible
variation of peptide or nucleotide sequence that could
20 be made by selecting combinations based on the
possible amino acid and codon choices made in
accordance with the standard triplet genetic code as
applied to the sequence of ZZA1 and ZZA2 and all such
variations are to be considered as being specifically
25 disclosed.

Although nucleotide sequences encoding ZZA1 and
ZZA2 and variants thereof are preferably capable of
hybridizing to the nucleotide sequences of ZZA1 and
ZZA2, respectively, under stringent conditions, it may
30 be advantageous to produce nucleotide sequences
encoding ZZA1 and ZZA2 (or ZZA1 and ZZA2 variants)
possessing a substantially different coding sequences.
Codons can be selected for use in a particular
expression host organism in accordance with the
35 frequency with which a particular codon is utilized by

the host to increase the rate at which expression of the peptide occurs. Other reasons for substantially altering the nucleotide sequence encoding a protein without altering the amino acid sequence include the
5 production of RNA transcripts having more desirable properties, e.g., greater half-life, than transcripts produced from the sequence ZZA1 and ZZA2 and the like.

By providing for the nucleotide sequences of ZZA1 and ZZA2, the present invention enables numerous
10 genetic manipulations of yeast, e.g., chromosomal insertions, gene inactivations, and the like, by employing well known yeast genetics techniques. These techniques may be applied to numerous yeast species, including, S. cerevisiae and P. pastoris. Description
15 of suitable yeast genetic manipulation techniques can be found, among other places, in Guthrie, C., and Fink, G.R. Methods in Enzymology, Volume 194: Yeast Genetics and Molecular Biology, Academic Press, Inc., San Diego, CA. (1991); Ito, H., Fukuda, Y., Murata,
20 K., and Kimura, A. J. Bacterial. 153:163-168 (1983); Rodriguez, R.L., and Tait, R.C. Recombinant DNA Techniques: An introduction, Addison-Wesley, Reading, Mass. (1983); and Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and
25 Struhl, K. Current protocols in Molecular Biology. Volume II, Greene Publishing Associates and Wiley-Interscience, New York, NY.

Nucleotide sequences of interest for joining to ZZA1, ZZA2, or fragments thereof, include numerous
30 widely available cloning vectors, e.g., plasmids, cosmids, A phage derivatives, phasmids, and the like, that are in the public domain. Vectors of interest include expression vectors, replication vectors, hybridization probe generation vectors, sequencing
35 vectors, and the like. In general, vectors of

interest may contain an origin of replication functional in at least one organism, convenient restriction endonuclease digestion sites, and selectable markers for the host cell. Nucleotide sequences encoding ZZA1 and ZZA2 may be joined to a variety of other nucleotide sequences of interest by means of well established recombinant DNA techniques (see, for example, Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor (1989)).

Expression vectors may be used to produce useful quantities of ZZA1 and ZZA2 in a variety of cell hosts. Detailed descriptions of many expression vectors and their use can be found, for example in Goeddel, Gene Expression Technology Methods in Enzymology, Vol. 185 Academic Press, San Diego, CA (1990). Expression vectors contain promoters functional in the host of interest. The promoter may be operably linked to the coding sequence of a gene of interest so as to produce a translatable mRNA transcript encoding ZZA1 and ZZA2. Expression vectors preferably have convenient restriction sites located near the promoter sequence so as to provide for the insertion of nucleic acid sequences encoding heterologous proteins. The promoters in suitable expression vectors may be either constitutive or inducible. In addition to having promoter sequences, expression vectors may contain various enhancer sequences and the like, included for the purpose of maximizing expression of ZZA1 and ZZA2.

ZZA1 and ZZA2 may be purified from a variety of cells. Suitable cell sources for the production of purified ZZA1 or ZZA2 include cells naturally producing these proteins, cells not naturally encoding ZZA1 or ZZA2, but genetically modified to do so, and

cells naturally producing ZZA1 or ZZA2, but genetically modified so as to produce elevated levels of these proteins. Cells from which ZZA1 and ZZA2 may be isolated include both prokaryotic and eukaryotic
5 cells. Preferred cellular sources for the isolation of variants are yeast cells genetically modified to overproduce ZZA1 or ZZA2, and naturally occurring strains of Pichia Dastoris.

It will be appreciated that an advantage of the
10 subject invention is to apply recombinant DNA techniques so as to provide for cellular lysates that contain ZZA1 and/or ZZA2 in significantly higher, at least 2-fold, preferably at least 10-fold, higher concentrations than found in naturally occurring cells
15 that have not been genetically modified. Since ZZA1 and ZZA2 variants are not naturally produced, it is apparent that cells from which these variants can be isolated do not naturally encode the variants, but are genetically modified to do so.

20 Numerous methods are known for purifying enzymes with alcohol oxidase activity. In general, routine variations of known methods for purifying alcohol oxidases may be used to purify ZZA1 and ZZA2. Methods for purifying alcohol oxidases have been described.
25 Descriptions of these methods can be found, among other places, in U.S. patent 4,619,898, and Giuseppin et al., Appl. Microbial. Biotechnol. 28:14-19 (1988)

The 5' non-coding region of ZZA1 contains nucleotide sequences known to be involved in the
30 transcription and translation of eukaryotic genes (figure 9) (SEQ ID NO:1). At position -45, the sequence TATAAA formed a putative Goldberg-Hogness or TATA box The nucleotides flanking the first methionine codon (ATG, position +112) conformed to a preferred
35 translation initiation sequence (ACCATGG) associated

with many highly expressed eukaryotic mRNAs and the sequence CAAAAACAA (position +87-95) resembled the CA-rich nucleotides found 10-30 bp upstream of the translation initiation site in many S. cerevisiae genes. Two sequences TGGTTTG (position -631) and GGTTTG (position -361) were very similar to the animal core enhancer, GTGG^{TTT}_{AAA}G. (Laimens et al., pgs. 28-3, in Enhancers and Eukaryotic Gene Expression: Current Communications in Molecular Biology, Gluzman and Shenk eds., Cold Spring Harbor Press, Cold Spring Harbor, New York (1983).

The nucleotide sequences of ZZA1 and AOX1 5' non-coding regions were compared, and the homology between the promoters was 64% (Fig. 10). Primer extension assays (see example 10) indicated that the transcription initiation site for ZZA1 is located 111 bp upstream of the start codon. This transcription initiation site is present in a highly conserved region, i.e., 39/40 nucleotides are identical between ZZA1 and AOX1.

Previous nucleotide sequence comparisons of methanol-regulated promoters from Pichia pastoris and Hansenula polymorpha failed to reveal any important regions of homology, see Ellis et al., Mol. Cell. Biol., 5:1111-1121, Koutz et al, Yeast, 5:167-177, Ledebor et al., Nucl. Acids Res., 13:3063-3082. The nucleotide sequence analysis provided herein identifies several conserved regions in methanol regulated genes (fig. 11). A core motif consisting of the sequence TTCCAA occurred in three conserved regions in ZZA1 (-316, -120, +35). The TTCCAA core consensus sequence is surrounded by additional regions of homology that are spaced approximately ten base pairs apart, i.e., (TTGNNNNNTTCCAANNNTGGT (SEQ ID NO:6) and AATTNNNNNTGGTTCCANNNNNNA (SEQ ID NO:7)). A

pyrimidine box containing the conserved sequence CCYCTTTTG is present in ZZA1 at position -163. The TTCCAA and the pyrimidine boxes are present in all of the methanol-regulated genes analyzed.

- 5 There are several unique features in the ZZA1 51 upstream region which might be involved in controlling gene expression. The upstream region has two sequences TGGTTTG (position -631) and GGTTTG (position -361) which are very similar to the animal core
- 10 enhancer, GTGG^{TTT}_{AAA}G and a putative alcohol dehydrogenase regulatory protein (ADR1) binding site GGAGA in a 30 bp imperfect palindrome GATGGATTTCAGGAGAAATTGTTCTGCCATC (SEQ ID NO:8) (-787). The ADR1 protein activates
- 15 transcription of the catabolite repressed alcohol dehydrogenase 2 gene (ADH2) and acts as a positive regulator for the peroxisomal catalase A gene. The ZZA1 gene contains an upstream repression site (URS) GTTTCCTCAAGGCAAGAACTCC (SEQ ID NO:3) (position -712) which shares homology to a URS in ENOI
- 20 (GATTCCTCAAGGTATGCCTCTCC) (SEQ ID NO:9). Deletion of this site in ENO1 resulted in a 10-fold increase in gene expression in cells grown on glucose, as noted in Cohen et al., Mol. Cell. Biol., 7:2753-2761. Similarly, deleting the ZZA1 URS from the ZZA1
- 25 regulatory region increases the expression of genes under the control of the ZZA1 regulatory region and removes glucose repression control.

The subject invention provides for a novel conserved nucleotide sequence,

- 30 TTGNNNGCTTCCAANNNNTGGT (SEQ ID No: 2), that may be used to increase the expression of genes of interest in a variety of cells. The conserved nucleotide sequence TTGNNNGCTTCCAANNNNTGGT (SEQ ID No: 2) is present in the 51 non-coding region of at least 6
- 35 methanol regulated genes from methylotrophic yeast.

The conserved nucleotide sequence contains a consensus core sequence 5'-TTCCAA-3'. The addition of one or more copies of the TTCCAA consensus core sequence at locations proximate to the natural location of TTG-N3-
5 GCTTCCAA-N5-TGGT, i.e., TTGNNNGCTTCCAANNNNTGGT (SEQ ID NO:2) sequence, with respect to the transcription initiation site, may be used to increase the expression level of genes under the regulatory control of ZZA1 or ZZA2 regulatory sequences. The
10 sequence TTGNNNGCTTCCAANNNNTGGT (SEQ ID NO:2) may also be used to confer some of the regulatory behavior of methanol regulated genes on heterologous proteins for expression.

The regulatory regions of ZZA1, ZZA2, or portions
15 thereof, may be operably linked to nucleotide sequences encoding heterologous protein i.e., heterologous with respect to the regulatory region, so as to place, at least in part, the expression of the heterologous protein under the regulation of the ZZA1
20 or ZZA2 regulatory sequences. The expression of a heterologous gene of interest may be driven by a promoter located within the ZZA1 or ZZA2 regulatory sequence. Also, by including a sufficient portion of ZZA1 or ZZA2 regulatory region nucleotide sequence in
25 genetic constructions comprising the regulatory region, the expression heterologous genes of interest may be regulated similarly to the regulation of ZZA1, ZZA2, or other alcohol oxidase genes. Many heterologous proteins of interest may be expressed
30 from nucleotide sequences under the regulatory control of ZZA1 or ZZA2 regulatory regions. These heterologous proteins include enzymes, hormones, short peptide, lymphokines, regulatory proteins, structural proteins, antigens (for use in inducing a specific
35 immune response), and the like. Additionally, ZZA1

and ZZA2 regulatory regions may be used to control the transcription anti-sense RNA.

The promoter sequences within the ZZA1 and ZZA2 regulatory regions may be used to provide for the
5 expression of heterologous and homologous proteins in a variety of cells, preferably eukaryotic cells. ZZA1 and ZZA2 promoters may be used to drive expression of genes in many cells from many species of yeast, including yeast species that are not methylotrophic,
10 such as Saccharomyces and Kluyveromyces species. The high level of expression exhibited by the ZZA1 promoter in S. cerevisiae is of particular interest, given the substantial differences between S. cerevisiae cells and P. pastoris cells.

15 The subject invention provides for alcohol oxidase promoters that can direct high level expression of rice α -amylase in the brewer's yeast Saccharomyces cerevisiae and in other yeast species. The ZZA1 regulatory region contains several nucleotide
20 sequences involved in promoter activity. These sequences include a TATAA box, located 45 bp upstream of the putative transcription initiation site. In order to find additional regulatory sites, the nucleotide sequence of ZZA1 promoter was compared to
25 AOX1, AOX2, MOX1, AOD1, and DHAS genes. The highly conserved regions between these genes may be involved in binding methanol specific transacting factors. A region in the core consensus sequence of ZZA1 (GCTTCCA, position -123) is identical to the sequence
30 which is adjacent the S. cerevisiae regulatory activating protein (RAPL) binding site, as noted in Buchanan et al., Mol. Cell. Biol., 8:2102-25, Brindle et al., Mol. Cell. Biol., 10:4872-4885. Deletion of the terminal G of the GCTTCCA sequence causes a loss
35 in ENO1 gene expression. The CTTCC motif may also

serve as a binding site for the transacting factor GCR1. This regulatory protein is required for high-level expression of several glycolytic genes in S. cerevisiae Huie, et al. Mol. Cell. Biol., 12:26900-
5 2700 (1992).

Both AOX1 and AOX2 show differences in the levels of alcohol oxidase mRNA accumulation during methanol induction, Cregg et al., Mol. Cell. Biol., 9:1316-1323. It is possible that differences in nucleotide sequence
10 near the TTCAA core consensus sequences might result in different binding efficiencies of positive regulatory proteins. Only AOX2 has a putative HAP2/HAP3 binding site (TGGTTGGT) next to the TTCAA core consensus sequence, Olesen et al., Cell, 51:953-
15 961. The TGGTTGGT sequence acts as a upstream activator of the CYC1 gene, Hahn and Guarente, Science, 240:317-321. Since the heme induced HAP2/HAP3 protein binds both CCAAT-boxes and TGGTTGGT elements, the AOX2 promoter might be regulated by heme
20 (the relevant AOX2 sequences are TTCCAATTGGTTGGT (SEQ ID NO:10)). In Saccharomyces cerevisiae heme regulates the expression of catalase A, a peroxisomal enzyme.

Another aspect of the subject invention is to
25 provide for methods of isolating genes encoding proteins that modulate the expression of genes under the regulatory control of alcohol oxidase gene regulatory regions. These genes under the regulatory control of alcohol oxidase gene regulatory regions are
30 referred to as AOER (alcohol oxidase expression regulator) genes. The gene product of an AOER gene is referred to as an AOER protein. The subject invention provides for isolated AOER genes and purified AOER proteins.

The methods of isolating AOER genes provided for are based upon assays that detect interactions between AOER proteins and AOER protein binding sequences. Nucleotide sequences containing putative binding sites
5 for AOER proteins to be used in AOER protein and gene isolation procedures may be derived from the regulatory region of ZZA1 or ZZA1, and preferably contain the conserved sequence TTGNNNGCTTCCAANNNTGGT (SEQ ID NO:2). Numerous assays to detect interactions
10 between gene expression regulatory proteins and target nucleotide binding sequences are known to the person of average skill in the art of molecular biology. These techniques are described, among other places, in Molecular CloningA Laboratory Manual, 2nd Ed.,
15 Sambrook et al., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), Guide to Molecular Cloning Technique: Methods in Enzymology Volume 152, Berger and Kimmel, Academic Press, San Diego, CA (1990), and the like. Providing the information that a specific
20 nucleotide sequence is a binding site for a protein that regulates gene expression enables the use of numerous well known methods for purifying proteins that bind to the specific nucleotide sequence and also enables the use of methods for isolating the genes
25 encoding such gene regulatory proteins. These protein purification and gene isolation techniques may be conveniently applied to AOER proteins and genes encoding AOER proteins. Examples of such methods include the construction of fusions between putative
30 regulatory protein binding nucleotide sequences and "reporter" genes encoding proteins conferring an easily detected phenotype, e.g., 9-galactosidase, α -amylase, and the like. When a genetic library containing an AOER encoding nucleotide sequence is
35 moved into host cells containing reporter gene

fusions, host cells containing library clones expressing proteins that modulate, i.e., increase or decrease, the expression of the reporter genes may readily be detected by means appropriate for the given

5 reporter gene and host cell, e.g., blue colony color for S. cerevisiae cells containing a Lac Z. reporter and grown on X-gal containing media, halo formation in iodine stained starch media for S. cerevisiae containing an α -amylase reporter. Genetic libraries

10 to be screened by the above-described screening assay are preferably prepared from nucleic acids isolated from an organism naturally encoding alcohol oxidase genes, more preferably a methylotrophic yeast, and even more preferably from Pichia pastoris.

15 AOER proteins may be purified using assays and/or separation techniques based on the ability of the proteins to specifically bind to nucleotide sequences within the regulatory regions of ZZA1 and ZZA2 genes. These techniques include gel mobility shift assays

20 such as those described in Protein Function: A Practical Approach, Ed. Creighton, IRL Press, NY, NY (1990). Other approaches for purifying AOER proteins include the technique of affinity chromatography using nucleotide sequences containing AOER binding sites.

25 Affinity chromatography of DNA binding proteins is described, among other places, in the book Guide to Protein Purification, Methods in Enzymology Vol. 182, ed. Deutscher, Academic Press, San Diego, CA (1990).

Many well known cloning techniques may be used to

30 isolate an AOER gene after AOER protein has been purified. The well known cloning techniques including, hybridization with oligonucleotide probes prepared on the basis of N-terminal sequence analysis of proteins, screening of expression libraries with

35 antisera prepared against the protein of interest, PCR

amplification using primers based on amino acid sequence analysis, and the like.

The invention having been described above, may be better understood by reference to the following examples. These examples are offered for the purpose of illustrating the subject invention, and should not be interpreted as limiting the invention.

EXAMPLES

10

Example 1

Pichia pastoris chromosomal DNA preparation

Yeast cells were grown in 30 mls of YEPD at 30°C for 16 hours. The cells were pelleted using a table top centrifuge at 3,000 rpm for 5 minutes. The pellet was washed once in double distilled H₂O, once in 1M sorbitol, once in SED (SED is 1 M Sorbitol, 25 mM EDTA, 50 mM DTT), and then resuspended in 5 ml of 0.1 M tris-HCL, pH 7.0, and 1 M sorbitol. The cells were mixed with 400 µl Zymolyase 100,000 (Seikagaku), 10 Al of beta-mercaptoethanol and incubated at 30°C for 1 hour. The resulting spheroplasts were then centrifuged at 1,000 g for 10 minutes and gently lysed with 4 ml of lysis buffer (0.1% SDS, 10 mM Tris-HCL, pH 7.4, 5 mM EDTA, 50 mM NaCl). 100 µl of Pronase (5 mg/ml, Boehringer Mannheim) and 100 ml RNase A (10 mg/ml, Sigma) were each added and the solution was incubated at 37°C for 2 hours. The DNA was then extracted in 10 mls of chloroform containing isoamyl alcohol (24:1 v/v), and the phases were separated by centrifugation at 9,000 rpm for 10 minutes. The upper (aqueous) phase was transferred to a fresh tube and one volume of phenol/chloroform/isoamyl alcohol was added. The phases were separated by centrifugation (9,000 rpm for 15 minutes) and the upper (aqueous)

phase was placed in a fresh tube. Two volumes of cold ethanol was added and the chromosomal DNA were spooled out using a glass rod. The DNA was then rinsed in 70% ethanol and immediately dissolved in 500 μ l of TE
5 buffer.

Example 2

Southern hybridization with *Pichia Pastoris* genomic DNA Total chromosomal DNA was isolated as
10 described in Example 1. 6.6 μ g of total DNA was digested with the restriction endonucleases BamHI, PstI, EcoRI, and BglII and the DNA fragments were separated by 1% agarose gel electrophoresis. The DNA was then denatured, the gel was neutralized, and the
15 DNA was transferred to nitrocellulose. The filters were then prehybridized at 42°C for 16 hours in 5X SSPE, 5X Denhart's, 0.4 mg/ml denatured calf thymus DNA, 0.5% SDS. Hybridization was carried out at 42°C for 16 hours in 5X SSPE, 5X Denhart's, 0.2 mg/ml
20 denatured calf thymus DNA, 0.5% SDS. The ³²P-labeled probe was prepared by annealing a 45-mer oligonucleotide
(5'CACCACCTAGAACTAGGATATCAAACCTCTTCGGGGATAGCCATCG3')
(SEQ ID NO:11) to a 18-mer oligonucleotide
25 (5'CGATGGCTATCCCCGAAG 3') (SEQ ID:12) and filling in the 27 base overhang with DATP, DGTP, dTTP, and (α -³²P) dCTP. The labelled mixture was chromatographed on a Sephadex G-25 column. The labelled DNA fractions were pooled and boiled for 5 minutes and immediately added
30 to the hybridization buffer. Hybridization was carried out at 42°C for 18 hours. After hybridization the filters were washed twice for 15 minutes at room temperature in 5X SSPE, 0.2% SDS, and one time at 42°C in 1X SSPE, 0.2% SDS. The filters were then
35 autoradiographed for 14 days.

Example 3

Construction of a partial *Pichia pastoris* genomic library.

Two micrograms of *Pichia pastoris* genomic DNA
5 prepared as described in Example 1 were digested with
PstI and BamHI. After fractionation by
electrophoresis on a 1% low melt agarose gel, DNA
fragments of approximately 1100 base pairs were
isolated and subcloned into YEp352. This *Pichia*
10 *pastoris* library was maintained in *Escherichia coli*
strain C600. The transformed cells were plated on LB
plates containing 100 µg/ml ampicillin. The plates
were incubated at 37°C for 24 hours and stored at
4° C. Approximately 500 independent transformants
15 were transferred onto nitrocellulose and incubated for
12 hours on LB containing 200 µg/ml chloramphenicol.
The colonies were then gently lysed by placing the
filters on Whatman paper soaked in 1.5 M NaCl and 0.5
M NaOH for 5 minutes. The filters were then
20 neutralized by transferring them to Whatman paper
soaked in 1 M Tris-HCL, pH 7.4 for 5 minutes and the
transferred into holding buffer (1.5 M NaCl, 0.5 M
Tris-HCL, pH 7.4). The filters were then dried at
650°C for 2 hours.

25

Example 4

Colony hybridization

The nitrocellulose filters containing the partial
genomic library prepared as described in Example 3
30 were prehybridized at 42°C for 16 hours in 5X SSPE,
5X Denhart's, 0.4 mg/ml denatured calf thymus DNA,
0.5% SDS. Hybridization were carried out at 42°C for
18 hours in 5X SSPE, 5X Denhart's, 0.4 mg/ml denatured
calf thymus DNA, 0.5% SDS. The ³²P-labeled probe was
35 prepared by annealing a 45-mer oligonucleotide

(5'CACCACCTAGAACTAGGATATCAAACCTCTTCGGGGATAGCCATCG3')
(SEQ ID NO:11) to a 18-mer oligonucleotide
(5'CGATGGCTATCCCCGAAG 3') (SEQ ID NO:12) and filling
in the 27 base overhang with dATP, dGTP, dTTP, and (α -
5 ³²P) dCTP. The labelled mixture was chromatographed on
a Sephadex G-25 column. The labelled DNA fractions
were pooled and boiled for 5 minutes and immediately
added to the hybridization buffer. After
hybridization the filters were washed twice for 15
10 minutes at room temperature, and one time at 42°C.
The filters were then autoradiographed for 3 days.

Example 5

Isolation of Alcohol Oxidase ZZA1 promoter

15 Two micrograms of Pichia pastoris genomic DNA
were digested with PstI and BamHI. After
fractionation by electrophoresis on a 1% low melt
agarose gel DNA fragments of approximately 1100 base
pairs were isolated and subcloned into YEp352 (Fig.1)
20 This Pichia pastoris library was maintained in
Escherichia coli strain C600. Approximately 500
independent transformants were screened by colony
hybridization using a ³²P-labeled probe prepared by
annealing a 45-mer oligonucleotide (5'
25 CACCACCTAGAACTAGGATATCAAACCTCTTCGGGGATAGCCATCG 3') (SEQ
ID NO:11) to a 18-mer oligonucleotide
(5'CGATGGCTATCCCCGAAG 3') (SEQ ID NO:12) and filling in
the 27 base overhang with dATP, dGTP, dTTP, and (α -³²P)
DCTP. Hybridizations were carried out at 42°C for 16
30 hours in 5X SSPE. 5X Denhart's, 0.2 mg/ml denatured
calf thymus DNA, 0.5% SDS. A putative alcohol oxidase
genomic clone 352 AO (Fig. 2) was identified and
characterized by restriction mapping.

Example 6Isolation of Alcohol Oxidase ZZA1 gene.

Three hundred nanograms of Pichia Pastoris genomic DNA digested with PstI and XhoI were amplified
5 by PCR using a 25-mer oligonucleotide
(5'TTGCACTCTGTTGGCTCATGACGAT 3') (SEQ ID NO:13)
corresponding to the nucleotide sequence of AOX1
promoter and a 26-mer oligonucleotide
(5'CAAGCTTGACAAACGAACGTCTCAC 3') (SEQ ID NO:14)
10 corresponding to a nucleotide sequence derived from
the AOX1 terminator. The PCR conditions using Thermus
aquaticus DNA polymerase (2.5 U; Perkin-Elmer Cetus)
consisted of an initial 2 minute incubation at 97°C
followed by two cycles at 97°C (1 min.), 45°C
15 (1 min.), 60°C (1 min.), thirty-five cycles at 94°C
(1 min.), 45°C (1 min.), 60°C (1 min.), and a final
DNA polymerase extension at 60°C for 7 min. The 3273
bp fragment containing ZZA1 gene was phenol/chloroform
treated and precipitated with ammonium
20 acetate/ethanol. After digestion with SacI the
fragment was purified by 1% low melt agarose
electrophoresis and subcloned into the SacI/EcoRV
sites in pBluescript KS-. A putative alcohol oxidase
genomic clone KS- A07'8' (Fig. 4) was partially
25 characterized by restriction mapping and dideoxy
nucleotide sequencing. Comparison of 172 bp of ZZA1,
AOX1, and AOX2 3' regions reveals that the ZZA1 and
AOX1 terminators are 66% homologous to each other
while ZZA1 and AOX2 terminators are only 24%
30 homologous (Fig. 5). In order to obtain the entire
nucleotide sequence of the ZZA1 gene an EcoRI/XhoI
fragment was subcloned into pBluescript SK- (SKA07'8',
Fig. 6).

Example 7

DNA sequencing

A 1095 bp PstI, BamHI fragment containing the entire putative alcohol oxidase ZZA1 promoter was subcloned into pBluescript KS+ and Bluescript KS- (Stratagene, La Jolla, Calif.). The nucleotide sequencing of KS+/AO (fig. 7) and KS-/AO (fig. 8) was carried out by dideoxy termination using single stranded templates. Nucleotide sequence analysis and amino acid sequence comparisons were performed using PCGENE™ and DNA Inspector™ IIE programs.

Example 8

Construction of ZZA1 regulatory region- rice α -amylase
15 gene fusions

A 990 bp PstI, HindIII fragment containing the alcohol oxidase promoter was subcloned into YEp352 (American Type Culture Collection # 37673). This resultant yeast expression vector 352H AO (fig.12) has a unique HindIII where foreign genes can be inserted. A 1.6 kb HindIII fragment from pUC18/103 was subcloned into the HindIII site of 352H AO. Plasmid pBGC103 (Fig. 13) contains the Pichia pastoris alcohol oxidase promoter, rice α -amylase cDNA pOS103 (GenBank #M24286), yeast 2 μ m ori, yeast URA3, and part of pUC18 plasmid. The vector pBGC103 has the following nucleotide sequence downstream of the putative transcription start point (tsp):

30 .tsp HindIII

5' CATCATTATCAGCTTACTTTCATAATTGTGACTGGTTCCAACCGACAAGC

3' GTAGTAATAGTCGAATGAAAGTATTAACACTGACCAAGGTTGGCTGTTCG

NcoI

TTGC ATG CAG GTG CTG AAC ACC ATG GTG AAC AAA CAC TTC 3'
AACG TAC GTC CAC GAC TTG TGG TAC CAC TTG TTT GTG AAG 5'
35 MET GLN VAL LEU ASN THR MET VAL ASN LYS HIS PHE
(SEQ ID NO:15)

Example 9Starch-clearing plate assay for rice α -amylase expression

Yeast cells containing pBGC103 were grown on YEP,
5 1% starch, 5 mM CaCl₂ and incubated at 30°C for three
days. The plates were stained with iodine vapors and
clear halos formed around yeast cells secreting active
 α -amylase. In order to directly select for amylolytic
colonies the yeast cells containing pBGC103 were grown
10 on DFM, pH 5.5, 1% starch, 5 mM CaCl₂ and incubated at
30°C for three days. The plates were stained with
iodine vapors and clear halos formed around yeast
cells secreting active α -amylase.

15 Example 10

RNA Isolation and analysis 1403-ER40[pBGC103] was
grown in YNBD and transferred to YEP, 0.5% methanol.
Total RNA was isolated at 0, 2, 4, and 6 hr after
methanol induction. The Pichia pastoris cells were
20 grown in YEPD and transferred to YEP, 0.5% methanol.
Total RNA was isolated from cells after 4 hr. The RNA
was fractionated on a 1.1% glyoxal gel and transferred
to nitrocellulose. The blot was probed with γ -³²P-dATP
end-labeled 31-mer oligonucleotide (5'CGATGTATCCCCGAAG
25 3') (SEQ ID NO:16) complementary to the leader region
of ZZA1. The transcription start point was determined
using primer extension analysis of total RNA isolated
from Pichia pastoris cells grown in YEP, 0.5%
methanol. A 31-mer oligonucleotide
30 (5'CGATGTATCCCCGAAG 3') (SEQ ID NO:16) complementary
to the leader region of ZZA1 was end-labeled with γ ³²-P
DATP and T4 polynucleotide kinase and hybridized to
RNA at 42°C in a 10 μ l reaction containing 400 mM KCl,
10 mM HEPES (pH 7.5), 1 mM EDTA for 12 hr. The
35 reverse transcription reaction was initiated by the

addition of 90 μ l of 50 mM Tris pH 8.0, 10 mM $MgCl_2$, 0.4 mM DTT, 1 mM of each deoxynucleotide triphosphates, 1.25 units/ μ l of RNasin and 30 units of AMV reverse transcriptase (Seikagaku). After a 42°C incubation for 90 min the final EDTA concentration was changed to 20 mM. 10 μ g of RNase was then added and the sample was incubated for 30 min at 37°C. The reaction mixture was brought to a final NH_4Ac concentration of 2M, extracted with phenol:chloroform:isoamyl alcohol (25:24:1), and precipitated with 2.5 volumes of ethanol. The cDNA was analyzed on an 8% polyacrylamide gel and the size of the extended product was determined by comparing it to a sequencing ladder of ZZA1 generated from the same end-labeled oligonucleotide primer.

Example 11

Effects of carbon source utilization on α -amylase expression In order to quantify the expression of rice α -amylase, a 2.5 ml inoculum from strain 1403-ER40[pBGC103] grown in YNBD + trp (20 μ g tryptophan per ml) was used to inoculate shake flasks containing 25 ml of YEP, 5 mM $CaCl_2$, 50 mM sodium acetate, pH 5.0, and various carbon sources. The cultures were incubated at 30°C with agitation (300 rpm). Samples were removed from the cultures after 44 hours and briefly centrifuged. A microassay was used to measure enzyme kinetics. The reaction was initiated by adding 50 μ l of supernatant to 1000 μ l of substrate (Sigma Kit #576-3). One unit of (α -amylase activity is defined as the amount of enzyme required to liberate 1 mmol of NADH/min. Each sample was assayed for enzyme activity at 25°C. Total soluble protein concentrations were determined using BSA as a standard. 12 μ l of supernatant from each sample was

analyzed on a 0.1% SDS, 12.5% polyacrylamide gel and transferred by electroblotting for 1 hour to a nitrocellulose membrane. The blotted membrane was incubated for 1 hour with a 2000-fold dilution of rabbit anti- α -amylase antiserum. The enhanced chemiluminescence horseradish peroxidase-linked, goat anti-rabbit IgG (Cappel) was developed according to the manufacturer's (Amersham) specifications. The autoradiogram was exposed for <30 seconds. The quantity of secreted α -amylase was estimated by comparing the crude extract autoradiogram signal to the signal obtained from known quantities of purified α -amylase. In order to study the effects of various carbon sources on the expression of rice α -amylase, an inoculum from strain 1403-ER40[pBGC103] grown in YNBD + trp (20 μ g tryptophan per ml) was used to inoculate shake flasks containing YEP, 5 mM CaCl₂, and various carbon sources at a final optical density of A₆₀₀ = 0.5/ml. The cultures were incubated at 30°C with agitation (300 rpm) and samples were removed from the cultures at 0, 2, 4, and 6 hr after inoculation. Equal volumes from each sample was subjected to Western analysis on a 12.5% PAGE as previously described with the following modifications. The blotted membrane was incubated for 1 hour with a 3000-fold dilution of rabbit anti- α -amylase antiserum and the autoradiogram was exposed for 5 minutes. Results of the experiment can be seen in Figure 17.

30 Example 12

Purification of alcohol oxidase

Pichia pastoris cells grown in YEP, 0.5% methanol were frozen in liquid nitrogen, and vortexed with glass beads in 20 mM Tris-HCL, pH 8.0. The suspension was centrifuged and the supernatant was bound on a

FPLC Mono Q ion-exchange column (Pharmacia). The sample was eluted with a linear NaCl gradient (0-1 M NaCl in 20 mM Tris-HCL, pH 8.0), and the fractions were assayed for alcohol oxidase activity. The N-terminal sequence of the purified protein was obtained using an automated protein sequenator. The N-terminal sequence NH₃-Ala-Ile-Pro-Glu-Glu-Phe-Asp-Ile-Ile-Val-Cys-Gly -Gly-Gly-Ser-COOH (SEQ ID NO:7) obtained from the purified protein was identical to the amino acid sequence deduced from the ZZA1 genomic clone.

Example 13

Regulation of α -amylase production in Saccharomyces cerevisiae

In order to study the effects of various carbon sources on the expression of rice α -amylase, a 2.5 ml inoculum from strain 1403-ER40[pBGC103] grown in YNBD + trp (20 μ g tryptophan per ml) was used to inoculate shake flasks containing 25 ml of YEP, 5 mM CaCl₂, 50 mM sodium acetate, pH 5.0 and various carbon sources. The cultures were incubated at 30°C with agitation (300 rpm). Samples were removed from the cultures after 44 hours and briefly centrifuged. A microassay was used to measure enzyme kinetics. The reaction was initiated by adding 50 μ l of supernatant to 1000 μ l of substrate (Sigma® Kit #576-3). One unit of α -amylase activity is defined as the amount of enzyme required to liberate 1 mmol of NADH/min. Each sample was assayed for enzyme activity at 25°C. The results are summarized in the Table I.

TABLE I

Rice α -amylase activity in *Saccharomyces cerevisiae* transformants

5	<u>Strain</u>	<u>Carbon Source</u>	<u>α-amylase activity</u>
	(U/ L)		
	1403-ER40	5% glucose	0
	1403-ER40[pBGC103]	0	7.1
	1403-ER40[pBGC103]	2% glucose	49.4
10	1403-ER40[pBGC103]	5% glucose	2.4
	1403-ER40[pBGC103]	2% sucrose	101.2
	1403-ER40[pBGC103]	5% sucrose	82.4
	1403-ER40[pBGC103]	34 glycerol	72.9
	1403-ER40[pBGC103]	24 ETOH	58.8
15	1403-ER40[pBGC103]	5% ETOH	25.9

Total soluble protein concentrations were determined using BSA as a standard. 12 μ l of supernatant from each sample was analyzed on a 0.1% SDS, 12.5% polyacrylamide gel and transferred by electroblotting for 1 hour to a nitrocellulose membrane. The blotted membrane was incubated for 1 hour with a 2000-fold dilution of rabbit anti- α -amylase antiserum. The enhanced chemiluminescence horseradish peroxidase-linked, goat anti-rabbit IgG (Cappel) was developed according to the manufacturer's (Amersham) specifications. The autoradiogram was exposed for <30 seconds. The quantity of secreted α -amylase was estimated by comparing the crude extract autoradiogram signal to the signal obtained from known quantities of purified α -amylase. Western blot analysis of secreted proteins was performed on 1403-ER40[pBGC103] grown in various carbon sources. A Western blot was performed in which the following lanes were run. (Lane 1: .15 μ g of purified rice α -amylase; 2: .75 μ g of purified

α -amylase; 3: Cells containing pBGC103 were grown in YEP, 5 mM CaCl₂, 50 mM NaAcetate, pH 5.0 and 3% glycerol; 4: 2% ethanol; 5: 2% glucose; 6: 2% sucrose; 7: 3% starch (ammonium sulfate precipitated extracellular proteins); 8: 3% starch (phenyl superose FPLC purified protein)). The Western blot results indicated that the α -amylase produced by ER40[pBGC103] was secreted in excess of 25 μ g/ml. Alpha amylase was produced in the presence of ethanol.

10

Example 14

Construction of ZZA1 BstXI site directed mutant regulatory region- rice α -amylase gene fusions

A rice α -amylase expression vector, pBGC103 BstXIL, containing an additional TTCAA core consensus sequence in the ZZA1 promoter was constructed. A 30 bp BstXI fragment, made by annealing a 30-mer oligonucleotide (5'CTGGTCTAGACATTGTATGCTTCCAAGTTT 3')(SEQ ID NO:18) to a 30-mer oligonucleotide (5'TTGAAGCATACAATGTCTAGACCAGAAAC 3')(SEQ ID NO:19), was subcloned into the BstXI site of pBGC103. The resultant new plasmid, pBGC103 BstXIL (Fig. 14), contains a mutant *Pichia pastoris* alcohol oxidase promoter containing an additional TTCAA core consensus sequence, rice α -amylase cDNA pOS103 (GenBank #M24286), yeast 2 μ m ori, yeast URA3, and part of pUC18 plasmid. The *Pichia -pastoris* alcohol oxidase ZZA1 promoter has the following nucleotide sequence:

30

NsiI	BstXI
5' ATGCATTGTTCTCCACATTGTATGCTTCCAAGTTTCTGG 3'	
3' TACGTAACAAGAGGTGTAACATACTAAGGTTCAAAGACC 5'	

(SEQ ID NO: 20)

35

Two oligonucleotides used in the site directed mutagenesis of the ZZA1 promoter have the following nucleotide sequence:

5 BstXI XbaI BstXI

5'CTGGTCTAGACATTGTATGCTTCCAAGTTT 3' (SEQ ID NO:21)
3'CAAAGACCAGATCTGTAACATACGAAGGTT 5' (SEQ ID NO:22)

In order to compare the relative expression of rice α -amylase from Saccharomyces cerevisiae ER40[pBGC103] to ER40[pBGC103 BstXILI], 2.5 ml inoculum from strain ER40[pBGC103] and strain

5 ER40[pBGC103 BstXILI] grown in YNBD + trp (20 μ g tryptophan per ml) were used to inoculate shake flasks containing 50 ml of YEPD. The cultures were incubated at 30°C with agitation (300 rpm). Samples were removed from the cultures after 48 hours and briefly

10 centrifuged. A microassay was used to measure enzyme kinetics. The reaction was initiated by adding 50 μ l of supernatant to 1000 μ l of substrate (Sigma Kit #576-3). One unit of α -amylase activity is defined as the amount of enzyme required to liberate 1 mmol of

15 NADH/min. Each sample was assayed for enzyme activity at 25°C. The results are summarized in the Table II.

TABLE II

Rice α -amylase activity in Saccharomyces cerevisiae

20 transformants

	<u>Strain</u>	<u>Carbon Source</u>	<u>α-amylase activity</u>
	(U/ L)		
	1403-ER40[pBGC103]	2% glucose	35.3
25	1403-ER40[pBGC103 <u>BstXILI</u>	2% glucose	96.5

28 μ l of supernatant from each sample was analyzed on a 0.1% SDS, 12.5% polyacrylamide gel and transferred by electroblotting for 1 hour to a nitrocellulose

30 membrane. The blotted membrane was incubated for 1 hour with a 2000-fold dilution of rabbit anti- α -amylase antiserum. The enhanced chemiluminescence horseradish peroxidase-linked, goat anti-rabbit IgG (Cappel) was developed according to the manufacturer's

35 (Amersham) specifications. The autoradiogram was

exposed for <30 seconds. The results of the Western blot analysis of secreted proteins from 1403-ER40[pBGC103] and 1403-ER40[pBGC103 BstXILI are in Figure 17.

5

Example 17

Purification of α -amylase from *S. cerevisiae*

A 50 ml inoculum from strain ER40[pBGC103] grown in YNB + trp (0.67% yeast nitrogen base without aa, 10 .2% glucose, 20 μ g tryptophan per ml) was used to inoculate.

500 ml of YEP, 3% potato starch, 5 mM CaCl_2 (pH 5.0). The culture was grown in a shake flask for 40 hrs at 30°C with agitation (300 rpm). After spinning 15 down the cells the supernatant was concentrated by ammonium sulfate precipitation (60%). The pellet was resuspended in approximately 10 mls of 10 mM Tris pH 7.4, 1M ammonium sulfate. The crude enzyme was then purified using phenyl superose HR 10/10 column 20 chromatography (Pharmacia FPLC, 1-0 M ammonium sulfate gradient). The α -amylase was eluted in fraction 66 and its enzyme activity was analyzed (Sigma Kit 576-3).

25 Example 18

Gel mobility shift assays

A *Pichia pastoris* nuclear extract is prepared by growing the yeast cells in 250 mls of YNB (minimal) 30 media plus 0.5% methanol (induction media) or 5% glucose (repression media) for 48 hours at 30°C. The cells are pelleted by centrifugation at 3,000 rpm for 5 minutes. The pellets are washed once in ddH₂O, once in SED, once in im sorbitol, and then they are 35 resuspended in 25 ml of 0.1 M tris-HCL, pH 7.0, and 1

M sorbitol. The cells are mixed with 750 μ l Zymolyase 100,000 (Seikagaku) and are incubated at 30°C for 1 hour. The resulting spheroplasts are centrifuged at 1,000 g for 5 minutes and resuspended in a glass bead
5 disruption buffer containing 50 mM Tris-HCL, pH 7.9, 5 mM MgCl₂, 1 mM dithiothreitol, 0.2 mM EDTA, 2.5 mM phenylmethylsulfonyl fluoride, 10% dimethyl sulfoxide, and 20% glycerol. The whole-cell extracts are quickly frozen by dropwise addition into liquid nitrogen. In
10 some preparations 2 μ M pepstatin A and 0.6 μ M leupeptin are included.

DNA-binding assays are performed in 5 μ l reaction mixes containing 15 mM Tris-HCL (pH 7.4), 50 mM KCL, 1 mM EDTA, 0.5 mM dithiothreitol, 50 μ g of bovine serum
15 albumin per ml, 0.05% Nonidet P-40, 25 μ g of poly (dI-dC) per ml, 6% glycerol, 2 fmol of 5'-end-labeled DNA probe (specific activity, 5 X 10⁶ dpm/pmol), and 1 to 20 μ g of whole-cell extract protein. The reactions are incubated for 2 to 20 minutes at 21°C, then 15 μ l
20 of 10 mM Tris-HCL (pH 7.4), 1 mM EDTA, 0.05% bromophenol blue, 0.0596 xylene cyanol blue, and 6% glycerol is added. The binding reactions are subjected to electrophoresis for 2 hours at 4 V/cm at 4°C on a 1.5% agarose gel in buffer containing 50 mM
25 Tris-borate (pH 8.3) and 1 mM EDTA. After electrophoresis the gel is dried on Whatman DE81 chromatography paper and subjected to autoradiography.

Competition DNA-binding assays are performed as described above in the absence of poly(dI-dC).
30 Unlabeled competitor DNA is added to the binding reaction mixture. The oligonucleotide probes used in this experiment have the following nucleotide sequences:

5' TTGTATGCTTCCAAGTTTCTGGT 3' (SEQ ID NO:24)
5' ACACCCGCTTTTGG 3' (SEQ ID NO:25)

5 Example 19

Isolation of methanol induced trans-acting factor

A random primed cDNA library is constructed using poly(A) RNA prepared from Pichia pastoris cells grown
10 in 250 mls of YNB (minimal) media plus 0.5% methanol (induction media). The amplified library is screened with a labeled oligonucleotide fragment that consists of the TTCAA motif.

15 Example 20

Isolation of methanol specific transacting factor by using an alcohol oxidase promoter shuttle vector

A PstI-KpnI fragment containing the ZZA1 promoter
20 and the rice α -amylase cDNA OS103 is subcloned into YIp5. This vector is then integrated into 334 by Leu2+ complementation.

A random primed cDNA library is constructed in 352H AO using poly(A) RNA prepared from Pichia
25 pastoris cells grown in 250 mls of YNB (minimal) media plus 0.5% methanol (induction media). The amplified cDNA library is transformed into 334[YIp5 AO 103] and recombinant yeast cells are selected on DFM plus 1% starch. Putative clones coding for methanol trans-
30 acting factors are identified by the production of large halos.

Example 21

Transformation and analysis of strain

35 1403-ER40[pBGC103]

1403-ER40, an ethanol tolerant strain of S. cerevisiae (MAT α , gal3, gal4, mel, SUC+, MAL4, MGL3, trp1, ura3), was transformed with pBGC103. Cultures of ER40[pBGC103] grown in YEP 3% starch were analyzed for rice α -amylase activity and ETOH production (Sigma® kits, catalogue numbers 576-3 and 332, respectively). Yeast DNA was isolated, analyzed by PCR, and transformed into E. coli C600. Restriction digests of miniscreen DNA isolated from the transformants indicate that pBGC103 was maintained in an extrachromosomal state and had not undergone any detectable intramolecular rearrangements.

Example 22

15

Construction of ZZA1 regulatory region- rice α -amylase gene fusions for expression in *Hansenula polymorpha*

A 2512 bp PstI, SmaI fragment containing the alcohol oxidase promoter and the rice α -amylase cDNA OS103 was subcloned into YEp351. This yeast expression vector is called 351HAO 103 (Fig. 16) and contains the Pichia -pastoris alcohol oxidase promoter, rice α -amylase cDNA pOS103 (GenBank #M24286), yeast 2 μ m ori, yeast LEU2, and part of the pUC18 plasmid. The vector 351HAO 103 has the following nucleotide sequence downstream of the putative transcription start point (tsp):

30 **.tsp** **HindIII**

5' CATCATTATCAGCTTACTTTTCATAATTGTGACTGGTTCCAACCGACAAGC
3' GTAGTAATAGTCGAATGAAAGTATTAACACTGACCAAGGTTGGCTGTTCG

NCOI

TTGC ATG CAG GTG CTG AAC ACC ATG GTG AAC AAA CAC TTC 3'
AACG TAC GTC CAC GAC TTG TGG TAC CAC TTG TTT GTG AAG 5'
MET GLN VAL LEU ASN THR MET VAL ASN LYS HIS PHE

35

(SEQ ID NO:15)

Example 23Transformation and analysis of Hansenula polymorpha
[351HAO 1031

5 A 24 hour 50 ml culture of the methylotrophic
yeast, Hansenula polymorpha (leul-1), was grown in
YEPD at 30°C. The cells were pelleted, washed once
with LITE (0.1 M Lithium acetate, 10 mM Tris pH 7.6, 1
mM EDTA), resuspended in 1 ml LITE, and incubated at
10 30°C for 1 hour with agitation. The competent
Hansenula polymorpha cells were transformed with
351HAO 103. Cultures of Hansenula polymorpha [351 HAO
103] #18 grown in YEP 0.5 & methanol, 5 mM CaCl₂ were
analyzed for rice α -amylase activity (Sigma® kits
15 #576-3).

Example 24Construction of ZZA1 alcohol oxidase expression
vector

20 A 2512 bp HindIII fragment containing the alcohol
oxidase ORF was subcloned into 352H AO. This yeast
expression is called pBGC104 and contains the Pichia
pastoris alcohol oxidase gene ZZA1, yeast 2 μ m ori,
yeast URA3, and part of the plasmid pUC18.

25

Example 25Isolation of Alcohol Oxidase ZZA2 terminator

1.5 micrograms of Pichia pastoris genomic DNA
digested with PstI and XhoI was amplified by PCR using
30 a 24 -mer oligonucleotide (5' TCG ACC CAG GTT TCA TGA
ACG ATG 3') (SEQ ID NO:4) corresponding to the
nucleotide sequence of AOX1 and AOX2 ORF and a 31-mer
oligonucleotide (5'TCC TGC AGC AAC CAA TGA GGA GAA TGA
CAA C 3') (SEQ ID NO:5) corresponding to a nucleotide
35 sequence derived from the AOX2 terminator. The PCR

conditions using Thermos aquaticus DNA polymerase (2.5 U; Perkin-Elmer Cetus) consisted of an initial 2 minute incubation at 97°C followed by five cycles at 97°C (1 min.), 45°C (1 min.), 60°C (2.5 min.), thirty-
5 five cycles at 94°C (1 min.), 45°C (0.5 min.), 60°C (2.5 min.), and a final DNA polymerase extension at 60°C for 7 min. The 910 bp fragment containing ZZA02 gene was phenol/chloroform treated and was precipitated with ammonium acetate/ethanol. After
10 digestion with PstI the fragment was purified by 1% low melt agarose electrophoresis and subcloned into the EcoIV/PstI sites in pBluescript KS+. The resultant plasmid, KS+ ZZA2T (Fig. 3), is a putative alcohol oxidase genomic clone containing a terminator
15 sequence.

Biological Deposits

On September 15, 1992, Applicants have deposited
20 with the American Type Culture Collection, Rockville, Md., USA (ATCC) the plasmid pBGC103, in S. cerevisiae, described herein, under ATCC accession no. 74185. On January 6, 1993, Applicants have deposited with the American Type Culture Collection, Rockville, Md., USA
25 (ATCC) the plasmid pBGC104, in E. coli, described herein, under ATCC accession no. 69181. On March 11, 1994, Applicants have deposited with the American Type Culture Collection, Rockville, Md., USA (ATCC) the plasmid KS+/ZZA2T, in E. coli JM109, described herein,
30 under ATCC accession no. 69585. These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the purposes of patent procedure and the Regulations thereunder (Budapest Treaty).
35 This assures maintenance of a viable culture for 30

years from date of deposit. The organisms will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Applicants and ATCC which assures unrestricted availability upon
5 issuance of the pertinent U.S. patent. Availability of the deposited strains is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

10

Equivalents

All publications and patents mentioned in the above specification are herein incorporated by
15 reference. The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those
20 skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

25

30

35

International Application No: PCT/

MICROORGANISMS	
Optional Sheet in connection with the microorganism referred to on page 88, lines 1-12 of the description *	
A. IDENTIFICATION OF DEPOSIT *	
Further deposits are identified on an additional sheet *	
Name of depositary institution *	
American Type Culture Collection	
Address of depositary institution (including postal code and country) *	
12301 Parklawn Drive Rockville, MD 20852 US	
Date of deposit * <u>September 15, 1992</u> Accession Number * <u>74185</u>	
B. ADDITIONAL INDICATIONS * (leave blank if not applicable). This information is continued on a separate attached sheet	
C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE * (if the indications are not all designated States)	
D. SEPARATE FURNISHING OF INDICATIONS * (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later * (Specify the general nature of the indications e.g., "Accession Number of Deposit")	
E. ☒ This sheet was received with the International application when filed (to be checked by the receiving Office)	
 YVETTE E. SIMMS INTERNATIONAL DIVISION (Authorized Officer)	
☐ The date of receipt (from the applicant) by the International Bureau *	
was _____	
(Authorized Officer)	

International Application No: PCT/ /

Form PCT/RO/134 (cont.)

American Type Culture Collection

12301 Parklawn Drive
Rockville, MD 20852
US

Accession No.

69181

69585

Date of Deposit

January 6, 1993

March 11, 1994

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising at
5 least at least a 15 nucleotide portion of ZZA1 or
ZZA2.

2. A polynucleotide according to claim 1
comprising the regulatory region of alcohol oxidase
10 gene ZZA1 or ZZA2.

3. A polynucleotide according to claim 1,
wherein said polynucleotide comprises the coding
region of ZZA1 or ZZA2.
15

4. An isolated polynucleotide capable of
hybridizing to ZZA1 or ZZA2 wherein said
polynucleotide encodes an enzyme possessing alcohol
oxidase activity.
20

5. An isolated polynucleotide according to Claim
4, wherein hybridization takes place under stringent
hybridization conditions.

25 6. An expression vector comprising a promoter
for heterologous gene expression, wherein said
promoter is a polynucleotide according to claim 2
joined in functional combination with a heterologous
gene for expression.

30 7. An expression vector according to Claim 6,
wherein said promoter does not contain the ZZA1 URS
polynucleotide sequence.

35

8. An expression vector according to Claim 7,
wherein said heterologous gene is rice α -amylase.
9. A host cell transformed by a vector
5 according to Claim 6.
10. A host cell according to Claim 9 wherein
said host cell is a cell selected from the group
consisting of yeast cells, fungal cell, and plant
10 cells.
11. Purified ZZA1 or ZZA2.
12. Purified ZZA1 or ZZA2 according to Claim 11,
15 wherein the purified protein is isolated from cells
genetically modified to produce increased levels of
ZZA1.
13. A method of isolating an AOER protein said
20 method comprising the step, assaying a cellular
extract for proteins that can bind to polynucleotides
selected from the group consisting of the regulatory
region of ZZA1, the regulatory region of ZZA2, and
polynucleotides comprising the polynucleotide
25 TTGNNNGCTTCCAANNNNTGGT (SEQ ID NO:2), wherein said
cellular extract is prepared from an organism
naturally producing an alcohol oxidase.
14. A method of isolating an AOER gene, said
30 method comprising the step, screening a genetic
library for isolates encoding a protein that can bind
to the polynucleotides selected from the group
consisting of the regulatory region of ZZA1, the
regulatory region of ZZA2, and polynucleotides
35 comprising the polynucleotide TTGNNNGCTTCCAANNNNTGGT

(SEQ ID NO:2), wherein said genetic library is prepared from an organism naturally producing an alcohol oxidase.

- 5 15. A method of converting starch into ethanol comprising the step,
growing yeast cells in a culture medium, wherein said yeast cells comprise a genetic construction, said genetic construction comprising a starch degrading
10 enzyme encoding nucleotide sequence in functional combination with the regulatory region of an alcohol oxidase gene isolated from a methylotrophic yeast.
16. A method according to Claim 15, wherein said
15 starch degrading enzyme is α -amylase.
17. A method according to Claim 15, wherein said regulatory region comprises at two least copies of the nucleotide sequence TTGNNNGCTTCCAANNNNTGGT (SEQ ID
20 NO:2).
18. A method according to claim 15, wherein said regulatory region does not contain the nucleotide sequence GTTTCCTCAAGGCAAGAACTCC (SEQ ID No:3).
25
- 30
- 35

19. A method of producing secreted α -amylase comprising the steps,

- growing yeast cells in a culture medium, wherein
5 the yeast cells comprise a genetic construction, said genetic construction comprising an α -alpha amylase encoding nucleotide sequence in functional combination with a regulatory region resented on a polynucleotide according to Claim 4, and
10 recovering secreted α -amylase from the medium.

20. A method of producing secreted α -amylase according to claim 19, wherein the α -amylase is rice α -amylase.

15

20

25

30

35

Fig. 1

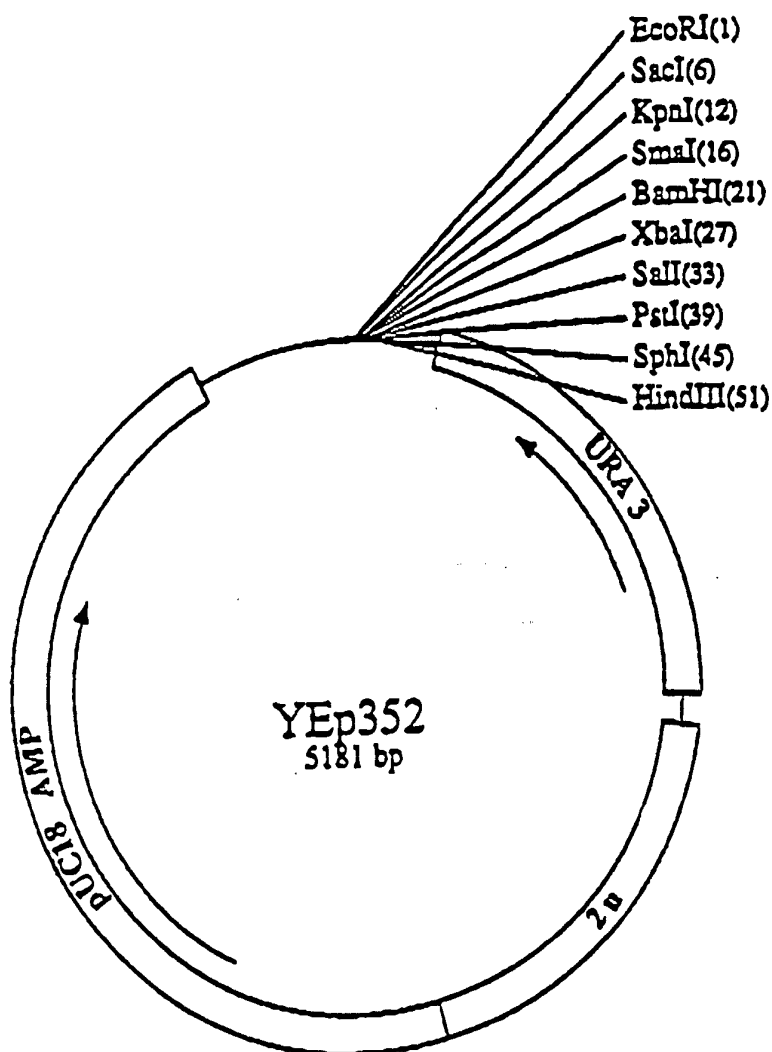


Fig. 2

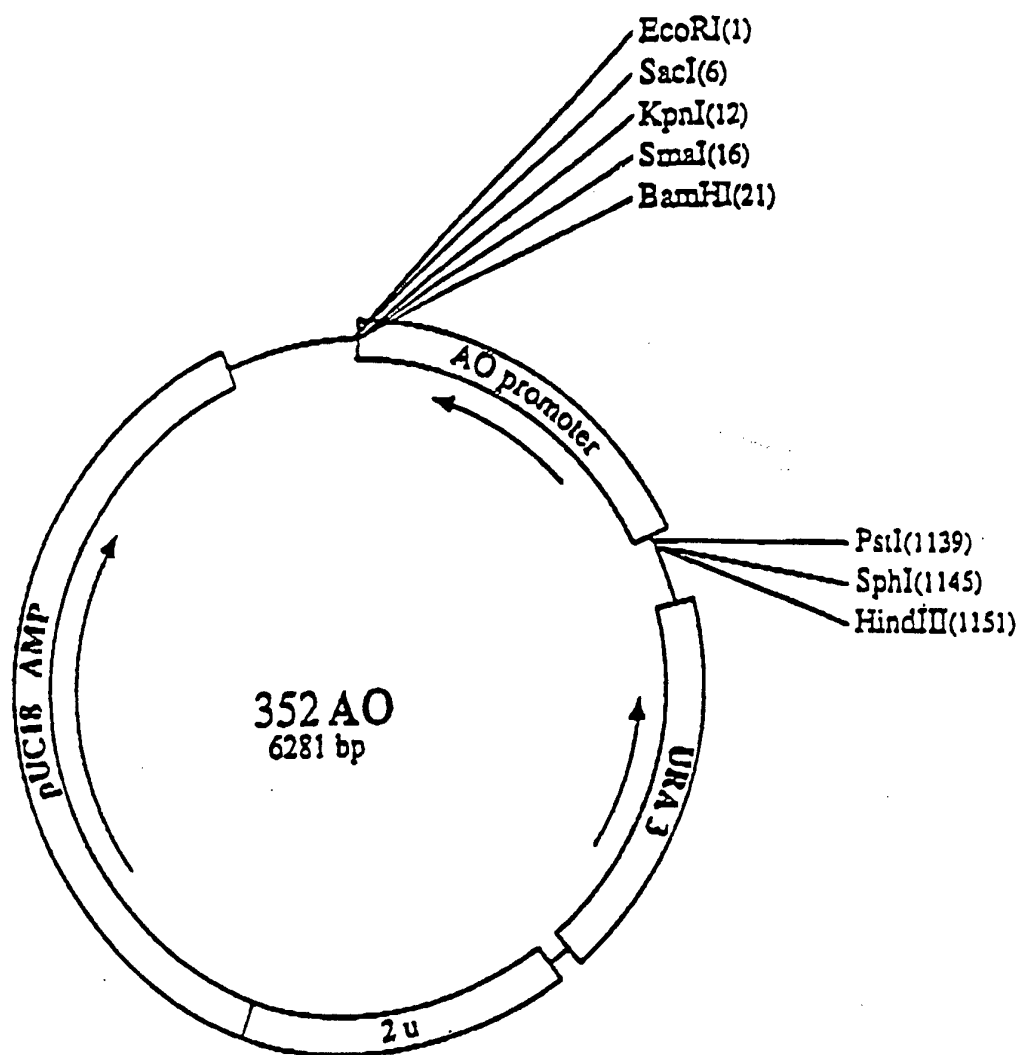


Fig. 3

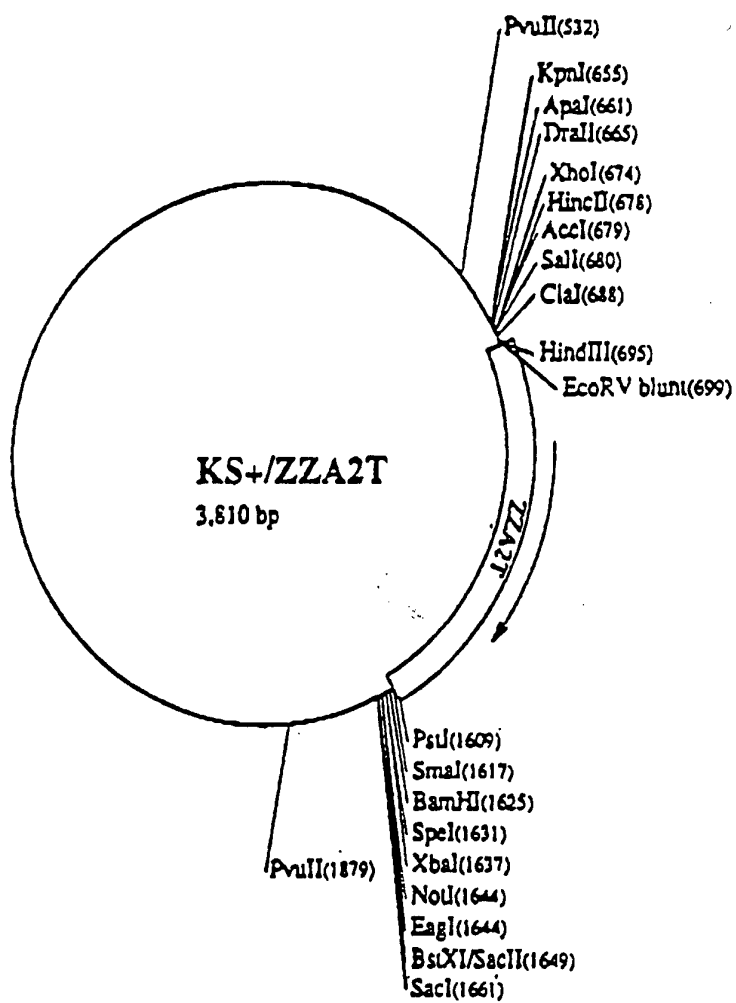
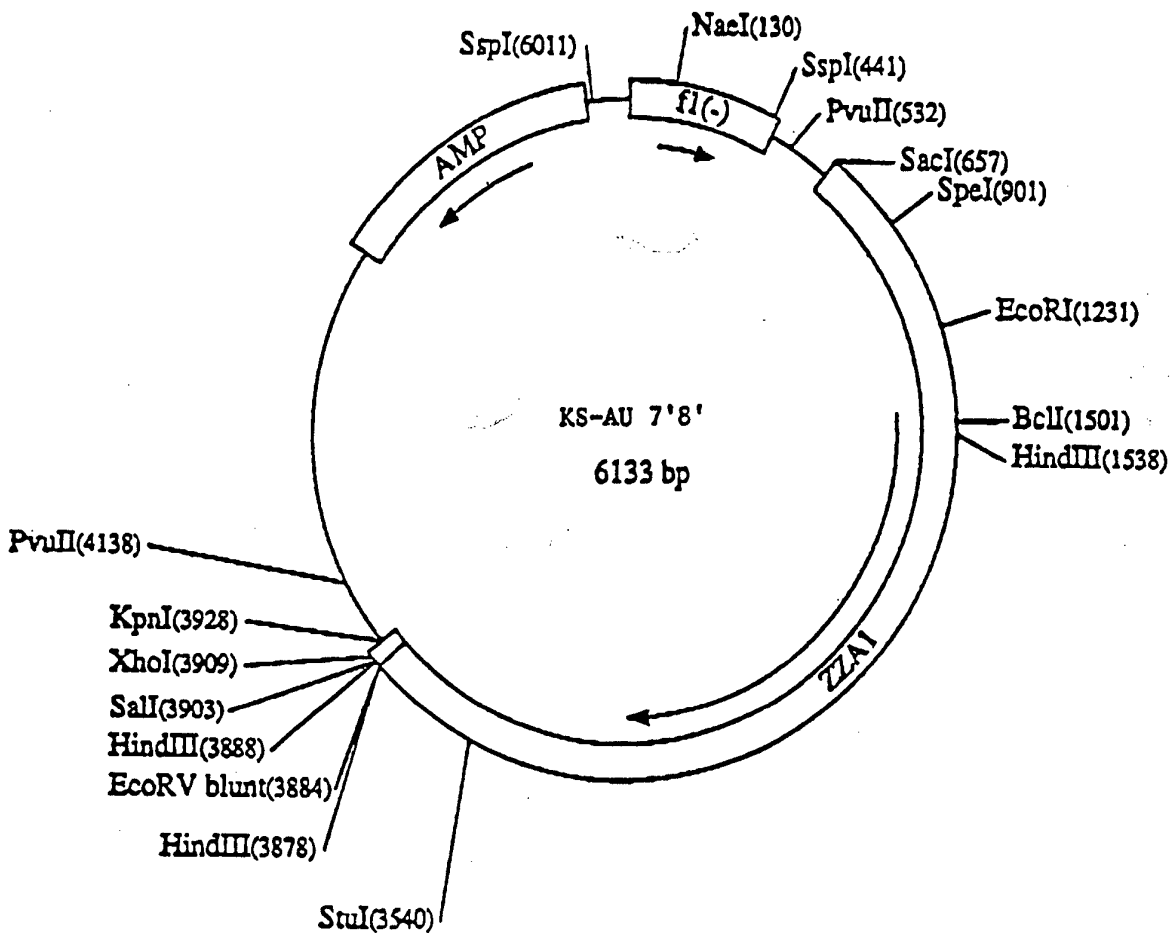


Fig. 4



TTCTCTCTGGCACTATTTTGCTTCTTATCAGTCTATCTTTGAGTTGGTGAATATCTTGAGA	ZZA1
TTTCTTCTCGTACGAGCTTGCCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGG	AOX1
CATGGGCTTGGGGAAATCATTTGATTTCGAAGTTTTGCTTGGTAGTTGACATTCTTCTTC	ZZA1
TAGGGGTTTGGGAAATCATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTC	AOX1
GGAGTATAAAAGATTTAGTGAGACGTTTCGTTTGTGCAAGCTT	ZZA1
AGAGTACAGAAGATTAAAGTGAGACGTTTCGTTTGTGCAAGCTT	AOX1

B. Comparison of ZZA1 and AOX2 3' regions:

TTCTCCTGGCACTATTTTGCTTCTTATCAGTCTATCTTTGAGTTGGTGAATATCTTGAGA	ZZA1
TTTCATTTCTTTTATACGTACGTATATGTACTAGATGAAGAATGCGACAAGGCCGACCA	AOX2
CATGGGCTTGGGGAAATCATTTGATTTCGAAGTTTGGCTTGGTAGTTGACATTCTTCTTC	ZZA1
	-
ACAGCAATGGTGCTTGGTACCAAAGTTTGGAAAGGTGCTACCGAATTGGCCGATGATATTG	AOX2
GGAGTATAAAAGATTTAGTGAGACGTTTCGTTTGTGCAAGCTT	ZZA1
AGTGGAGTTGTCATTCT	AOX2

Note: The sequence in bold is part of the A08' primer. Comparison of the two *Pichia pastoris* alcohol oxidase 3' regions indicate that they are 24% homologous to each other.

Fig. 6

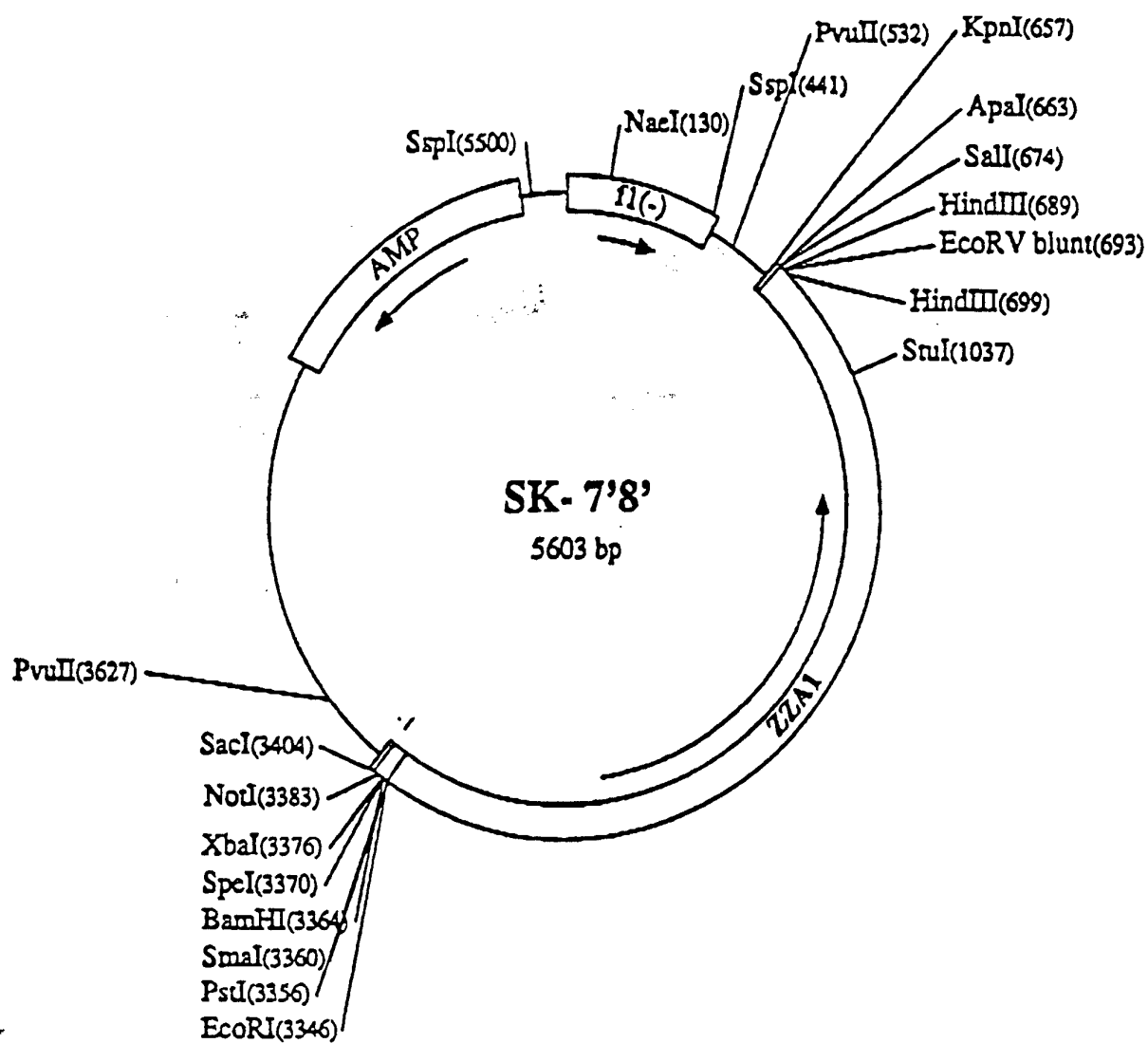


Fig. 7

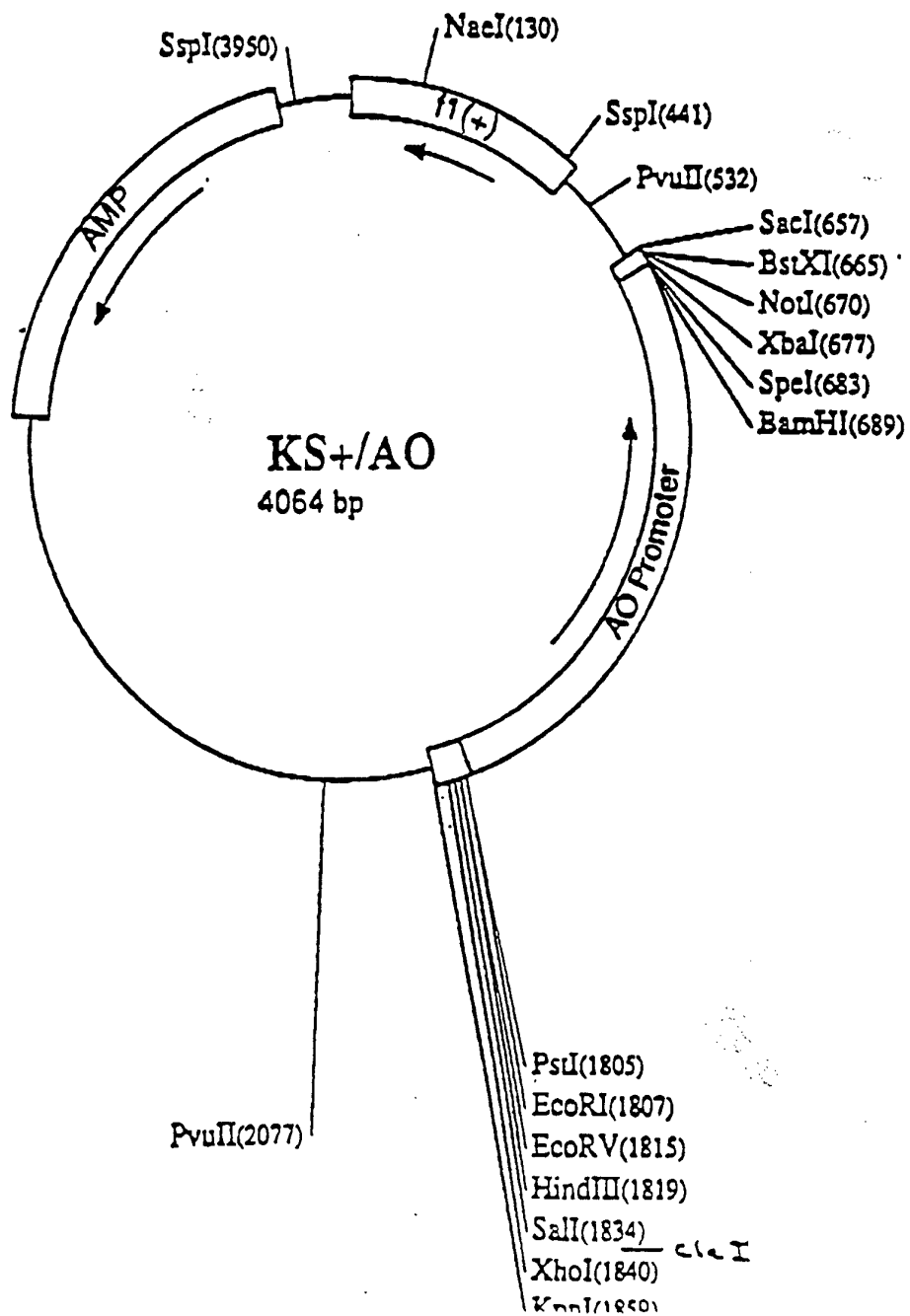


Fig. 8

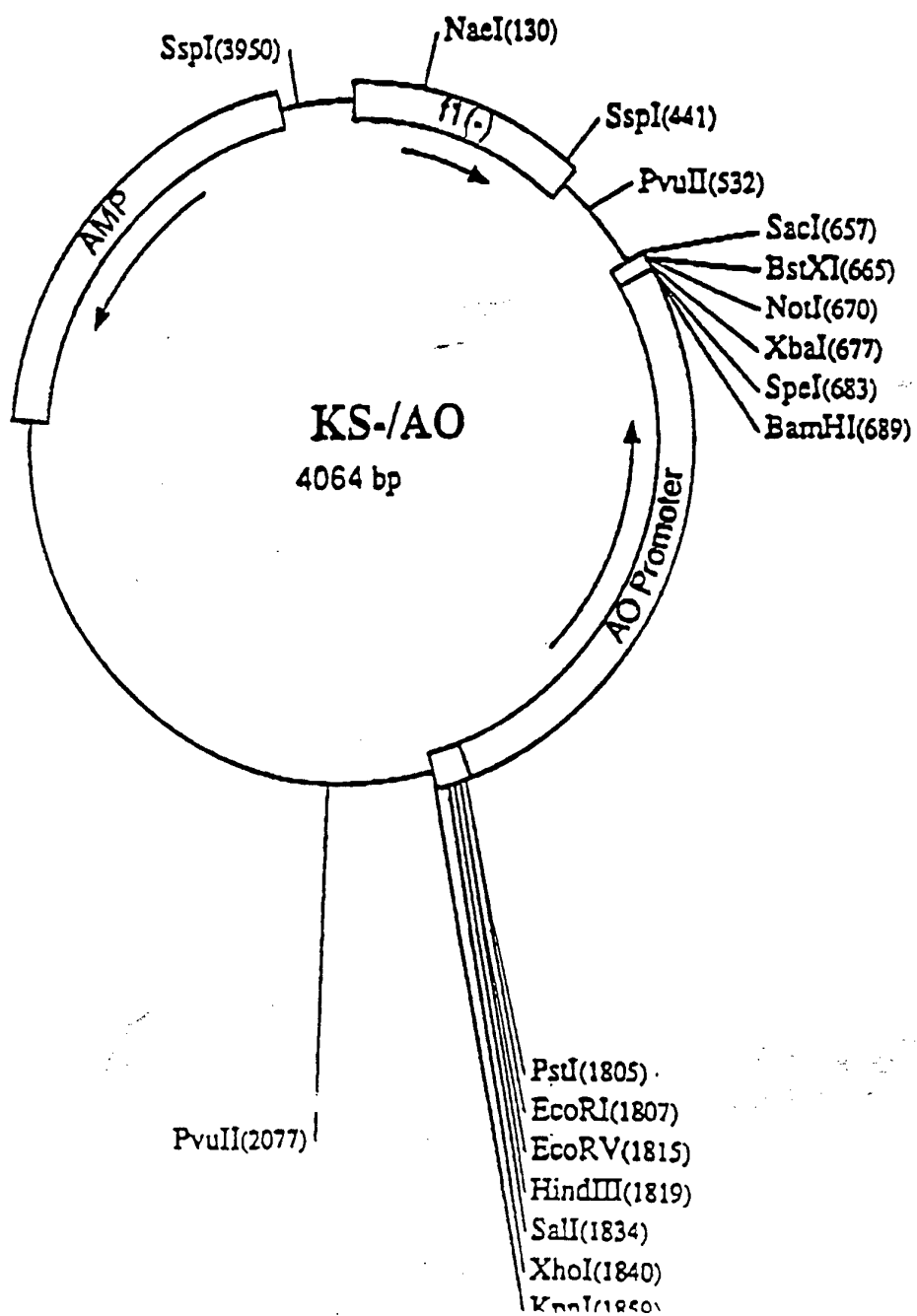


Fig. 9

PstI
 CTGCAGCTTTGATACCTGAAATTCCTGAGCCTATAATAATGACTTTTGCACTCTGTTGGCTCATGACGAT -867

SacI
 TTTGTTGAAATGAATCTTCACAAGAAGAGCTCAATTGAGTAGAGATAATTAGTAAGTGAGATCCAACACC -797

 CAGGAACGAGATGGATTGAGGAGAAATTGTTCTGCCATCCGACATCGACAAGTTAGACACAATAGTGCCA -727

 AATGCAGAGGGGACGTTTCCTCAAGGCAAGAAGTCCACTTTATTCTCCTCAAACACCCGCCCTCGCCGT -657

SpeI
 TAAAAACCAGCCCAGTTACTAACATGGTTTGGACTCTCTCTAATCCACTTTGTTAGGCTACTAGTAGCA -587

 TTATTTTCTTAGCCTGTCTATATGGTTTCCTTGCGAGTTTTTAATTTTATTCTATTTCGGAATGTAAGT -517

 ACTCCGCATTCCATCCCAACACCAGAAAGTTGAGGGTTTTTGTGAGTGTGGGGTCGGTAACAGTTTCATG -447

 TTCCCCCAATGGCCTAAAATTGACACTTTAGACGCCCTGTTCAAACCTCAAATTGACAAAAGCGTGATCTC -377

 ATCAGAGATGAACTAGGTTTGGTTTCGATCAAAAGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAATGT -307

EcoRI
 CGGCATACCGTTTGTTCGTTTGACCCGACAATTGATGTTGAAGAATTCCCTCTTACACTTAGCGCAGCC -237

 TTTATTTTGCTTGGGGTCTCGCTGCGCTTGGGTCTCGGTGTGCTTGTGACCGGAAACGCAAATGGGGAAA -167

NsiI **BstXI**
 CACCCGCTTTTGGATGATTATGCATTGTTCTCCACATTGTATGCTTCCAAGTTTCTGGTGGGAATACTG -97

 ATAGCCTAACGTTTCATGATCAAACTAATGTCTTCCCTACTTGAACAGCAATATATATAACAGAAGAAGAT -27

 TTCCTTTCTAAGGTCTTTTTTTTTATCATCATTATCAGCTTACTTTTCATAATTGTGACTGGTTCCAATTG +44

HindIII **BstBI**
 ACAAGCTTTTGATTCTAACGACTTTAACGACAACCTAAAGAACAAAAACAATAATTATTCGAAACAATG +114
 Met

BamHI
 GCT ATT CCC GAA GAA TTT GAT ATT ATC GTC TGT GGT GGT GGA TCC +159
 Ala Ile Pro Glu Glu Phe Asp Ile Ile Val Cys Gly Gly Gly Ser

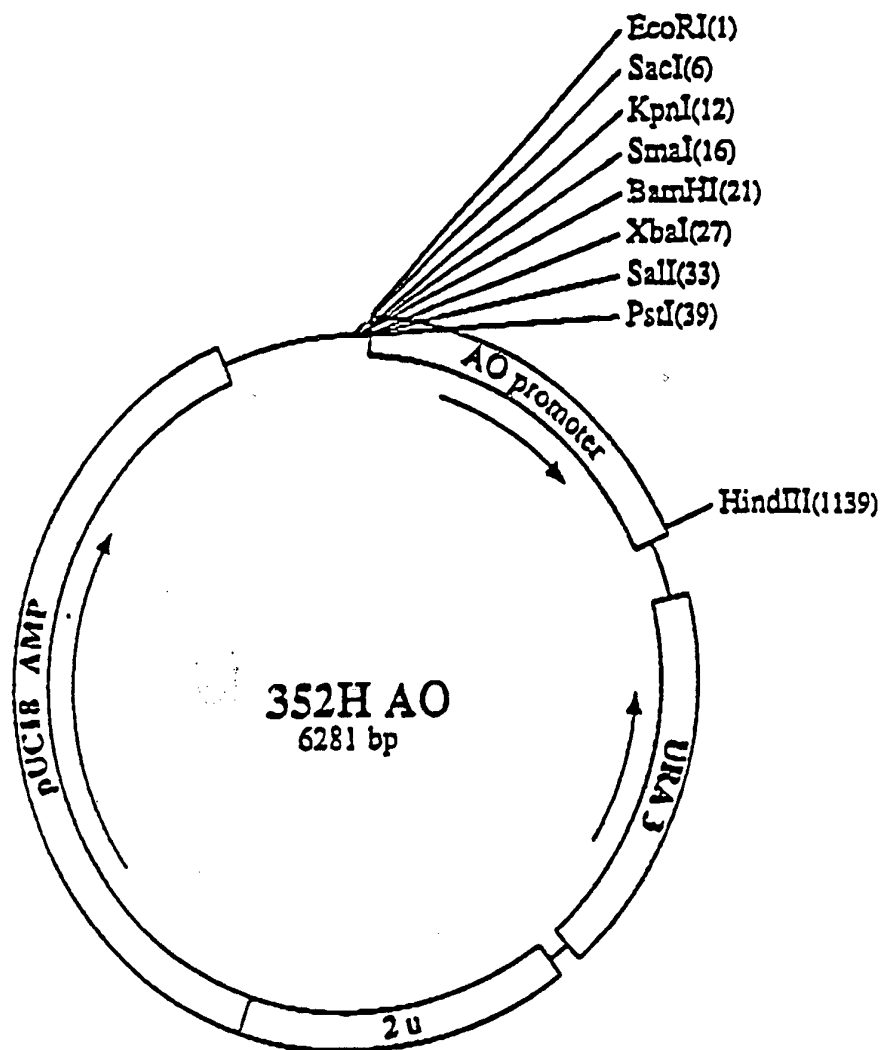
Fig. 10

[illegible]

Fig. 11

ZZA1 (P)	-316	TCAAAAAGAACTTCCAATGTGGCATA
AOX1 (P)	-314	TCAAAAAGAACTTCCAAAAGTCGCCATA
ZZA1 (P)	-141	ACATTGTATGCTTCCAAGTTCTGGTGG
AOX1 (P)	-140	ACATTGTATGCTTCCAAGATTCTGGTGG
AOX2 (P)	-275	ACAGTTGGGAGTTTCCAATTGGTTGGTTTT
MOX (H)	-409	TGTGCTGGATGC.ACCAATTAATTGTTGC
MOX (H)	-658	GGCTTTGGTCATTT.CAATGTTGTCGTC
AOD (C)	-583	AAAATGCTCTTTTCCATCATCATCATC
DAS (H)	-751	GTCTTTGATGTCTTCCA.CCATCTGCAGAT
Consensus		TTG.....TTCCAA.....TGGT
ZZA1 (P)	+35	AATTGTGACTGGTTCCAATTGACAAGCTT
AOX1 (P)	+25	AATTGCGACTGGTTCCAATTGACAAGCTT
MOX (H)	-61	AATTCT.TATGCTACCGTGCAGCGACTC
DAS (H)	-96	AATTTAGCCTCGTTCCAGCCATTCAAGG
Consensus		AATT...G.CTGGTTCCA.....A.
ZA1 (P)	-160	ACACCCGGCTTTTTTG
.OX1 (P)	-169	CAACCCGGCTTTTTTG
.OX2 (P)	-127	AAACCCCTTTTATG
IOX (H)	-24	ATCCCCAGTTTTTGC
.OD(C)	-694	AATTACTCTTTTTG
IAS (H)	-251	CGCAAGCCTTTTTGC
Consensus		CC.CTTTTTG

Fig. 12



.tsp HindIII
CATCATTATCAGCTTACTTTTCATAATTGTGACTGGTTCCAATTGACAAGCTTGC

Start codon Rice α -amylase Signal peptide
ATG CAG GTG CTG AAC ACC ATG GTG AAC AAA CAC TTC
MET GLN VAL LEU ASN THR MET VAL ASN LYS HIS PHE

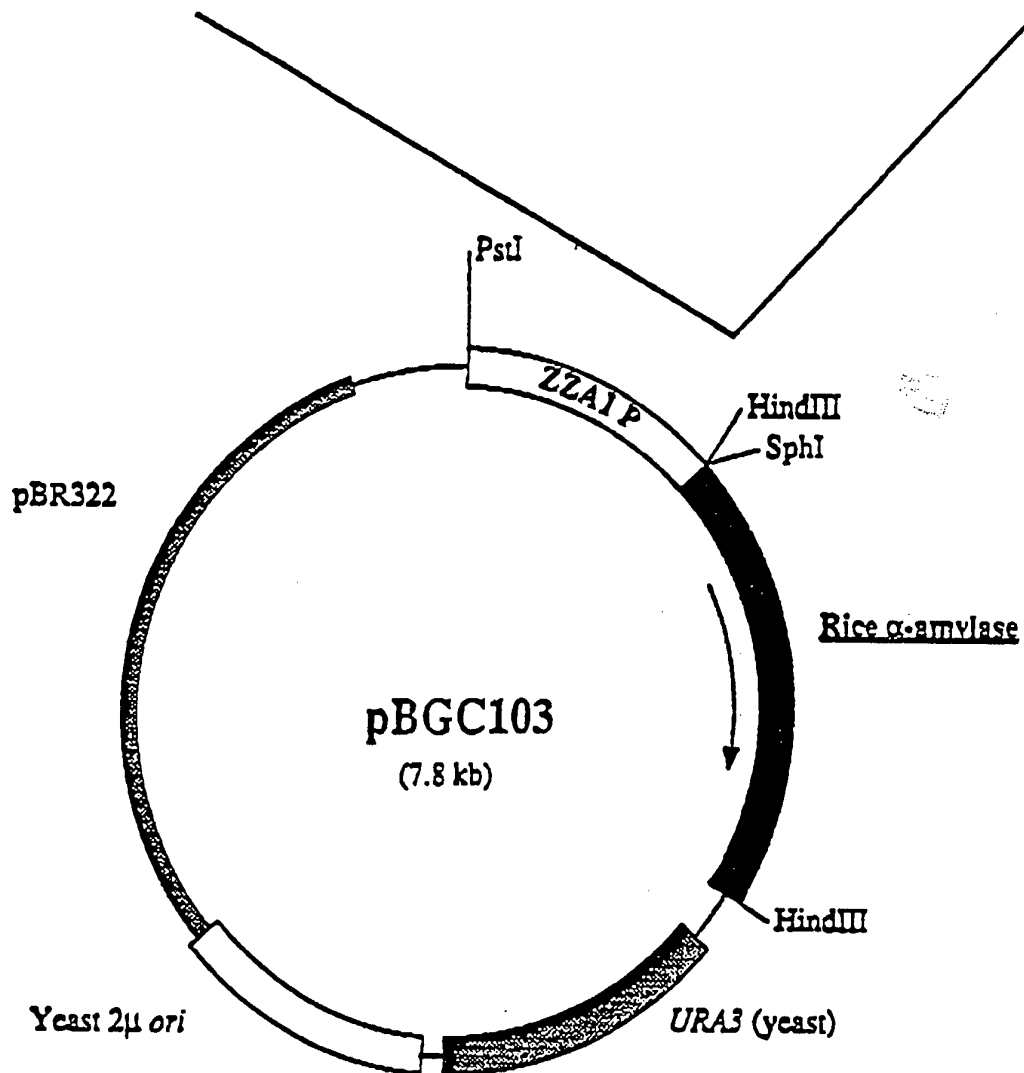


Fig. 14

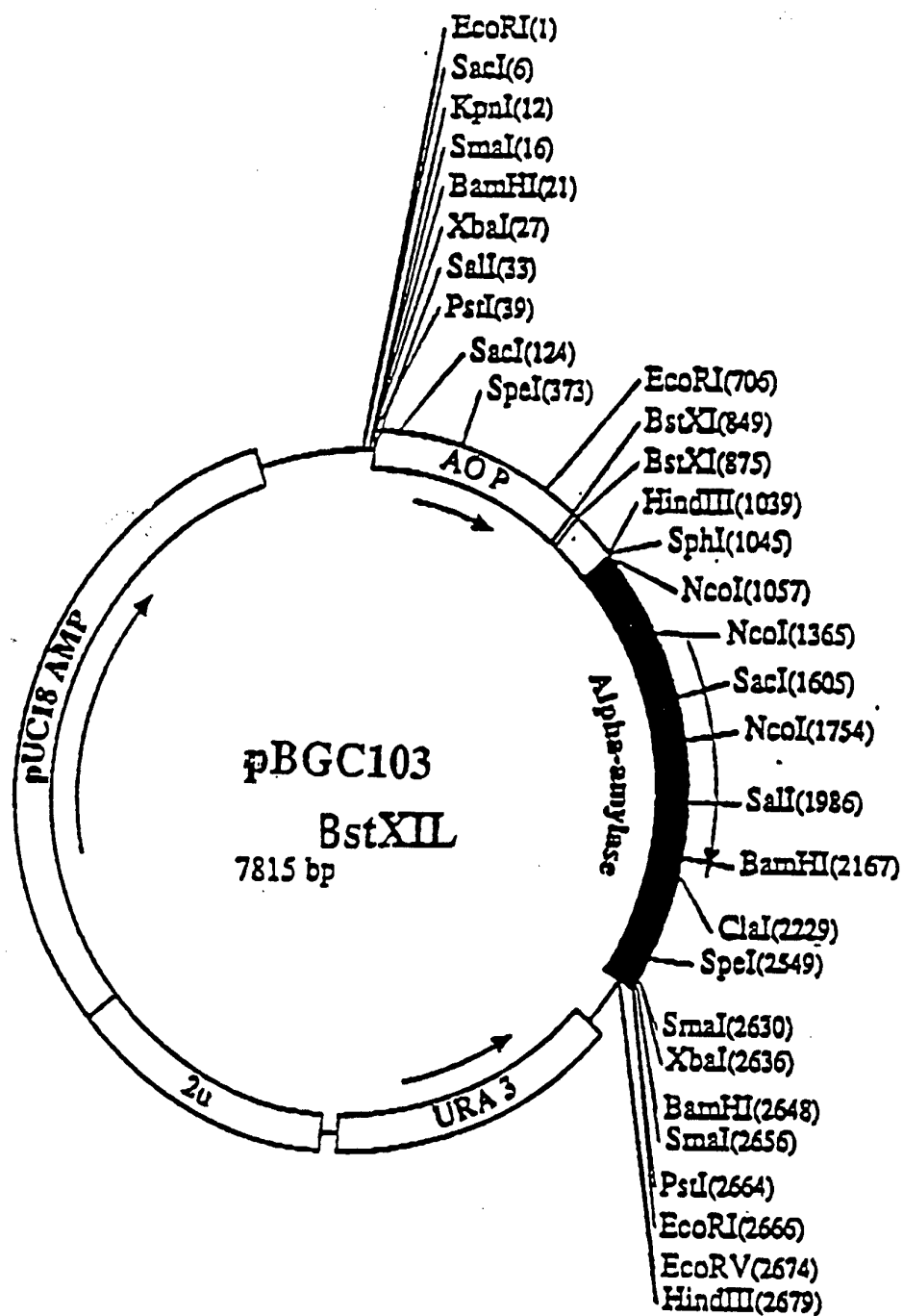


Fig. 15

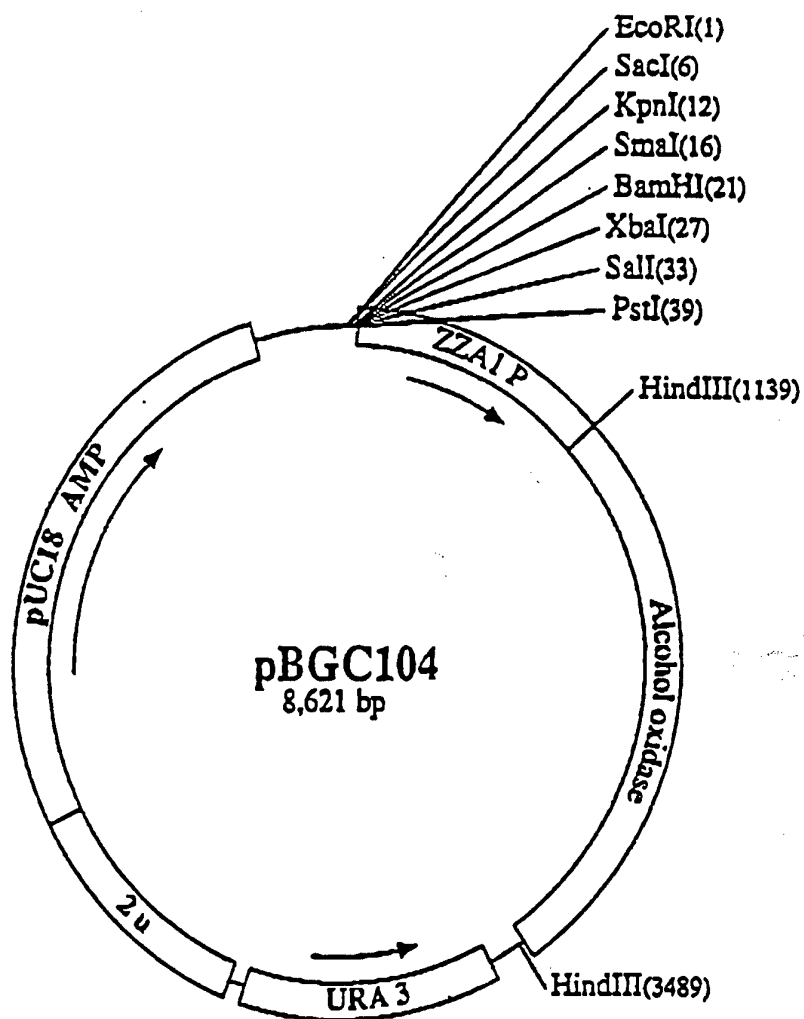


Fig. 16

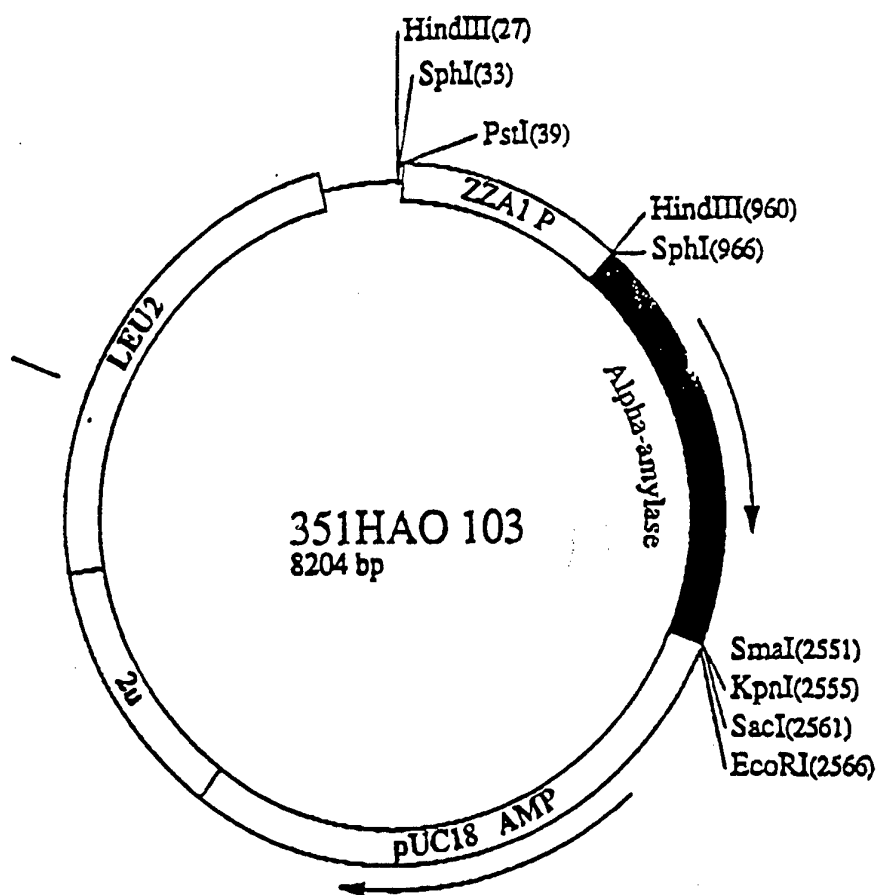
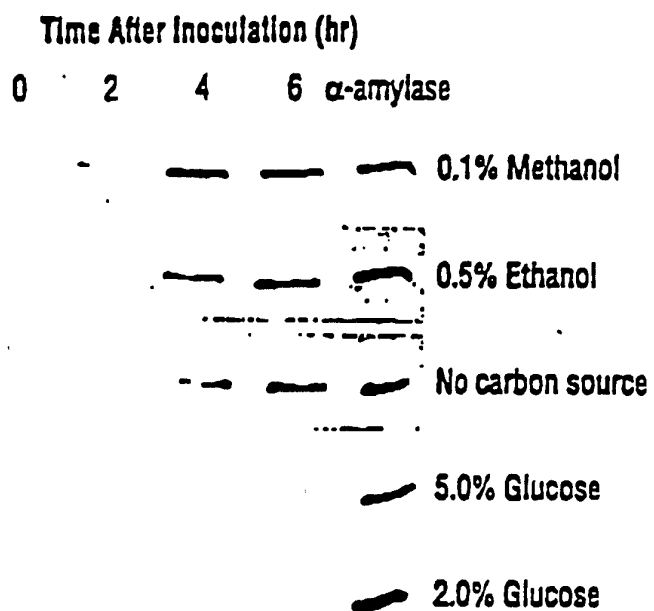


Fig. 17



A. Comparison of ZZA1 ZZA2, AOX1, and AOX2 3' regions:

ZZA2T	CAAAACCCCTCCCACTCTGCGCTGCTCAACACTTACGATACCTCTCGCTCATC	80
AOX2	CAAAAGACTGCCCTCTTGGCTTGGTTGAAGACTTAGGATACACAGC.TGAGT	
AOX1	CAAAAGACTGCCACTTTGCTTGGAGAGACTTAAAGATACCTCTGG.TGAGT	
	GLULysThrAlaThrLeuValGlyGluAspLeuGlyTyrSerGly GluA	
ZZA2T	CCCTCGACATGACTGTTCCCTCACTTCAAGTTCCGCACTTACCGCAAGACT	100
AOX2	CCTTAGACATGACTGTACCTCACTTCAAGTTCCGCACTTACCGAAGACT	
AOX1	CCTTAGACATGACTGTTCCCTCACTTCAAGTTCCGCACTTACCGAAGACT	
	IsLeuAspMetThrValIleGlnPheLysLeuGlyThrTyrGluLysThr	
ZZA2T	GGTCTTGCTAGATTCTAGTTTGTCTGTCTACGAAATATTTTTTAAATGA	150
AOX2	GGTCTTGCTAGATTCTAGTTTGTCTGTCTATGATATGATATTTTTTACAT	
AOX1	GGTCTTGCTAGATTCTATTCAGAGGATGTACGAAATGCAATTTTCCTGAG	
	GlyLeuAlaArgPheEnd	
ZZA2T	TTGATTTTTCTGTCACATTTGTTAGTTATTTATGTCGTTATAGTCACT	200
AOX2	GGTGGATTTGTTCCAACTAGTTGTTAGTTATTTGTCGTTATAGTCACT	
AOX1	AAATGCAATTTCTATTTTGTAGTTTATTTTATTTTAACTATATAGTAT	
ZZA2T	AAACAGTTTAGCTTCTTAGCAATTTTCTCTCTTACCAACAG.....	245
AOX2	AATACAGTTTACCATCGAATCTTTCTATTTCTTTTATACGTAAGTATAT	
AOX1	AGGATTTTTTTCTCATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
ZZA2T	TTCTCTGCACTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	
ZZA2T	ATACTAGATTAGAAATGCACTTACAGCGACTAATACCAACGGAGCTTGAT	295
AOX2	GTACTAGATGAAGATGCGGACAAAGCCGAAACAAACAAAGCTTCTCTCT	
AOX1	AGCCTATCTCCAGCTAATAAATACTTGTGGTAGGGGTTTGGGAAATC	
ZZA2T	AGTCTATCTTTGAGTTGCTGAATATCTTGAACATGGGCTTGAAGAAATC	
ZZA2T	ACCAAGGTTTACGAGGTTGCTACCGAATTCGCCATGATATCTCCCTAGAC	345
AOX2	ACCAAGGTTTGAAGGTTGCTACCGAATTCGCCATGATATTTGAAGTGGAG	
AOX1	ATTCAGGTTTGAATTTTCTTGGTATTTCCCACTCTCTCTCAGAGTACA	
ZZA2T	ATTTGATTTGAAAGTTTCTCTGCTAGTTGACATCTCTCTTCCGACTATA	

INTERNATIONAL SEARCH REPORT

Intern al Application No
PCT/US 94/03213

A. CLASSIFICATION OF SUBJECT MATTER

IPC 5 C12N15/81 C12N9/04 C12P7/10 C12N9/32 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 C12N C12P C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 183 071 (PHILLIPS PETROLEUM COMPANY) 4 June 1986 see the whole document ---	4,5
Y	EP,A,0 299 108 (RHEIN BIOTECH GESELLSCHAFT FÜR BIOTECHNOLOGISCHE PROZESSE UND PRODUKTE) 18 January 1989 see page 10, line 11 - page 10, line 30 ---	15,16, 19,20
Y	BIOTECHNOLOGY vol. 5 , December 1987 , NEW YORK US pages 1305 - 1308 J.F. TSCHOPP ET AL. 'Highlevel secretion of glycosylated invertase in the methylotrophic yeast, Pichia Pastoris' see the whole document --- -/--	15,16, 19,20

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

19 August 1994

Date of mailing of the international search report

31. 08. 94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Van Putten, A

INTERNATIONAL SEARCH REPORT

Intern al Application No

PCT/US 94/03213

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 257 115 (HEINEKEN TECHNISCH BEHEER B.V.) 2 March 1988 see claims 1-5 ---	15,16, 19,20
A	GENE vol. 94 , 13 November 1990 , AMSTERDAM NL pages 209 - 216 M.H. KUMAGAI ET AL. 'Expression and secretion of rice alpa-amylase by Saccharomyces cerevisiae' cited in the application ---	15-20
P,X	BIOTECHNOLOGY vol. 11 , May 1993 , NEW YORK US pages 606 - 610 M.H. KUMAGAI ET AL. 'Conversion of starch to ethanol in a recombinant Saccharomyces cerevisiae strain expressing rice alpha-amylase from a novel Pichia pastoris alcohol oxidase promoter' -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 94/03213

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0183071	04-06-86	US-A-	4808537	28-02-89
		US-A-	4855231	08-08-89
		AU-B-	572002	28-04-88
		AU-A-	4875285	12-06-86
		DE-A-	3586889	21-01-93
		EP-A-	0483115	29-04-92
		JP-A-	6098789	12-04-94
		JP-A-	61173781	05-08-86

EP-A-0299108	18-01-89	AT-T-	105858	15-06-94
		DE-D-	3789866	23-06-94

EP-A-0257115	02-03-88	AU-B-	621810	26-03-92
		AU-A-	7685087	16-06-88
		EP-A-	0260404	23-03-88
		JP-A-	63133984	06-06-88
		US-A-	5100794	31-03-92
		US-A-	5151354	29-09-92
		ZA-A-	8706134	29-02-88
