(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 January 2004 (08.01.2004)

PCT

(10) International Publication Number
WO 2004/003194 A2

(51) International Patent Classification[7]: C12N 9/24, 15/81, C12P 21/00, C12N 9/10

(21) International Application Number: PCT/EP2003/006711

(22) International Filing Date: 25 June 2003 (25.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/185,475 26 June 2002 (26.06.2002) US

(71) Applicant: **FLANDERS INTERUNIVERSITY INSTITUTE FOR BIOTECHNOLOGY** [BE/BE]; Rijvisschestraat 120, B-9052 Zwijnaarde (BE).

(72) Inventors: **CONTRERAS, Roland**; Molenstraat 53, B-9820 Merelbeke (BE). **CALLEWAERT, Nico, L., M.**; Ketelbuiserstraat 99, B-8810 Lichtervelde (BE). **GEYSENS, Steven, C., J.**; Steenbeekstraatje 8, B-9770 Kruishoutem (BE). **KAIGORODOV, Vladimir**; Ledeganckstraat 35, B-9000 Gent (BE). **WOUTER, Vervecken**; Vanhoorebekstraat 90, B-9040 Gent-Ledeberg (BE).

(74) Agents: **HOFFMANN EITLE** et al.; Arabellastrasse 4, 81925 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROTEIN GLYCOSYLATION MODIFICATION IN PICHIA PASTORIS

(57) Abstract: The present invention provides genetically engineered strains of methylotrophic yeast including *Pichia* and especially *Pichia pastoris* capable of producing proteins with reduced or modified glycosylation. Methods of producing glycoproteins with reduced and/or modified glycosylation using such genetically engineered strains of *Pichia* are also provided. Vectors, which comprise coding sequences for α-1,2-mannosidase I, glucosidase II, GlcNAc-tranferase I and mannosidase II or comprising OCH1 disrupting sequence, for transforming methylotrophic yeasts are contemplated by the present invention. Kit for providing the contemplated vectors are also included in this invention.

WO 2004/003194 A2

Protein Glycosylation Modification in *Pichia Pastoris*

Field of the Invention

The present invention relates to methods and vectors useful for genetically modifying the glycosylation process in methylotrophic yeast strains, including *Pichia* and especially *Pichia pastoris*, for the purpose of producing glycoproteins with reduced or modified glycosylation. The present invention further relates to methylotrophic yeast strains generated using the present methods and vectors, as well as glycoproteins produced from such genetically modified strains.

Background of the Invention

The methylotrophic yeasts including *Pichia pastoris* have been widely used for production of recombinant proteins of commercial or medical importance. Many pharmaceutical compositions consist glycoproteins produced in methylotrophic yeasts including *Pichia pastoris*. However, production and medical applications of some therapeutic glycoproteins can be hampered by the differences in the protein-linked carbohydrate biosynthesis between these yeasts and the target organisms such as a mammalian subject.

Protein N-glycosylation originates in the endoplasmic reticulum (ER), where the precursor N-linked oligosaccharide of 14 sugars ($Glc_3Man_9GlcNAc_2$) is assembled on a dolichol (a lipid carrier intermediate), and it is transferred to the appropriate Asn of growing nascent polypeptides. This is an event common to all eukaryotic N-linked glycoproteins. These glycans are then subjected to extensive modification as the glycoproteins mature and move through the ER via the Golgi complex to their final destinations inside and outside the cell. Three terminal glucose residues are trimmed away by glucosidase I and II, and one terminal α-1,2-linked mannose residue is removed by one or more different mannosidase in the ER, such as ER-mannosidase, resulting in the oligosaccharide $Man_8GlcNAc_2$. This glycoprotein is

then transported to the Golgi apparatus where the sugar moiety undergoes various modifications. There are significant differences in the modifications of the sugar branches in the Golgi apparatus between yeasts and higher eukaryotes.

In mammalian cells, the modification of the sugar branches in the Golgi apparatus proceeds via three different pathways depending on the protein moieties to which the sugars are added. They are, (1) where the glycoprotein does not change; (2) where the glycoprotein is modified by adding the N-acetylglucosamine-1-phosphate moiety (GlcNAc-1-P) in UDP-N-acetyl glucosamine (UDP-GlcNAc) to the 6-position of mannose in the sugar branch, followed by removing the GlcNAc moiety to form an acidic sugar branch in the glycoprotein; or (3) where the N-linked glycan is first converted into $Man_5GlcNAc_2$ by removing three mannose residues by Golgi mannosidase I; $Man_5GlcNAc_2$ is further modified by adding one GlcNAc by N-acetylglucosamine transferase I (GlcNAc-Transferase I or GnTI) and removing two more mannose residues by mannosidase II. During subsequent terminal glycosylation there is addition of new terminal sugars including GlcNAc, galactose (Gal), fucose, and N-acetylneuraminic acid (also called sialic acid (NeuNAc)) to produce various hybrid or complex glycans (R. Kornfeld and S. Kornfeld, *Ann. Rev. Biochem.* 54: 631-664, 1985; Chiba et al *J. Biol. Chem.* 273: 26298-26304, 1998; Helenius A and Aebi M, *Science* 291:2364-2369, 2001).

In yeasts, the modification of the sugar branches in the Golgi apparatus involves a series of additions of mannose residues by different mannosyltransferases ("outer chain" glycosylation). The structure of the outer chain glycosylation is specific to the organisms, typically with more than 50 mannose residues in *S. cerevisiae*, and most commonly with structures smaller than $Man_{14}GlcNAc_2$ in *Pichia pastoris*. This yeast-specific outer chain glycosylation of the high mannose type is also denoted hyperglycosylation.

Hyperglycosylation is often undesired since it leads to heterogeneity of a recombinant protein product in both carbohydrate composition and molecular weight,

which may complicate the protein purification. The specific activity (units/weight) of the hyperglycosylated proteins may be lowered by the increased portion of carbohydrate. In addition, the outer chain glycosylation is strongly immunogenic which is undesirable in a therapeutic application. Moreover, the large outer chain sugar can mask the immunogenic determinants of a therapeutic protein. For example, the influenza neuraminidase (NA) expressed in *P. pastoris* is glycosylated with N-glycans containing up to 30-40 mannose residues. The hyperglycosylated NA has a reduced immunogenicity in mice, as the variable and immunodominant surface loops on top of the NA molecule are masked by the N-glycans (Martinet et al. *Eur J. Biochem.* 247: 332-338, 1997).

Therefore, it is desirable to genetically engineer methylotrophic yeast strains in which glycosylation of proteins can be manipulated and from which recombinant glycoproteins can be produced having a mammalian-like glycosylation pattern.

Summary of the Invention

The present invention is directed to methods and vectors useful for genetically modifying the glycosylation process in methylotrophic yeast strains to produce glycoproteins with reduced or modified glycosylation. Methylotrophic yeast strains generated using present methods and vectors, as well as glycoproteins produced from such genetically modified strains, are also provided.

In one embodiment, the present invention provides vectors useful for making genetically engineered methylotrophic yeast strains which are capable of producing glycoproteins with reduced or modified glycosylation.

In one aspect, the present invention provides "knock-in" vectors which are capable of expressing in a methylotrophic yeast strain one or more proteins whose enzymatic activities lead to a reduction or modification of glycosylation in glycoproteins produced by the methylotrophic yeast strain.

In a preferred embodiment, the knock-in vectors of the present invention include a nucleotide sequence coding for an α-1,2-mannosidase or a functional part thereof and are capable of expressing the α-1,2-mannosidase or the functional part in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the α-1,2-mannosidase of a fungal species, and more preferably, *Trichoderma reesei*. Preferably, the α-1,2-mannosidase expression vector is engineered such that the α-1,2-mannosidase or a functional part thereof expressed from the vector includes an ER-retention signal. A preferred ER-retention signal is HDEL. The α-1,2-mannosidase coding sequence can be operably linked to a constitutive or an inducible promoter, and a 3' termination sequence. The vectors can be integrative vectors or replicative vectors. Particularly preferred α-1,2-mannosidase expression vectors include pGAPZMFManHDEL, pGAPZMFManMycHDEL, pPICZBMFManMycHDEL, pGAPZmManHDEL, pGAPZmMycManHDEL, pPIC9mMycManHDEL and pGAPZmMycManHDEL.

In another preferred embodiment, the knock-in vectors of the present invention include a sequence coding for a glucosidase II or a functional part thereof and are capable of expressing the glucosidase II or the functional part in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the glucosidase II of a fungal species, and more preferably, *Saccharomyces cerevisiae*. Preferably, the glucosidase II expression vector is engineered such that the glucosidase II or a functional part thereof expressed from the vector includes an ER-retention signal. A preferred ER-retention signal is HDEL. The glucosidase II coding sequence can be operably linked to a constitutive or an inducible promoter, and a 3' termination sequence. The vectors can be integrative vectors or replicative vectors. Particularly preferred glucosidase II expression vectors include pGAPZAGLSII, pPICZAGLSII, pAOX2ZAGLSII, pYPTIZAGLSII, pGAPADEglsII, pPICADEglsII, pAOX2ADEglsII, pYPTIADEglsII, pGAPZAglsIIHDEL and pGAPADEglsIIHDEL.

Expression vectors which include both of an α-1,2-mannosidase expression unit and a glucosidase II expression unit are also provided by the present invention.

By "expression unit" is meant that a nucleotide sequence capable of expressing a gene of interest. In general, an express unit includes the gene to be expressed, which is operably linked to a promoter, a termination sequence and any other sequence that may be appropriate.

In a preferred embodiment, the knock-in vectors of the present invention include a nucleotide sequence encoding a N-acetylglucosamine transferase I (GlcNAc-transferase I or GnTI) or a functional part thereof and are capable of expressing the GlcNAc-transferase I or the functional part in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the GlcNAc-transferase I of a mammalian species, e.g., human. Preferably, the GlcNAc-transferase I expression vector is engineered such that the GlcNAc-transferase I or a functional part thereof expressed from the vector includes a yeast Golgi localization signal. A preferred yeast Golgi localization signal is a peptide derived from the N-terminal part of *S. cerevisiae* Kre2 (ScKre2). The GlcNAc-transferase I coding sequence can be operably linked to a constitutive or an inducible promoter, and a 3' termination sequence. The vectors can be integrative vectors or replicative vectors. Particularly preferred GlcNAc-transferase I expression vectors include pPIC6AKreconGnTI.

In another preferred embodiment, the knock-in vectors of the present invention include a sequence encoding a mannosidase II or a functional part thereof and are capable of expressing the mannosidase II or the functional part thereof in a methylotrophic yeast strain. A preferred nucleotide sequence is a nucleotide sequence encoding the mannosidase II of a mammalian species, e.g., human. Preferably, the mannosidase II expression vector is engineered such that the mannosidase II or a functional part thereof expressed from the vector includes a yeast Golgi localization signal. A preferred Golgi localization signal is a peptide derived from the N-terminal part of ScKre2. The mannosidase II coding sequence can be operably linked to a

constitutive or an inducible promoter, and a 3' termination sequence. The vectors can be integrative vectors or replicative vectors. Particularly preferred mannosidase II expression vectors include pGAPKreManII.

Expression vectors which include both of a GlcNAc-transferase I expression unit and a mannosidase II expression unit are also provided by the present invention.

Expression vectors which include two or more of an α-1,2-mannosidase expression unit, a glucosidase II expression unit, a GlcNAc-transferase I expression unit and a mannosidase II expression unit, are also provided by the present invention.

In another aspect, the present invention provides inactivation vectors which, when introduced into a methylotrophic yeast strain, inactivate or disrupt a gene thereby facilitating the reduction in the glycosylation of glycoproteins produced in the methylotrophic yeast strain.

In one embodiment, the present invention provides an inactivation vector which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the OCH1 gene. A preferred OCH1 inactivation vector is pBLURA5'PpOCH1and pZMFManHDEL5'PpOch1Mut.

Still another embodiment of the present invention provides "knock-in-inactivation" vectors which include at least a "knock-in unit" and at least a "inactivation unit". By "knock-in unit" is meant an expression unit which is capable of expressing in a methylotrophic yeast strain one or more proteins whose enzymatic activities lead to a reduction or modification of glycosylation in glycoproteins produced by the methylotrophic yeast strain. By "inactivation unit" is meant an expression unit which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the endogenous OCH1 gene or other mannosyltransferase genes. When a "knock-in-and inactivation" vector is introduced to a methylotrophic yeast strain, potent exdogenous enzyme expression and endogenous mannosyltransferase

disruption can be achieved simultaneously. A preferred knock-in-and-inactivation vector is pZMFManHDEL5'PpOCH1Mut.

Another embodiment of the present invention provides methods of modifying the glycosylation in a methylotrophic yeast strain by transforming the yeast with one or more vectors of the present invention.

Strains of a methylotrophic yeast which can be modified using the present invention include, but are not limited to, yeast strains capable of growing on methanol, such as yeasts of the genera *Candida, Hansenula, Torulopsis,* and *Pichia.* Preferred methylotrophic yeasts are of the genus *Pichia.* Especially preferred are *Pichia pastoris* strains GS115 (NRRL Y-15851), GS190 (NRRL Y-18014), PPF1 (NRRL Y-18017), PPY120H, YGC4, and strains derived therefrom. Methylotrophic yeast strains which can be modified using the present methods also include those methylotrophic yeast strains which have been engineered to express one or more heterologous proteins of interest. The glycosylation on the heterologous proteins expressed from these previously genetically engineered strains can be reduced by transforming such strains with one or more of the vectors of the present invention.

Methylotrophic yeast strains which are modified by practicing the present methods are provided in another embodiment of the present invention.

A further aspect of the present invention is directed to methods of producing glycoproteins with reduced or modified glycosylations.

In accordance with such methods, a nucleotide sequence capable of expressing a glycoprotein can be introduced into a methylotrophic yeast strain which has previously been transformed with one or more of the vectors of the present invention. Alternatively, a methylotrophic yeast strain which has been genetically engineered to express a glycoprotein can be transformed with one or more of the vectors of the present invention. Moreover, if a methylotrophic yeast strain is not transformed with a nucleotide sequence encoding a glycoprotein of interest or any of the vectors of the present invention, such yeast strain can be transformed, either

consecutively or simultaneously, with both a nucleotide sequence capable of expressing a glycoprotein and one or more vectors of the present invention. Additionally, a methylotrophic yeast strain can be transformed with one or more of the present knock-in and/or inactivation vectors which also include a nucleotide sequence capable of expressing a glycoprotein in the methylotrophic yeast strain.

Glycoprotein products produced by using the methods of the present invention, i.e., glycoproteins with reduced or modified N-glycosylation, are also part of the present invention.

Kits which include one or more of the vectors of the present invention, or one or more strains modified to produce glycoproteins with reduced or modified glycosylation, are also provided.

Brief Description of the Drawings

Figure 1 depicts vectors carrying an HDEL-tagged *Trichoderma reesei* α-1,2-mannosidase expression cassette and describes the way in which these vectors were constructed according to methods known in the art. Abbreviations used throughout construction schemes: 5' AOX1 or AOX1 P: *Pichia pastoris* AOX1 promoter sequence; Amp R: ampicillin resistance gene; ColE1: ColE1 origin of replication; 3'AOX1: 3' sequences of the *Pichia pastoris* AOX1 gene; HIS4: HIS4 gene of *Pichia pastoris*. AOX TT: transcription terminator sequence of the *Pichia pastoris* AOX1 gene; ORF: open reading frame; S: secretion signal; P TEF1: the promoter sequence of the *Saccharomyces cerevisiae* transcription elongation factor 1 gene; P EM7: synthetic constitutive prokaryotic promotor EM7; Zeocin: Zeocin resistance gene; CYC1 TT: 3' end of the *S. cerevisiae* CYC1 gene; GAP: promoter sequence of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene; PpURA3: *Pichia* pastoris URA3 gene. As can be seen in this figure, the *Trichoderma reesei* α-1,2-mannosidase was operably linked to the coding sequence for the *S. cerevisiae* α-mating factor secretion signal sequence and further operably linked at the

3' terminus of the coding sequence to the coding sequence for an HDEL peptide. The whole fusion construct was operably linked to either the *P. pastoris* AOX1 promoter (in pPIC9MFManHDEL) or to the *P. pastoris* GAP promotor (in pGAPZMFManHDEL).

Figure 2 depicts vectors carrying an HDEL-tagged *Mus musculus* α-1,2-mannosidase IB expression cassette and describes the way in which these vectors were constructed according to methods known in the art. As can be seen in this figure, the catalytic domain of the *Mus musculus* α-1,2-mannosidase IB was operably linked to the coding sequence for the *S. cerevisiae* α-mating factor secretion signal sequence and further operably linked at the 3' terminus of the coding sequence to the coding sequence for an HDEL peptide. The whole fusion construct was operably linked to either the *P. pastoris* AOX1 promoter (in pPIC9mManHDEL) or to the *P. pastoris* GAP promotor (in pGAPZmManHDEL). Furthermore, variants of the expression cassette were made in which the coding sequence for a cMyc epitope tag was inserted between the coding sequence for the *S. cerevisiae* α-mating factor secretion signal sequence and the coding sequence for the catalytic domain of the *Mus musculus* α-1,2-mannosidase IB. This expression cassette was also operably linked to either the *P. pastoris* AOX1 promoter (in pPIC9mMycManHDEL) or to the *P. pastoris* GAP promotor (in pGAPZmMycManHDEL).

Figure 3 depicts vectors carrying a MycHDEL tagged *Trichoderma reesei* α-1,2-mannosidase and the way in which these vectors were obtained. The resulting fusion construction was again operably linked to either the *P. pastoris* AOX1 promoter (in pPICZBMFManMycHDEL) or to the *P. pastoris* GAP promotor (in pGAPZMFManMycHDEL).

Figure 4 demonstrates the intracellular localization of the MycHDEL-tagged *Trichoderma reesei* α-1,2-mannosidase and indicates ER-targeting by immunofluorescence analysis. ***Panel A*** Western blotting. Yeast strains were grown in

10 ml YPG cultures to an OD_{600}=10, diluted fivefold and grown in YPM for 48 h. 1/50th of the culture medium and 1/65th of the cells were analysed by SDS-PAGE and Western blotting with the mouse monoclonal 9E10 anti-Myc antibody. The position of molecular weight marker proteins are indicated with arrows. Lanes 1-5: cellular lysates. 1 and 2: pGAPZMFManMycHDEL transformants. 3: non-transformed PPY12OH (negative control). 4 and 5: pPICZBMFManMycHDEL transformants. Lanes 6-10: culture media. 6: non transformed PPY12OH (negative control). 7 and 8: pGAPZMFManMycHDEL transformants. 9 and 10: pPICZBMFManMycHDEL transformants. ***Panel B*** Immunofluorescence microscopy. 1: phase contrast image of a *P. pastoris* cell (strain PPY12OH transformed with pGAPZMFManHDEL) at 1000x magnification. The nucleus is visible as an ellipse in the lower right quadrant of the cell. 2: same cell as in 1, but in fluorescence microscopy mode to show localization of the *T. reesei* mannosidase-Myc-HDEL protein. The protein is mainly localized in a circular distribution around the nucleus (nuclear envelope), which is typical for an endoplasmic reticulum steady-state distribution. 3: phase contrast image of a *P. pastoris* cell (strain PPY12OH transformed with pGAPZMFManHDEL) at 1000x magnification. 4: same cell in fluorescence microscopy to show localization of the Golgi marker protein OCH1-HA in *P. pastoris* strain PPY12OH. The dot-like distribution throughout the cytoplasm, with 3-4 dots per cell is typical for cis-Golgi distribution in *P. pastoris.*

Figure 5 depicts the co-sedimentation of mannosidase-MycHDEL with Protein Disulfide Isomerase in sucrose density gradient centrifugation. The top panel shows the distribution over the different fractions of the sucrose gradient of the OCH1-HA Golgi marker protein. The middle panel shows this distribution for the Protein Disulfide Isomerase endoplasmic reticulum marker protein. Finally, the bottom panel shows the distribution of the MycHDEL-tagged *Trichoderma reesei* α-1,2-mannosidase over the same fractions. It is concluded that the mannosidase-

MycHDEL almost exactly matches the distribution of the ER marker PDI and thus mainly resides in the ER of the *Pichia pastoris* yeast cells.

Figure 6 depicts the N-glycan analysis of *Trypanosoma cruzi trans*-sialidase coexpressed with *Trichoderma reesei* mannosidase-HDEL. Panel A: malto-oligosaccharide size reference ladder. Sizes of the glycans are expressed in Glucose Units (GU) by comparison of their electrophoretic mobility to the mobility of these malto-oligosaccharides. Panel B: N-glycans derived from recombinant *Trypanosoma cruzi trans*-sialidase expressed in *Pichia pastoris*. The peak at GU=9,2 corresponds to $Man_8GlcNAc_2$. Panel C: same analytes as panel 2, but after overnight treatment with 3U/ml purified recombinant *T. reesei* α-1,2-mannosidase. Panel D: N-glycans derived from recombinant *trans*-sialidase co-expressed in *Pichia pastoris* with *T. reesei* mannosidase-HDEL (under control of the GAP promotor). The peak at GU=7,6 corresponds to the $Man_5GlcNAc_2$ peak in the profile of RNase B (Panel F). Panel E: same analytes as panel D, but after overnight treatment with 3 mU/ml purified recombinant *T. reesei* α-1,2-mannosidase. Panel F: N-glycans derived from bovine RNase B. These glycans consist of $Man_5GlcNAc_2$ to $Man_8GlcNAc_2$. Different isomers are resolved, accounting for the number of peaks for $Man_7GlcNAc_2$.

Figure 7 depicts the processing of influenza haemagglutinin N-glycans by HDEL-tagged *Trichoderma reesei* α-1,2-mannosidase and the HDEL-tagged catalytic domain of murine α-1,2-mannosidase IB. The $Man_5GlcNAc_2$ reference oligosaccharide runs at scan 1850 in this analysis (not shown). Panel 1: malto-oligosaccharide size reference ladder. Panel 2: N-glycans derived from recombinant influenza haemagglutinin expressed in *Pichia pastoris*. The peak at scan 2250 corresponds to $Man_9GlcNAc_2$. Panel 3: N-glycans derived from recombinant haemagglutinin co-expressed in *Pichia pastoris* with *T. reesei* mannosidase-HDEL (under control of the GAP promotor). The peak at scan 1950 corresponds to $Man_6GlcNAc_2$. Panel 4: Same analytes as for panel 3, but after overnight treatment with 30 mU purified recombinant *T. reesei* α-1,2-mannosidase. Panel 5: N-glycans

derived from recombinant haemagglutinin co-expressed in *Pichia pastoris* with mouse mannosidase IB-HDEL (under control of the GAP promotor). Panel 6: same analytes as for panel 5, but after overnight treatment with 30 mU purified recombinant *T. reesei* α-1,2-mannosidase.

Figure 8 graphically depicts vector pBLURA5'PpOCH1 and the way in which it was constructed.

Figure 9 depicts the scheme for disrupting the *Pichia pastoris* OCH1 gene by single homologous recombination using pBLURA5'PpOCH1.

Figure 10 depicts the cell wall glycoprotein N-glycan analysis of the OCH1-inactivated clone and three clones derived from this OCH1-inactivated clone by transformation with pGAPZMFManHDEL. Panel 1 shows the analysis of a mixture of malto-oligosaccharides, the degree of polymerisation of which is given by the numbers on the very top of the figure. This analysis serves as a size reference for the other panels. On the vertical axis of all panels, peak intensity in relative fluorescence units is given. Panel 2-6: analysis of the cell wall glycoprotein N-glycans of the following strains: Panel 2, non-engineered *P. pastoris* strain YGC4; Panel 3, YGC4 transformed with pBLURA5'PpOCH1; 4-6, three clones of the strain of Panel 3, supplementarily transformed with pGAPZMFManHDEL. Panel 7: the N-glycans derived from bovine RNaseB, consisting of a mixture of $Man_{5-9}GlcNAc_2$. As can be seen from comparison between panel 2 and 3 and reference to panel 7, transformation with pBLURA5'PpOCH1 leads to a strongly increased abundance of the $Man_8GlcNAc_2$ substrate N-glycan (named peak 1 in Panel 2) of OCH1p. Peak 2 represents the $Man_9GlcNAc_2$ product of OCH1p. Furthermore, upon supplementary transformation of pGAPZMFManHDEL, the major glycan on the cell wall glycoproteins of three independent clones is the $Man_5GlcNAc_2$ end product (peak 3 in panel 4) of *T. reesei* α-1,2-mannosidase digestion of the $Man_8GlcNAc_2$ substrate.

Figure 11 depicts the analysis of exactly the same glycan mixtures as in Figure 10, but after an *in vitro* digest with 3mU/ml purified *Trichoderma reesei* α-1,2-

mannosidase, overnight in 20 mM sodium acetate pH=5.0. Axis assignment is the same as in Figure 10. More $Man_5GlcNAc_2$ is formed in the pBLURA5'PpOCH1 transformed strain (Panel 3) than in the parent strain (Panel 2). Peaks in all panels before scan 3900 come from contaminants and should be ignored in the analysis.

Figure 12 depicts the expression vector pGAPZAGLSII (SEQ ID NO: 18). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 13 depicts the expression vector pAOX2ZAGLSII (SEQ ID NO: 16). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. AOX2 P: promotor of the *P. pastoris* AOX2 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 14 depicts the expression vector pPICZAGLSII (SEQ ID NO: 20). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: origin of replication. AOX1 P: promotor of the *P. pastoris* AOX1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 15 depicts the expression vector pYPT1ZAGLSII ((SEQ ID NO: 22). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7: synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: origin of replication. P YPT1: promotor of the *P. pastoris* YPT1 gene. GLS2: *S. cerevisiae*

glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 16 depicts the expression vector pGAPADE1glsII (SEQ ID NO: 19). Amp R: Ampillicin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene

Figure 17 depicts the expression vector pAOX2ADE1glsII (SEQ ID NO: 17). Amp R: Ampillicin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. AOX2 P: promotor of the *P. pastoris* AOX2 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 18 depicts the expression vector pPICADE1glsII (SEQ ID NO: 21). Amp R: Ampillicin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. AOX1 P: promotor of the *P. pastoris* AOX1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 19 depicts the expression vector pYPT1ADE1glsII (SEQ ID NO: 23). Amp R: Ampillicin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. P YPT1: promotor of the *P. pastoris* YPT1 gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 20 depicts the expression vector pGAPADE1glsIIHDEL (SEQ ID NO: 25). Amp R: Ampillicin resistance marker gene. ADE1: *P. pastoris* ADE1 selection marker gene. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 21 depicts the expression vector pGAPZAglsIIHDEL (SEQ ID NO: 24). P TEF1: promotor of *S. cerevisiae* transcription elongation factor gene. P Em7:

synthetic prokaryotic promotor. Zeocin: zeocine resistance marker gene. CYC1 TT: transcription terminator of *S. cerevisiae* cytochrome C1 gene. Col E1: bacterial origin of replication. GAP: promotor of the *P. pastoris* GAP gene. GLS2: *S. cerevisiae* glucosidase II gene. AOX1 TT: transcription terminator of the *P. pastoris* AOX1 gene.

Figure 22 depicts the test of the GLSII activity assay using a commercially available yeast alpha-glucosidase (Sigma: Cat. No. G-5003). The assay mixture contains phosphate-citrate buffer pH 6.8, mannose, 2-deoxy-D-glucose, the substrate 4-methylumbellyferyl-alpha-D-glucopyranoside and alpha-glucosidase from Sigma. 1: assay mixture illuminated with UV-light after overnight incubation at 37 ° C; 2: same as 1, but this time, the assay mixture lacks the alpha-glucosidase; 3: same as 1, but this time, the assay mixture lacks the substrate.

Figure 23 depicts the results of the activity of recombinantly expressed GLSII from *Pichia pastoris*. All assay mixtures were incubated overnight at 37 °C and afterwards illuminated with UV-light. 1: assay with yeast alpha-glucosidase (Sigma: Cat. No. G-5003); 2: assay with the purified medium of strain 18 (PPY12-OH transformed with pGAPZAGLSII); 3: assay with purified medium of the WT PPY12-OH strain; 4: assay with the purified medium of strain H3 (PPY12-OH transformed with pGAPZAglsIIHDEL).

Figure 24 depicts vectors carrying a ScKre2-tagged human GlcNAc-transferase I (GnTI) expression cassette and describes the way in which these vectors were constructed according to methods known in the art. Abbreviations used throughout the construction schemes: 5' AOX1 or AOX1 P: *Pichia pastoris* AOX1 promoter sequence; Amp R: ampicillin resistance gene; ColE1: ColE1 origin of replication; 3'AOX1: 3' sequences of the *Pichia pastoris* AOX1 gene; AOX TT: transcription terminator sequence of the *Pichia pastoris* AOX1 gene; ORF: open reading frame; Blastisidin R: blastisidin resistance gene; CYC1 TT: 3' end of the *S. cerevisiae* CYC1 gene. As can be seen in this figure, the human GnTI was operably

linked at the 5' terminus of the coding sequence to the coding sequence for a peptide derived from the N-terminal part of ScKre2. The whole fusion construct was operably linked to the *P. pastoris* AOX1 promoter.

Figure 25 depicts an analysis of native N-glycans from proteins secreted by *P. pastoris* transformants with GlcNAc-transferase I, separated by means of DSA-FACE. Panel 1 shows the analysis of Dextran. This analysis serves as a size reference for the other panels. On the vertical axis of all panels, peak intensity in relative fluorescence units is given. Panel 2-5 (T1-T4) shows the results from the analysis of the N-glycans from transformants 1-4. The analysis on a WT or untransformed organism is shown in Panel 6.

Figure 26 depicts an analysis of N-glycans from secreted proteins of *P. pastoris* transformants with GlcNAc-transferase I digested with Jack Bean beta-N-acetylhexosaminidase.

Figure 27 depicts an analysis of N-glycans from Transsialidase (TS) produced in three different strains. Panel 1: oligomaltose reference. Panel 2: Glycans from TS produced in GS115. The two main peaks are from $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$. While TS produced in GS115 which was transformed with MannosidaseHDEL presented mainly $Man_5GlcNAc_2$, there were still some $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$ peaks, as shown in panel 3. But glycans from TS produced in GS115 strain which was transformed with both MannosidaseHDEL and Kre2-GlcNAc-transferase I showed that almost all $Man_5GlcNAc_2$ was converted to $GlcNAcMan_5GlcNAc_2$ (panel 4).

Figure 28 depicts a lectin blot of Transsialidase (TS). Above: lectin blot of TS from 3 different strains. Lane 1: TS from GS115, lane 2: TS from GS115 ManHDEL, lane 3: GS115 ManHDEL Kre2GnTI. Below: Same blot stained with Ponceau red. Only the band in lane 3 is stained by the lectin, indicating that only there terminal GlcNAc is present.

Figure 29 depicts an N-glycan analysis of a GnTI-transformant treated with Jack Bean β-N-acetylhexosaminidase.

Figure 30 depicts vectors carrying a SrKre2-tagged human mannosidase II expression cassette and describes the way in which these vectors were constructed. Abbreviations used throughout the construction schemes: 5' AOX1 or AOX1 P: *Pichia pastoris* AOX1 promoter sequence; Amp R: ampicillin resistance gene; ColE1: ColE1 origin of replication; 3'AOX1: 3' sequences of the *Pichia pastoris* AOX1 gene; ARG4: ARG4 gene of *Pichia pastoris*. AOX TT: transcription terminator sequence of the *Pichia pastoris* AOX1 gene; ORF: open reading frame; P EM7: synthetic constitutive prokaryotic promotor EM7; Zeocin: Zeocin resistance gene; CYC1 TT: 3' end of the *S. cerevisiae* CYC1 gene; GAP: promoter sequence of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene; PpURA3: *Pichia pastoris* URA3 gene. As can be seen in this figure, the human mannosidase II is operably linked at the 5' terminus of the coding sequence to the coding sequence for a peptide derived from the N-terminal part of ScKre2. The whole fusion construct was operably linked to the *P. pastoris* GAP promoter.

Figure 31 depicts a measurement of intracellular class I mannosidase activity in *Pichia pastoris* cells transformed with pGAPZMFManHDEL. The hydrolysis of DNP-Man was measured by spectrophotometry at 405 nm. In all transformants, a three-to fourfold increase in activity was measured as compared to the GS115 parent strain.

Figure 32 depicts the scheme of the genomic rearrangement upon transformation with pZ5'PpOCH1Trunc. pZ5'PpOCH1Trunc contains the promoterless coding sequence for amino acid 25-155 of *Pichia* OCH1, followed by the 3' AOXI transcription terminator. This vector is linearized with BstBI in the OCH1-homologous part, inducing specific homologous recombination into the OCH1 locus. After integration, two OCH1 sequence parts are generated: the first OCH1 sequence consists of the genomic OCH1 sequence up to the site of recombination with the

vector, followed by the portion of OCH1 sequence that is present on the vector. This first OCH1 sequence is now under control of the OCH1 promoter of the host methylotrophic yeast, yet cannot produce an active OCH1 protein as such OCH1 sequence codes for no or an inactive fragment of the OCH1 protein, as described hereinabove. The second OCH1 sequence is a truncated OCH1 coding sequence (lacks the first 24 amino acids). Moreover this truncated OCH1 coding sequence is not operably linked to any known promoter sequence and thus, no messenger is expected to be formed for synthesis of an active OCH1 protein.

Figure 33 depicts the expression vector pZMFManHDEL5'PpOCH1Mut (SEQ ID NO: 46)

Figure 34 depicts genomic analysis of GS115 clones transformed with pZ5'PpOCH1Trunc. For clones that have correctly integrated the plasmid in the OCH1 locus, the expected product length for the primer couple Kai10/Kai 7 is 1050 bp, whereas for the primer couple Kai11/Kai 5, it is 2100 bp. It is clear that clones 1,2,3 and 8 fulfill this requirement. We conclude that the OCH1 locus has been disrupted in these clones.

Detailed Description of the Invention

It has been established that the majority of N-glycans on glycoproteins leaving the endoplasmic reticulum (ER) of methylotrophic yeasts, including *Pichia* and especially *Pichia pastoris*, have the $Man_8GlcNAc_2$ oligosaccharide structure. After the proteins are transported from the ER to the Golgi apparatus, additional mannose residues are added to this core sugar moiety by different mannosyltransferases, resulting in glycoproteins with oligosaccharide structures consisting of a high manose core, or extended branched mannan outer chains. Such hyperglycosylation of recombinant glycoproteins is undesirable in many instances. Accordingly, the present invention provides methods and vectors for genetically modifying methylotrophic yeast strains to produce glycoproteins with reduced or

modified glycosylation. Methylotrophic yeast strains generated using the present methods and vectors, as well as glycoproteins produced from such genetically modified strains are also provided.

By the term "reduced or modified glycosylation", it is meant that a host yeast strain used for the production of a protein of interest has been genetically modified, such that the protein expressed from the genetically modified strain bears oligosaccharide side branches with fewer mannose residues, as compared to a protein from unmodified yeast strains.

Yeast strains genetically engineered in accordance with the present invention are also capable of producing proteins with oligosaccharide side branches that are "mammalian-like". By the term "mammalian-like", it is meant that the oligosaccharide is of a structure ordinarily seen on a mammalian glycoprotein, e.g., glycoproteins comprising glycans with fewer terminal mannose residues or glycans with a hybrid- or a complex-type glycosylation pattern, as compared to glycans with high terminal mannose residues on yeast glycoproteins. Typical mammalian-like oligosaccharide structures start from $Man_5GlcNAc_2$, $GlcNAcMan_5GlcNAc_2$ and $GlcNAcMan_3GlcNAc_2$. Branched structures, e.g., bi-, tri, and tetra-antennary, are synthesized by the GlcNAc-transferase-catalyzed addition of GlcNAc to regions of the oligosaccharide residue. Subsequent to their formation, the antennary structures are terminated with different sugars including Gal, GalNAc, GlcNAc, Fuc and sialic acid residues.

In one embodiment, the present invention provides vectors useful for genetically modifying methylotrophic yeast strains to produce glycoproteins with reduced or modified glycosylation.

In one aspect, the present invention provides "knock-in" vectors which are capable of expressing in a methylotrophic yeast strain one or more proteins whose enzymatic activities lead to a reduction or modification of glycosylation in glycoproteins produced by the methylotrophic yeast strain. According to the present

invention, such proteins include, e.g., an α-1,2-mannosidase, a glucosidase II, a GlcNAc-transferase I (GnTI), a mannosidase II, or functional parts thereof.

In a preferred embodiment, the vectors of the present invention include a sequence coding for an α-1,2-mannosidase or a functional part thereof and are capable of expressing the α-1,2-mannosidase or the functional part thereof in methylotrophic yeast.

An α-1,2-mannosidase cleaves the α-1,2-linked mannose residues at the non-reducing ends of $Man_8GlcNAc_2$, and converts this core oligosaccharide on glycoproteins to $Man_5GlcNAc_2$. *In vitro*, $Man_5GlcNAc_2$ is a very poor substrate for any *Pichia* Golgi mannosyltransferase, i.e., mannose residues can not be added to this sugar structure. On the other hand, $Man_5GlcNAc_2$ is the acceptor substrate for the mammalian N-acetylglucosaminyl-transferase I and is an intermediate for the hybrid- and complex-type sugar chains characteristic of mammalian glycoproteins. Thus, by way of introducing an α-1,2-mannosidase into methylotrophic yeasts such as *Pichia*, glycoproteins with reduced mannose content glycans, such as $Man_5GlcNAc_2$, can be produced. The branching of mammalian-like oligosaccharide structures will then occur after trimming of the oligosaccharide to the $Man_5GlcNAc_2$.

According to the present invention, the nucleotide sequence encoding an α-1,2-mannosidase for use in the expression vector of the present invention can derive from any species. A number of α-1,2-mannosidase genes have been cloned and are available to those skilled in the art, including mammalian genes encoding, e.g., a murine α-1,2-mannosidase (Herscovics et al. *J. Biol. Chem.* 269: 9864-9871, 1994), a rabbit α-1,2-mannosidase (Lal et al. *J. Biol. Chem.* 269: 9872-9881, 1994) or a human α-1,2-mannosidase (Tremblay et al. *Glycobiology* 8: 585-595, 1998), as well as fungal genes encoding, e.g., an *Aspergillus* α-1,2-mannosidase (msdS gene), a *Trichoderma reesei* α-1,2-mannosidase (Maras et al. *J. Biotechnol.* 77: 255-263, 2000), or a *Saccharomyces cerevisiae* α-1,2-mannosidase. Protein sequence analysis has revealed

a high degree of conservation among the eukaryotic α-1,2-mannosidases identified so far.

Preferably, the nucleotide sequence for use in the present vectors encodes a fungal α-1,2-mannosidase, more preferably, a *Trichoderma reesei* α-1,2-mannosidase, and more particularly, the *Trichoderma reesei* α-1,2-mannosidase described by Maras et al. *J. Biotechnol.* 77: 255-63 (2000).

According to the present invention, the nucleotide sequence can also code for only a functional part of an α-1,2-mannosidase.

By "functional part" is meant a polypeptide fragment of an α-1,2-mannosidase which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant that at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length α-1,2-mannosidase is retained. For example, as illustrated by the present invention, the catalytic domain of the murine α-1,2-mannosidase IB constitutes a "functional part" of the murine α-1,2-mannosidase IB. Those skilled in the art can readily identify and make functional parts of an α-1,2-mannosidase using a combination of techniques known in the art. Predictions of the portions of an α-1,2-mannosidase essential to or sufficient to confer the enzymatic activity can be made based on analysis of the protein sequence. The activity of a portion of an α-1,2-mannosidase of interest, expressed and purified from an appropriate expression system, can be verified using *in vitro* or *in vivo* assays described hereinbelow.

In accordance with the present invention, an α-1,2-mannosidase or a functional part thereof expressed in a methylotrophic yeast strain preferably is targeted to a site in the secretory pathway where $Man_8GlcNAc_2$ (the substrate of α-1,2-mannosidase) is already formed on a glycoprotein, but has not reached a Golgi glycosyltransferase which elongates the sugar chain with additional mannose residues.

Accordingly, in a preferred embodiment of the present invention, the α-1,2-mannosidase expression vector is engineered as such that the α-1,2-mannosidase or a functional part thereof expressed from the vector includes an ER-retention signal.

An "ER retention signal" refers to a peptide sequence which directs a protein having such peptide sequence to be transported to and retained in the ER. Such ER retention sequences are often found in proteins that reside and function in the ER.

Multiple choices of ER retention signals are available to those skilled in the art, e.g., the first 21 amino acid residues of the *S. cerevisiae* ER protein MNS1 (Martinet et al. *Biotechnology Letters* 20: 1171-1177, 1998). A preferred ER retention signal for use in the present invention is peptide HDEL (SEQ ID NO: 1). The HDEL peptide sequence, found in the C-terminus of a number of yeast proteins, acts as a retention/retrieval signal for the ER (Pelham *EMBO J.* 7: 913-918, 1988). Proteins with an HDEL sequence are bound by a membrane-bound receptor (Erd2p) and then enter a retrograde transport pathway for return to the ER from the Golgi apparatus.

According to the present invention, an ER retention signal can be placed anywhere in the protein sequence of an α-1,2-mannosidase, but preferably at the C-terminus of the α-1,2-mannosidase.

The α-1,2-mannosidase for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG and His6 tags, which are well-known in the art. An epitope-tagged α-1,2-mannosidase can be conveniently purified, or monitored for both expression and intracellular localization.

An ER retention signal and an epitope tag can be readily introduced into a protein of interest by inserting nucleotide sequences coding for such signal or tag into the nucleotide sequence encoding the protein of interest, using any of the molecular biology techniques known in the art.

In another preferred embodiment, the vectors of the present invention include a sequence coding for a glucosidase II or a functional part thereof and are capable of expressing the glucosidase II or the functional part in the methylotrophic yeast strain.

It has been established that the initial N-linked oligosaccharide ($Glc_3Man_9GlcNAc_2$), transferred in the ER onto a protein, is cleaved in the ER by specific glucosidases to remove 3 glucose residues, and by a mannosidase to remove 1 specific α-1,2-linked mannose. It has been observed by the present inventors that some recombinant proteins expressed in *Pichia* have residual glucose residues on the sugar moiety when such proteins leave the ER for the Golgi apparatus. The residual glucose molecules present on the sugar structure prevent the complete digestion of the sugar moiety by an α-1,2-mannosidase, and the introduction of an exogenous glucosidase can facilitate the removal of these glucose residues.

According to the present invention, the nucleotide sequence encoding a glucosidase II can derive from any species. The Glucosidase II genes have been cloned from a number of mammalian species including rat, mouse, pig and human. The glucosidase II protein from these mammalian species consists of an alpha and a beta subunit. The alpha subunit is about 110 kDa and contains the catalytic activity of the enzyme, while the beta subunit has a C-terminal HDEL ER-retention sequence and is believed to be important for the ER localization of the enzyme. The glucosidase II gene from *S. cerevisiae* has also been cloned (ORF YBR229c, located on chromosome II). This gene encodes a protein of about 110 kDa, which shows a high degree of homology to the mammalian alpha subunits.

A preferred glucosidase II gene for use in the present vectors is from a fungal species such as *Pichia pastoris* and *S. cerevisiae*. An example of a fungal glucosidase II gene is the *S. cerevisiae* glucosidase II gene.

According to the present invention, the nucleotide sequence can also encode only a functional part of a glucosidase II. By "functional part" is meant a

polypeptide fragment of a glucosidase II which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant that at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length glucosidase II is retained. Functional parts of a glucosidase II can be identified and made by those skilled in the art using a variety of techniques known in the art.

In a preferred embodiment of the present invention, the glucosidase II protein is engineered to include an ER retention signal such that the protein expressed in a methylotrophic yeast strain is targeted to the ER and retains therein for function. ER retention signals are as described hereinabove, e.g., the HDEL peptide sequence.

The glucosidase II for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG, and His6 tag, which are well-known in the art.

According to the present invention, the "knock-in" vectors can include either or both of an α-1,2-mannosidase coding sequence and a glucosidase II coding sequence.

In another preferred embodiment, the vectors of the present invention include a sequence coding for a GlcNAc-Transferase I or a functional part thereof and are capable of expressing the GlcNAc-Transferase I or the functional part thereof in a methylotrophic yeast strain.

A GlcNAc-Transferase I is responsible for the addition of β-1,2-GlcNAc to a $Man_5GlcNAc_2$, and converts this core oligosaccharide on glycoproteins to $GlcNAcMan_5GlcNAc_2$. The mannose residues of $GlcNAcMan_5GlcNAc_2$ can be further trimmed by a mammalian Golgi mannosidase II. New sugars are added to form hybrid- or complex-type sugar branches characteristic of mammalian glycoproteins. Thus, by way of introducing a GlcNAc-transferase I into methylotrophic yeasts such as *Pichia*, glycoproteins with a mammalian-like glycosylation pattern, such as glycoproteins with glycans $GlcNAcMan_5GlcNAc_2$, can be produced.

According to the present invention, the nucleotide sequence encoding a GlcNAc-transferase I (GnTI) for use in the expression vector of the present invention can derive from any species, e.g., rabbit, rat, human, plants, insects, nematodes and protozoa such as *Leishmania tarentolae*, or can be obtained through protein engineering experiments. Preferably, the nucleotide sequence for use in the present vectors encodes a human GnTI, and more preferably, the GnTI having a sequence as set forth in SEQ ID NO: 34.

According to the present invention, the nucleotide sequence can also encode only a functional part of a GlcNAc-Transferase I.

By "functional part" is meant a polypeptide fragment of a GlcNAc-Transferase I which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant that at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length GlcNAc-Transferase I is retained. For example, as illustrated by the present invention, the catalytic domain of the human GnTI constitutes a "functional part" of the human GnTI. Those skilled in the art can readily identify and make functional parts of a GlcNAc-Transferase I using a combination of techniques known in the art. Predictions of the portions of a GlcNAc-Transferase I essential to, or sufficient to confer, the enzymatic activity can be made based on analysis of the protein sequence. The activity of a portion of a GlcNAc-Transferase I of interest, expressed and purified from an appropriate expression system, can be verified using *in vitro* or *in vivo* assays described hereinbelow.

In accordance with the present invention, a GnTI or a functional part thereof expressed in a methylotrophic yeast strain preferably is targeted to a site in the secretory pathway where $Man_5GlcNAc_2$ (the substrate of GnTI) is already formed on a glycoprotein. Preferably, the GnTI or a functional part thereof is targeted to the Golgi apparatus; and more preferably, to the Golgi apparatus of a yeast strain which is also transformed with α-1,2-mannosidase.

Accordingly, in a preferred embodiment of the present invention, the GnTI is engineered such that the GnTI or a functional part thereof expressed from the vector includes a yeast Golgi localization signal.

A yeast "Golgi localization signal" refers to a peptide sequence which directs a protein having such peptide sequence to be retained in the Golgi apparatus. Such Golgi localization sequences are often found in proteins that reside and function in the Golgi apparatus.

Choices of Golgi localization signals are available to those skilled in the art. A preferred Golgi localization signal for use in the present invention is a peptide derived from the N-terminal part of a *Saccharomyces cerevisiae* Kre2 protein (ScKre2). More preferably, the Kre2 gene has a sequence as set forth in SEQ ID NO: 30. A particularly preferred Golgi localization signal is a peptide having amino acid 1-100 (SEQ ID NO: 35) of the ScKre2 protein having SWISS-PROT Accession No. P27809 (SEQ ID NO: 38).

According to the present invention, a Golgi localization signal can be placed anywhere within the GnTI, but preferably at the terminus of the GnTI, and more preferably at the N-terminus of the GnTI.

The GnTI for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG and His6 tags well-known in the art. An epitope-tagged GnTI can be conveniently purified, or monitored for both expression and intracellular localization.

A Golgi localization signal and an epitope tag can be readily introduced into a protein of interest by inserting nucleotide sequences coding for such signal or tag into the nucleotide sequence encoding the protein of interest, using any of the molecular biology techniques known in the art.

In another preferred embodiment, the vectors of the present invention include a sequence coding for a mannosidase II or a functional part thereof and are

capable of expressing the mannosidase II or the functional part in a methylotrophic yeast strain.

It has been established that a GlcNAc-Transferase I adds a β-1,2-GlcNAc to a $Man_5GlcNAc_2$, and converts the core oligosaccharide $Man_5GlcNAc_2$ on glycoproteins to $GlcNAcMan_5GlcNAc_2$. $GlcNAcMan_5GlcNAc_2$ is the substrate of, and can be further processed by, a mammalian Golgi mannosidase II, to remove more mannose residues. The resulting oligosaccharide group can be further modified to produce hybrid or complex type sugar branches characteristic of mammalian glycoproteins. Thus, by way of introducing a mannosidase II into methylotrophic yeasts such as *Pichia*, and especially *Pichia pastoris*, glycoproteins with a reduced mannose content can be produced. The glycosylation pattern on the proteins so produced is also characteristic of mammalian glycoproteins.

According to the present invention, the nucleotide sequence encoding a Golgi mannosidase II can derive from any species. The Mannosidase II genes have been cloned from a number of mammalian species.

A preferred mannosidase II gene for use in the present vectors is from a mammalian species. An example of a mammalian mannosidase II gene is the human mannosidase II cDNA, as published in EMBL with Accession NO 31520 (SEQ ID NO: 34).

According to the present invention, the nucleotide sequence can also encode only a functional part of a mannosidase II. By "functional part" is meant a polypeptide fragment of a mannosidase II which substantially retains the enzymatic activity of the full-length protein. By "substantially" is meant that at least about 40%, or preferably, at least 50% or more of the enzymatic activity of the full-length mannosidase II is retained. Functional parts of a mannosidase II can be identified and made by those skilled in the art using a variety of techniques known in the art.

In a preferred embodiment of the present invention, the mannosidase II protein is engineered to include a Golgi localization signal such that the protein

expressed in a methylotrophic yeast strain is targeted to the Golgi compartment and retains therein for function. The Golgi localization signals which can be used in connection with a mannosidase II are as described hereinabove in connection with GnTI, e.g., a peptide derived from the N-terminal part of the ScKre2.

The mannosidase II for use in the present invention can be further modified, e.g., by insertion of an epitope tag to which antibodies are available, such as Myc, HA, FLAG, and His6 tag, which are well-known in the art.

According to the present invention, the "knock-in" vectors can include either or both of a GnTI coding sequence and a mannosidase II coding sequence.

Vectors including two or more of α-1,2-mannosidase coding sequence, glucosidase II coding sequence, GnTI coding sequence and mannosidase II coding sequence are also provided by present invention.

Further according to the present invention, the nucleotide sequence encoding the enzyme to be expressed (e.g., an α-1,2-mannosidase or a functional part thereof, or a glucosidase II or a functional part thereof, a GnTI or a functional part thereof, or a mannosidase II or a functional part thereof) can be placed in an operable linkage to a promoter and a 3' termination sequence.

Promoters appropriate for expression of a protein in a methylotrophic yeast can include both constitutive promoters and inducible promoters. Constitutive promoters include, e.g., the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase promoter ("the GAP promoter") and the *Pichia pastoris* YPT1 promoter (Sears et al, Yeast 14: 783-790, 1998). Examples of inducible promoters include, e.g., the *Pichia pastoris* alcohol oxidase I promoter ("the AOXI promoter") (U.S. Patent No. 4,855,231), the *Pichia pastoris* alcohol oxidase II ("the AOX2 promoter") (Ohi et al., Mol. Gen. Genet 243: 489-499, 1994), or the *Pichia pastoris* formaldehyde dehydrogenase promoter ("the FLD promoter") (Shen et al. Gene 216: 93-102, 1998).

3' termination sequences are sequences 3' to the stop codon of a structural gene which function to stabilize the mRNA transcription product of the gene to which the sequence is operably linked, such as sequences which elicit polyadenylation. 3' termination sequences can be obtained from *Pichia* or other methylotrophic yeast. Examples of *Pichia pastoris* 3' termination sequences useful for the practice of the present invention include termination sequences from the *AOX1* gene, *p40* gene, *HIS4* gene and *FLD1* gene.

The vectors of the present invention preferably contain a selectable marker gene. The selectable marker may be any gene which confers a selectable phenotype upon a methylotrophic yeast strain and allows transformed cells to be identified and selected from untransformed cells. The selectable marker system may include an auxotrophic mutant methylotrophic yeast strain and a wild type gene which complements the host's defect. Examples of such systems include the *Saccharomyces cerevisiae* or *Pichia pastoris HIS4* gene which may be used to complement *his4* *Pichia* strains, or the *S. cerevisiae* or *Pichia pastoris ARG4* gene which may be used to complement *Pichia pastoris arg* mutants, or the *Pichia pastoris* URA3 and ADE1 genes, which may be used to complement *Pichia pastoris* ura3 resp. ade1 mutants. Other selectable marker genes which function in *Pichia pastoris* include the Zeo^R gene, the $G418^R$ gene, blastisidin resistance gene, and the like.

The vectors of the present invention can also include an autonomous replication sequence (ARS). For example, U.S. Patent No. 4,837,148 describes autonomous replication sequences which provide a suitable means for maintaining plasmids in *Pichia pastoris*. The disclosure of U.S. Patent No. 4,837,148 is incorporated herein by reference.

The vectors can also contain selectable marker genes which function in bacteria, as well as sequences responsible for replication and extrachromosomal maintenance in bacteria. Examples of bacterial selectable marker genes include

ampicillin resistance (*Amp*^{r}), tetracycline resistance (*Tet*^{r}), neomycin resistance, hygromycin resistance and zeocin resistance (*Zeo*^{R}) genes.

According to the present invention, the nucleotide sequence encoding the protein to be expressed in a methylotrophic yeast can be placed in an integrative vector or a replicative vector (such as a replicating circular plasmid).

Integrative vectors are disclosed, e.g., in U.S. Patent No. 4,882,279 which is incorporated herein by reference. Integrative vectors generally include serially arranged sequences of at least a first insertable DNA fragment, a selectable marker gene, and a second insertable DNA fragment. The first and second insertable DNA fragments are each about 200 nucleotides in length and have nucleotide sequences which are homologous to portions of the genomic DNA of the species to be transformed. A nucleotide sequence containing a structural gene of interest for expression is inserted in this vector between the first and second insertable DNA fragments whether before or after the marker gene. Integrative vectors can be linearized prior to yeast transformation to facilitate the integration of the nucleotide sequence of interest into the host cell genome.

Replicative and integrative vectors carrying one or more of an α-1,2-mannosidase coding sequence, a glucosidase II coding sequence, a GnTI coding sequence, and a mannosidase II coding sequence can be constructed by standard techniques known to one of ordinary skill in the art and found, for example, in Sambrook et al. (1989) Molecular Cloning: A *Laboratory Manual*, or any of a myriad of laboratory manuals on recombinant DNA technology that are widely available.

Preferred vectors of the present invention carrying an α-1,2-mannosidase expression sequence include pGAPZMFManHDEL, pGAPZMFManMycHDEL, pPICZBMFManMycHDEL, pGAPZmManHDEL, pGAPZmMycManHDEL, pPIC9mMycManHDEL and pGAPZmMycManHDEL, which are further described in the Examples hereinbelow.

Preferred vectors of the present invention carrying a glucosidase II expression sequence include pGAPZAGLSII, pPICZAGLSII, pAOX2ZAGLSII, pYPTIZAGLSII, pGAPADE1glsII, pPICADE1glsII, pAOX2ADE1glsII, pYPTIADE1glsII, pGAPZAglsIIHDEL and pGAPADE1glsIIHDEL, which are further described in the Examples hereinbelow.

Preferred vectors of the present invention carrying a GlcNAc-transferase I expression sequence include pPIC6AKrecoGnTI, which are further described in the Examples hereinbelow.

Preferred vectors of the present invention carrying a mannosidase II expression sequence include pGAPKreManII, which are further described in the Examples hereinbelow.

In another aspect, the present invention provides inactivation vectors which, when introduced into a methylotrophic yeast strain, inactivate or disrupt a gene thereby facilitating the reduction in the glycosylation of glycoproteins produced in the methylotrophic yeast strain.

In one embodiment, the present invention provides a inactivation vector which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the OCH1 gene.

The *S. cerevisiae* QCH1 gene has been cloned (Nakayama et al. *EMBO J.* 11: 2511-2519, 1992). It encodes a membrane bound α-1,6-mannosyltransferase, localized in the early Golgi complex, that is functional in the initiation of α-1,6-polymannose outer chain addition to the N-linked core oligosaccharide ($Man_5GlcNAc_2$ and $Man_8GlcNAc_2$) (Nakanishi-Shindo et al. *J. Biol. Chem.* 268: 26338-26345, 1993).

A *Pichia* sequence has been described in Japanese Patent Application No. 07145005 that encodes a protein highly homologous to the *S. cerevisiae* OCH1. For purpose of the present invention, this sequence is denoted herein as "the *Pichia* OCH1 gene". Those skilled in the art can isolate the OCH1 genes from other methylotrophic yeasts using techniques well known in the art.

According to the present invention, a disruption in the OCH1 gene of a methylotrophic yeast can result in either the production of an inactive protein product or no product. The disruption may take the form of an insertion of a heterologous DNA sequence into the coding sequence and/or the deletion of some or all of the coding sequence. Gene disruptions can be generated by homologous recombination essentially as described by Rothstein (in *Methods in Enzymology*, Wu et al., eds., vol 101:202-211, 1983).

To disrupt the OCH1 gene by homologous recombination, an OCH1 inactivation vector can be constructed in such a way to include a selectable marker gene. The selectable marker gene is operably linked, at both 5' and 3' end, to portions of the OCH1 gene of sufficient length to mediate homologous recombination. The selectable marker can be one of any number of genes which either complement host cell auxotrophy or provide antibiotic resistance, including *URA3*, *ARG4*, *HIS4*, *ADE1*, *LEU2* and *HIS3* genes. Other suitable selectable markers include the *CAT* gene, which confers chloramphenicol resistance on yeast cells, or the *lacZ* gene, which results in blue colonies due to the expression of active β-galactosidase. Linearized DNA fragments of an OCH1 inactivation vector are then introduced into host methylotrophic yeast cells using methods well known in the art. Integration of the linear fragments into the genome and the disruption of the OCH1 gene can be determined based on the selection marker and can be verified by, for example, Southern Blot analysis.

Alternatively, an OCH1 inactivation vector can be constructed in such a way to include a portion of the OCH1 gene to be disrupted, which portion is devoid of any OCH1 promoter sequence and encodes none or an inactive fragment of the OCH1 protein. By "an inactive fragment", it is meant a fragment of the OCH1 protein which has, preferably, less than about 10% and most preferably, about 0% of the activity of the full-length OCH1 protein. Such portion of the OCH1 gene is inserted in a vector in such a way that no known promoter sequence is operably linked to the OCH1

sequence, but that preferably a stop codon and a transcription termination sequence are operably linked to the portion of the OCH1 gene. This vector can be subsequently linearized in the portion of the OCH1 sequence and transformed into a methylotrophic yeast strain using any of the methods known in the art. By way of single homologous recombination, this linearized vector is then integrated in the OCH1 gene. Two OCH1 sequences are produced in the chromosome as a result of the single homologous recombination. The first OCH1 sequence is the portion of the OCH1 gene from the vector, which is now under control of the OCH1 promoter of the host methylotrophic yeast, yet cannot produce an active OCH1 protein as such OCH1 sequence codes for no or an inactive fragment of the OCH1 protein, as described hereinabove. The second OCH1 sequence is a full or mutated OCH1 coding sequence, which is not operably linked to any known promoter sequence and thus, no active messenger is expected to be formed for synthesis of an active OCH1 protein. Preferably, an inactivating mutation is introduced in the OCH1 sequence, to the 5' end of the site of linearization of the vector and to the 3' end of the translation initiation codon of OCH1. By "inactivating mutation" it is meant a mutation introducing a stop codon, a frameshift mutation or any other mutation causing a disruption of the reading frame, or a 5' truncation of the OCH1 reading frame. A preferred inactivating mutation is a point mutation. Preferably, by introduction of a stop codon at the amino acid 12 at the very 5' end of OCH1 as described in Example 8. Such mutation can be introduced into an OCH1 sequence using any of the site directed mutagenesis methods known in the art. Another inactivating mutation is 5' truncation of the open reading frame. Inactivating mutations as described above ensure that no functional OCH1 protein can be formed even if there exist some promoter sequences 5' to the OCH1 sequence in the inactivation vector.

Preferred OCH1 inactivation vectors of the present invention include pBLURA5'PpOCH1, pZ5'PpOCH1Trunc and pZMFManHDEL5'PpOCH1Mut, which are further described in the Examples hereinbelow.

If desired, one or more of an α-1,2-mannosidase expression sequence, a glucosidase expression sequence, a GnTI expression sequence, or a mannosidase II expression sequence can be carried on the same plasmid used to disrupt the OCH1 gene to create a "knock-in-and-inactivation" vector.

In a preferred embodiment, the vectors of the present invention include a sequence coding for an α-1,2-mannosidase or a functional part thereof, and a sequence which is capable of disrupting or inactivating the endogenous OCH1 gene or a functional part thereof. Such vector(s) is/are capable of expressing the α-1,2-mannosidase or the functional part, and is/are capable of disrupting the OCH1 gene in methylotrophic yeast, including *Pichia*, especially *Pichia pastoris*.

Preferred knock-in-and-inactivation vectors of the present invention include pZMFManHDEL5'PpOCH1Mut, which are further described in the Examples hereinbelow.

Additionally, any of the above-described vectors can further include a nucleotide sequence capable of expressing a glycoprotein of interest in a methylotrophic yeast strain, including *Pichia*, especially *Pichia pastoris*.

Another aspect of the present invention is directed to methods of modifying methylotrophic yeast strains to reduce or modify glycosylations, to a mammalian-like pattern, on proteins produced by the methylotrophic yeast strains. In accordance with the present methods, methylotrophic yeast strains are modified by transforming these yeast strains with one or more, i.e., at least one, knock-in and/or inactivation vectors of the present invention as described herein above.

Methylotrophic yeast strains which can be modified using the present methods include but are not limited to yeasts capable of growing on methanol, such as yeasts of the genera *Candida, Hansenula, Torulopsis,* and *Pichia.* A list of species which are exemplary of this class of yeasts can be found in C. Anthony (1982), *The Biochemistry of Methylotrophs, 269. Pichia pastoris, Pichia methanolica, Pichia anomola, Hansenula polymorpha* and *Candida boidinii* are examples of

methylotrophic yeasts useful in the practice of the present invention. Preferred methylotrophic yeasts are of the genus *Pichia*. Especially preferred are *Pichia pastoris* strains GS115 (NRRL Y-15851); GS190 (NRRL Y-18014) disclosed in U.S. Patent No. 4,818,700; PPF1 (NRRL Y-18017) disclosed in U.S. Patent No. 4,812,405; PPY120H and YGC4; as well as strains derived therefrom.

Methylotrophic yeast strains which can be modified using the present methods also include those methylotrophic yeast strains which have been genetically engineered to express one or more heterologous glycoproteins of interest. The glycosylation on the heterologous glycoproteins expressed from these previously engineered strains can be reduced or modified to mammalian-like pattern by transforming such strains with one or more of the vectors of the present invention.

The vectors of the present invention can be introduced into the cells of a methylotrophic yeast strain using known methods such as the spheroplast technique, described by Cregg et al. 1985, or the whole-cell lithium chloride yeast transformation system, Ito et al. *Agric. Biol. Chem. 48*:341, modified for use in *Pichia* as described in EP 312,934. Other published methods useful for transformation of the plasmids or linear vectors include U.S. Patent No. 4,929,555; Hinnen et al. *Proc. Nat. Acad. Sci. USA* 75:1929 (1978); Ito et al. *J. Bacteriol.* 153:163 (1983); U.S. Patent No. 4,879,231; Sreekrishna et al. *Gene* 59:115 (1987). Electroporation and PEG1000 whole cell transformation procedures may also be used, as described by Cregg and Russel, *Methods in Molecular Biology: Pichia Protocols,* Chapter 3, Humana Press, Totowa, N.J., pp. 27-39 (1998).

Transformed yeast cells can be selected by using appropriate techniques including, but not limite, to culturing auxotrophic cells after transformation in the absence of the biochemical product required (due to the cell's auxotrophy), selection for and detection of a new phenotype, or culturing in the presence of an antibiotic which is toxic to the yeast in the absence of a resistance gene contained in the transformants. Transformants can also be selected and/or verified by integration of the

expression cassette into the genome, which can be assessed by, e.g., Southern Blot or PCR analysis.

In one embodiment, a methylotrophic yeast strain is transformed with a vector which includes a nucleotide sequence coding for an α-1,2-mannosidase or a functional part thereof. The nucleotide sequence is capable of expressing the α-1,2-mannosidase or the functional part in the methylotrophic yeast strain, and is, preferably, integrated into the genome of the methylotrophic yeast strain.

The expression of an α-1,2-mannosidase introduced in a methylotrophic yeast strain can be verified both at the mRNA level, e.g., by Northern Blot analysis, and at the protein level, e.g., by Western Blot analysis. The intracellular localization of the protein can be analyzed by using a variety of techniques, including subcellular fractionation and immunofluorescence experiments. An ER localization of an α-1,2-mannosidase can be determined by co-sedimentation of this enzyme with a known ER resident protein (e.g., Protein Disulfide Isomerase) in a subcellular fractionation experiment. An ER localization can also be determined by an immunofluorescence staining pattern characteristic of ER resident proteins, typically a perinuclear staining pattern.

To confirm that an α-1,2-mannosidase or a functional part thereof expressed in a methylotrophic yeast strain has the expected mannose-trimming activity, both *in vitro* and *in vivo* assays can be employed. Typically, an *in vitro* assay involves digestion of an *in vitro* synthesized substrate, e.g., $Man_8GlcNAc_2$, with the enzyme expressed and purified from a methylotrophic yeast strain, and assessing the ability of such enzyme to trim $Man_8GlcNAc_2$ to, e.g., $Man_5GlcNAc_2$. Alternatively, the mannosidase activity in *Pichia* cell lysates can be assayed using a chromogenic substrate for the enzyme, such as DNP-α-D-mannopyranoside (T. Desmedt, N. Callewaert, R. Contreras and M. Claeyssens, Anal. Biochem., in press). In *in vivo* assays, the α-1,2-mannosidase or a functional part thereof is co-expressed in a

methylotrophic yeast with a glycoprotein known to be glycosylated with N-glycans bearing terminal α-1,2-linked mannose residues in such yeast. The enzymatic activity of such an α-1,2-mannosidase or a functional part thereof can be measured based on the reduction of the number of α-1,2-linked mannose residues in the structures of the N-glycans of the glycoprotein. In both *in vitro* and *in vivo* assays, the composition of a carbohydrate group can be determined using techniques that are well known in the art and are illustrated in the Examples hereinbelow.

In another embodiment, a methylotrophic yeast strain is transformed with a vector which includes a nucleotide sequence coding for a glucosidase II or a functional part thereof. The nucleotide sequence is capable of expressing the glucosidase II or the functional part in the methylotrophic yeast strain, and is, preferably, integrated into the genome of the methylotrophic yeast strain.

The expression of Glucosidase II introduced in a methylotrophic yeast strain can be verified both at the mRNA level, e.g., by Northern Blot analysis, and at the protein level, e.g., by Western Blot analysis. The intracellular localization of the protein can be analyzed by using a variety of techniques, including subcellular fractionation and immunofluorescence experiments.

The enzymatic activity of a glucosidase II or a functional part thereof expressed in a transformed methylotrophic yeast strain can be assessed using a variety of assays. For example, methylotrophic yeast cells transformed with a sequence encoding a glucosidase II or a functional part thereof can be set to grow on solid medium containing a substrate of the glucosidase, e.g., 5-bromo-4-chloro-3-indolyl-α-D-glucopyranoside or 4-MU-α–D-Glc. When the enzyme is expressed by the *Pichia* and secreted extracellularly, the substrate is acted upon by the enzyme, giving rise to detectable signals around the colonies such as blue color or fluorescent glow. Alternatively, liquid culture medium containing the expressed protein molecules can be collected and incubated in test tubes with a substrate, e.g., p-nitrophenyl-α-D-

glucopyranoside or 4-MU-α–D-Glc. The enzymatic activity can be determined by measuring the specific product released. Moreover, *in vivo* assays can be employed, where a glucosidase II is co-expressed in yeast with a glycoprotein known to be N-glycosylated with glucose residues, e.g., influenza neuraminidase. The enzymatic activity of the glucosidase II can be measured based on the reduction of the glucose content in the sugar chain(s) of the glycoprotein.

In one embodiment, a methylotrophic yeast strain, such as *Pichia* and particularly *Pichia pastoris*, is transformed with a vector which includes a nucleotide sequence coding for GlcNAc-transferase I or a functional part thereof. The nucleotide sequence is capable of expressing the GlcNAc-transferase I or the functional part in the methylotrophic yeast strain, and is, preferably, integrated into the genome of the methylotrophic yeast strain.

The expression of GlcNAc-transferase I introduced in a methylotrophic yeast strain can be verified both at the mRNA level, e.g., by Northern Blot analysis, and at the protein level, e.g., by Western Blot analysis. The intracellular localization of the protein can be analyzed by using a variety of techniques, including subcellular fractionation and immunofluorescence experiments.

To confirm that GlcNAc-transferase I or a functional part thereof expressed in a methylotrophic yeast strain has the expected $GlcNAcMan_5GlcNAc_2$, both *in vitro* and *in vivo* assays can be employed. In both *in vitro* and *in vivo* assays, the composition of a carbohydrate group can be determined using techniques that are well known in the art and are illustrated in Example 6 hereinbelow.

In another embodiment, a methylotrophic yeast strain, such as *Pichia* and particularly *Pichia pastoris*, is transformed with a vector which includes a nucleotide sequence coding for a mannosidase II or a functional part thereof. The nucleotide sequence is capable of expressing the mannosidase II or the functional part in the methylotrophic yeast strain, and is, preferably, integrated into the genome of the methylotrophic yeast strain.

The expression and enzymatic activity of a mannosidase II or a functional part thereof expressed in a transformed methylotrophic yeast strain can be assessed using a variety of techniques and assays, such as gemomic analysis and glycan assay.

In still another embodiment of the present invention, a methylotrophic yeast strain is transformed with an OCH1 inactivation vector. As a result of the transformation and integration of the vector, the endogenous genomic OCH1 gene in the yeast strains is disrupted.

In still another embodiment of the present invention, a methylotrophic yeast strain is transformed with a knock-in-and-inactivation vector, such as an α-1,2-mannosidase-disrupted OCH1 vector. As a result of the transformation and integration of the vector, the genomic OCH1 gene in the yeast strains is disrupted and a potent α-1,2-mannosidase is expressed in the yeast.

In a further embodiment of the present invention, a methylotrophic yeast strain, such as *Pichia* and particularly *Pichia pastoris*, is transformed with one or more of an α-1,2-mannosidase expression vector, a glucosidase II expression vector, a GnTI expression vector, a mannosidase II expression vector, an OCH1 inactivation vector or an α-1,2-mannosidase-disrupted OCH1 knock-in-and-inactivation vector. Such modification can be achieved by serial, consecutive transformations, i.e., introducing one vector at a time, or alternatively by co-transformation, i.e., introducing the vectors simultaneously.

Preferably, a strain transformed with a vector encoding GnTI is also transformed with a vector encoding α-1,2-mannosidase, either simultaneously or sequentially. Also preferably, a strain transformed with a vector encoding mannosidase II is also transformed with a vector encoding GnTI and α-1,2-mannosidase.

The modified methylotrophic yeast strains described hereinabove can be further modified if desired. For example, additional disruption of genes encoding any

other yeast mannosyltransferases can be made. Genes encoding mammalian enzymes can also be introduced to produce glycoproteins having hybrid- or complex-type N-glycans, if desired.

Methylotrophic yeast strains which are modified by using the present methods, i.e., by transforming with one or more of the vectors of the present invention, form another embodiment of the present invention.

It should be understood that certain aspects of the present invention, especially the introduction of an intracellularly expressed α-1,2-mannosidase activity, are also useful to obtain a reduced or modified glycosylation of the O-linked glycans on glycoproteins produced in a methylotrophic yeast, as it is known in the art that these O-linked glycans consist mainly of α-1,2-linked mannose residues. O-linked glycans as used herein refer to carbohydrate structures linked to serine or threonine residues of glycoproteins.

A further aspect of the invention is directed to methods of producing a glycoprotein with reduced or modified glycosylation, preferably a mammalian-like pattern of glycosylation with hybrid or complex glycans, in a methylotrophic yeast, especially a glycoprotein heterologous to the methylotrophic yeast.

"A glycoprotein" as used herein refers to a protein which, in methylotrophic yeasts, is either glycosylated on one or more asparagines residues having the consensus sequence "Asn-Xaa-Thr" or "Asn-Xaa-Ser" (where Xaa is any amino acid except proline), or on one or more serine (or threonine) residues, or on both asparagines and serine (or threonine) residues.

The term "reduced or modified glycosylation" refers to a reduced size of the carbohydrate moiety on the glycoprotein, particularly with fewer mannose residues or modified to desired, such as mammalian-like, complex glycans on the glycoprotein, when the glycoprotein is expressed in a methylotrophic yeast strain which has been modified in accordance with the present invention, as compared to a wild type, unmodified strain of the methylotrophic yeast.

In accordance with the present invention, the production of a glycoprotein of interest with reduced or modified glycosylation can be achieved in a number of ways. A nucleotide sequence capable of expressing a glycoprotein can be introduced into a methylotrophic yeast strain which has been previously modified in accordance with the present invention, i.e., a strain transformed with one or more of the vectors of the present invention and capable of producing glycoproteins with reduced and/or modified glycosylation. Alternatively, a methylotrophic yeast strain which has already been genetically engineered to express a glycoprotein can be transformed with one or more of the vectors of the present invention. Otherwise, if a methylotrophic yeast strain does not express a glycoprotein of interest, nor is the strain transformed with any of the vectors of the present invention, such yeast strain can be transformed, either consecutively or simultaneously, with both a nucleotide sequence capable of expressing the glycoprotein and one or more vectors of the present invention. Additionally, a methylotrophic yeast strain can be transformed with one or more of the present knock-in and/or inactivation and/or knock-in-and-inactivation vectors which also include a nucleotide sequence capable of expressing a glycoprotein in the methylotrophic yeast strain.

The nucleotide sequence capable of expressing a glycoprotein in a methylotrophic yeast can be made to include from 5' to 3', a promoter, a sequence encoding the glycoprotein, and a 3' termination sequence. Promoters and 3' termination sequences which are suitable for expression of a glycoprotein can include any of those promoters and 3' termination sequences described hereinabove.

The nucleotide sequence for expression of a glycoprotein can include additional sequences, e.g., signal sequences coding for transit peptides when secretion of a protein product is desired. Such sequences are widely known, readily available and include *Saccharomyces cerevisiae* alpha mating factor prepro (αmf), *Pichia pastoris* acid phosphatase (PHO1) signal sequence and the like.

The nucleotide sequence for expression of a glycoprotein can be placed on a replicative vector or an integrative vector. The choice and construction of such vectors are as described hereinabove.

The nucleotide sequence capable of expressing a glycoprotein can be carried on the same replicative plasmid as a plasmid carrying any number (at least one) of, an α-1,2-mannosidase, a glucosidase II, a GnTI, a mannosidase II, an inactived OCH1 and an α-1,2-mannosidase-inactivated OCH1 expression unit. Alternatively, the nucleotide sequence containing the glycoprotein coding sequence is carried on a separate plasmid or integrated into the host genome.

Glycoproteins produced can be purified by conventional methods. Purification protocols can be determined by the nature of the specific protein to be purified. Such determination is within the ordinary level of skill in the art. For example, the cell culture medium is separated from the cells and the protein secreted from the cells can be isolated from the medium by routine isolation techniques such as precipitation, immunoabsorption, fractionation or a variety of chromatographic methods.

Glycoproteins which can be produced by the methods of the present invention include, e.g., *Bacillus amyloliquefaciens* α-amylase, *S. cerevisiae* invertase, *Trypanosoma cruzi trans*-sialidase, HIV envelope protein, influenza virus A haemagglutinin, influenza neuraminidase, Bovine herpes virus type-1 glycoprotein D, human angiostatin, human B7-1, B7-2 and B-7 receptor CTLA-4, human tissue factor, growth factors (e.g., platelet-derived growth factor), tissue plasminogen activator, plasminogen activator inhibitor-I, urokinase, human lysosomal proteins such as α-galactosidase, plasminogen, thrombin, factor XIII and immunoglobulins. For additional useful glycoproteins which can be expressed in the genetically engineered *Pichia* strains of the present invention, see Bretthauer and Castellino, *Biotechnol. Appl. Biochem.* 30: 193-200 (1999), and Kukuruzinska et al. *Ann Rev. Biochem.* 56: 915-44 (1987).

Glycoproteins produced by using the methods of the present invention, i.e., glycoproteins with reduced or modified glycosylation, are also part of the present invention.

Still another aspect of the present invention provides kits which contain at least one of the knock-in vectors, inactivation vectors, or knock-in-and-inactivation vectors of the present invention described above. More particularly, a kit of the present invention contains at least one of the following vectors: a vector capable of expressing an α-mannosidase I in a methylotrophic yeast, a vector capable of expressing a glucosidase II in a methylotrophic yeast, a vector capable of expressing a GnTI in a methylotrophic yeast, a vector capable of expressing a mannosidase II in a methylotrophic yeast and a vector capable of inactivating the OCH1 gene in a methylotrophic yeast.

The kit can also include a nucleotide sequence which encodes and is capable of expressing a heterologous glycoprotein of interest. Such nucleotide sequence can be provided in a separate vector or in the same vector which contains sequences for knocking-in or inactivating as described hereinabove.

In addition, the kit can include a plasmid vector in which a nucleotide sequence encoding a heterologous protein of interest can be subsequently inserted for transformation into and expression in a methylotrophic yeast. Alternatively, the knock-in, inactivation or knock-in-and-inactivation vectors in the kits have convenient cloning sites for insertion of a nucleotide sequence encoding a heterologous protein of interest.

The kit can also include a methylotrophic yeast strain which can be subsequently transformed with any of the knock-in, inactivation or knock-in-and-inactivation vectors described hereinabove. The kit can also include a methylotrophic yeast strain which has been transformed with one or more of the knock-in, inactivation or knock-in-inactivation vectors. Furthermore, the kit can include a methylotrophic

yeast strain which has been transformed with a nucleotide sequence encoding and capable of expressing a heterologous glycoprotein of interest.

The present invention is further illustrated by the following examples.

Example 1

Introduction of α-1,2-Mannosidase to the ER-Golgi Border

1.1 Plasmids

Plasmid	Promoter	Enzyme	Tag
pGAPZMFManHDEL	GAP	*T. reesei* α-1,2-mannosidase	--
pGAPZMFManMycHDEL	GAP	*T. reesei* α-1,2-mannosidase	Myc
pPICZBMFManMycHDEL	AOX1	*T. reesei* α-1,2-mannosidase	Myc
pGAPZMFmManHDEL	GAP	mouse mannosidase IB catalytic domain	--
pGAPZMFmMycManHDEL	GAP	mouse mannosidase IB catalytic domain	Myc

The *Trichoderma reesei* α-1,2-mannosidase gene has been isolated and described by Maras et al. (*J. Biotechnol.* 77;255-263, 2000). The sequence of this gene is available at NCBI Genbank under Accession No. AF212153. A construction fragment was generated by PCR using the pPIC9MFmanase plasmid (same as pPP1MFmds1 described by Maras et al. (2000)) as the template and using the following oligonucleotide primers: 5'-GACTGGTTCCAATTGACAAGC-3' (SEQ ID NO:2) and 5'-AGTCTAGATTACAACTCGTCGTGAGCAAGGTGGCCGCCCCG TCG-3' (SEQ ID NO:3). The resulting product contained the 3' end of the *Pichia pastoris* AOXI promoter, the prepro-signal sequence of the *S. cerevisiae* α-mating factor, the open reading frame of the *Trichoderma reesei* α-1,2-mannosidase cloned in frame with the signal sequence, the coding sequence for HDEL, a stop codon and an *Xba* I restriction site. This fragment was digested with *Eco* RI and *Xba* I, removing the 5' sequences up to the mannosidase ORF, and then cloned into the vector pGAPZαA (Invitrogen, Baarn, The Netherlands) which had been digested with *Eco* RI

and *Xba* I, thus restoring the fusion with the *S. cerevisiae* α-mating factor signal sequence. The resulting plasmid was named pGAPZMFManHDEL and is graphically depicted in **Figure 1**. The ORF sequence of the MFManHDEL fusion in pGAPZMFManHDEL is set forth in SEQ ID NO: 14.

In order to introduce the coding sequence for a c-Myc tag between the catalytic domain and the HDEL-signal, the 3' end of the ORF of *T. reesei* α-1,2-mannosidase was PCR-amplified using a sense primer 5'-CCATTGAGGACGCATGCCGCGCC-3' (SEQ ID NO: 4) (containing an *Sph* I restriction site) and an antisense primer GTATCTAGATTACAACTCGTCGTGCAGATCCTCTTCTGAGATGAGTTTTTGT TCAGCAAGGTGGCCGCCCCGTCGTGATGATGAA (SEQ ID NO: 5) (containing the coding sequences of the c-Myc tag and the HDEL signal, followed by a stop codon and an *Xba* I restriction site). The resulting PCR product was digested with *Sph* I and *Xba* I, purified by agarose gel electrophoresis and inserted into pGAPZMFManHDEL which had been cut with the same restriction enzymes, resulting in plasmid pGAPZMFManMycHDEL. To put the ORF of pGAPZMFManMycHDEL under the control of the inducible AOXI promoter, the entire ORF was liberated from pGAPZMFManMycHDEL with *Bst* BI and *Xba* I, and cloned in pPICZB (Invitrogen, Baarn, The Netherlands), resulting in pPICZBMFManMycHDEL.

Cloning of the mouse mannosidase IB catalytic domain with concomitant addition of the coding sequence for a C-terminal HDEL-tag was done by PCR on a mouse cDNA library (mRNA isolated from the L929 cell line induced with cycloheximide and mouse Tumor Necrosis Factor. Average insert length of the cDNA library was 2000 bp). The PCR oligonucleotide primers used were: 5'AACTCGAGATGGACTCTTCAAAACACAAACGC3' (SEQ ID NO: 6) and 5'TTGCGGCCGCTTACAACTCGTCGTGTCGGACAGCAGGATTACCTGA3' (SEQ ID NO: 7). The product contained a 5' *Xho* I site and the coding sequence for C-terminal HDEL-site, followed by a stop codon and a *Not* I site at the 3' end. The

product was cloned in pGAPZαA via the *Xho* I /*Not* I sites in the PCR product and the vector, resulting in an in frame fusion of the mouse mannosidase catalytic domain with the *S. cerevisiae* α-mating factor signal sequence. The sequence of the entire open reading frame generated is set forth in SEQ ID NO: 15.

5

1.2 Yeast Transformation and Genomic Integration

Table 2

Parental strain	DNA transformed
GS115 (his4)	pGAPZMFManHDEL
	pPIC9MFManHDEL
	pPIC9mManHDEL
	pPIC9mMycManHDEL
	pGAPZmManHDEL
	pGAPZmMycManHDEL
GS115 (his4 complemented by pPIC9InfluenzaHA)	pGAPZMFManHDEL
	pGAPZmManHDEL
	pGAPZmMycManHDEL
PPY120H (his4 complemented by pPIC9sOCH1)	pGAPZMFManMycHDEL
	pPICZBMFManMycHDEL
yGC4 (his4 arg1 ade2 ura3 complemented by pBLURA5'PpOCH1)	pPIC9InfluenzaNeuraminidase
	pGAPZMFManHDEL
	pPIC9Glucoseoxidase

All transformations to *Pichia pastoris* were performed with electroporation according to the directions of Invitrogen. Transformants of vectors carrying the Zeocin resistance gene were selected on YPD containing 100 µg/ml Zeocine

(Invitrogen, Baarn, the Netherlands) and 1M sorbitol. Selection of transformants of pPIC9 derivatives was done on minimal medium lacking histidine and containing 1M sorbitol. Genomic integration of the expression cassettes was verified using PCR on genomic DNA purified from the *Pichia* strains using the Yeast Miniprep method (Nucleon). In all cases concerning the *Trichoderma reesei* gene fusions, the primers used were the sense primer 5'-CCATTGAGGACGCATGCCGCGCC-3' (SEQ ID NO: 8), which annealed to the 3' half of the mannosidase ORF, and the antisense primer 3' AOXI 5'-GCAAATGGCATTCTGACATCCT-3' (SEQ ID NO: 9), which annealed to the AOXI transcription terminator that was present in all our expression constructs. For the control of genomic integration of the mouse mannosidase transgenes, PCR was done using the sense primer 5'GAP 5'GTCCCTATTTCAATCAATTGAA3' (SEQ ID NO: 10, annealing to the GAP promoter or 5'AOXI 5'GACTGGTTCCAATTGACAAGC3' (SEQ ID NO:11), annealing to AOXI promoter), and the antisense primer 3'AOXI (above). For the expression constructs containing a Myc tagged *Trichoderma reesei* α-1,2-mannosidase expression unit, further evidence for genomic integration was obtained using Southern Blotting with the entire MFManMycHDEL ORF (^{32}P labelled using HighPrime, Boehringer Mannheim) as a probe.

1.3 Expression of α-1,2-mannosidase

Measurement of intracellular class I mannosidase activity.

Lysate preparation: yeast cells (*P. pastoris* GS115 wild type and clones of the same strain transformed with pGAPZMFManHDEL (Callewaert *et al.*, 2001)) were grown in YPD to an OD_{600}=2-2.5, in a volume of 10 ml. After pelleting of the cells by centrifugation at 3000 x g for 5 min, the cells were washed twice in sterile water and subsequently resuspended in 500 µl of the lysis buffer: 100mM NaAc, pH 5, containing EDTA-free Complete protease inhibitor cocktail (Boehringer) in the dilution specified by the supplier and 1% Triton X100. After transfer of the cell

suspension to a 1.5 ml eppendorf tube, glass beads were added to the meniscus of the tube and the tubes were shaken in a Retsch cell disrupter for 5 x 1 min., with 1 min. pauses on ice in between the vortex steps, to avoid over-heating. Finally, the cell debris was precipitated by centrifugation at 5000 rpm in a microcentrifuge and the supernatant was used for protein concentration determination (Pierce BCA kit, the sample was diluted tenfold prior to measurement in order to lower the concentration of Triton, which may interfere with this protein concentration measurement procedure). The protein concentration in lysates obtained as described above should average about 3.5 mg/ml.

Activity measurement: samples containing 375 µg protein were volume-adjusted to 150 µl with 100 mM NaAc, pH 5 and incubated at 50°C for 1h. This denatured most of the proteins in the preparation, yielding a milky white appearance of the solution after incubation. The *Trichoderma reesei* α-1,2-mannosidase was heat-stable under these conditions. The denatured proteins were precipitated by centrifuging at full speed for 10 min in a microcentrifuge (20000 x g). The supernatant was transferred to wells in a standard microtiter plate. Subsequently, 50 µl of a working solution was added, containing 250 µM swainsonine (inhibitor for a specific class II mannosidases), 4 mM DNP-mannose and 2 mM $CaCl_2$. As blank control, 150 µl of the lysis buffer was used, supplemented with 50 microliter of the working solution. After incubation at 50°C for 1h, the absorbance was measured at 405 nm in a standard microtiter plate reader. The reported values in **Figure 31** are blank-substracted.

Expression of an α-1,2-Mannosidase in GS115 strains expressing influenza virus haemagglutinin was verified by qualitative Northern blot. Expression of an α-1,2-Mannosidase in PPY120H strains was verified by anti-Myc Western blot.

Qualitative Northern Blot -- Total RNA was purified from *Pichia* strains and the yield was determined spectrophotometrically. Northern blotting was

performed according to standard procedures and an estimate of the quantity of RNA loaded was made using methylene blue staining of the blot, visualizing the rRNA bands. The blot was probed with a *ClaI/Nar*I fragment of the mannosidase, labelled with ^{32}P using HighPrime (Boehringer Mannheim).

SDS-PAGE and Western Blotting -- Total yeast cell lysates were prepared by washing the cells twice with PBS, followed by boiling in 1 volume of 2x concentrated Laemmli loading buffer for 5 min. The lysate was spun briefly in a microcentrifuge prior to gel loading and only the supernatant was loaded. For the analysis of proteins secreted into the growth media, the proteins were precipitated from 200 µl of these media using desoxycholate/trichloroacetic acid according to standard procedures. The pellet was redissolved in 2x concentrated Laemmli loading buffer and the solutions were pH-corrected using Tris. SDS-PAGE was performed and followed by semidry electroblotting to nitrocellulose membranes. For Western Blotting, the 9E10 anti-Myc and the anti-HA mouse monoclonals (Boehringer Mannheim) were used at a concentration of 1µg/ml, and the rabbit anti-PDI antiserum (Stressgen) was used at a dilution of 1/500. The secondary antibodies were goat anti-mouse IgG conjugated to alkaline phosphatase for the monoclonals and goat anti-rabbit IgG conjugated to peroxidase for the polyclonal (secondary antibodies from Sigma). Detection was performed using the NBT/BCIP system for alkaline phosphatase and the Renaissance substrate (NENBiosciences) for peroxidase. Imaging of the latter blot result was done on a Lumilager imaging device (Boehringer Mannheim).

The results shown in **Figure 4** indicated that the great majority of the HDEL-tagged protein was retained intracellularly, both when expressed from the strong constitutive GAP promoter and when expressed from the strong inducible AOXI promoter.

1.4 Localization of α-1,2-Mannosidase

Isopycnic sucrose density gradient centrifugation -- To determine the localization of the HDEL-tagged mannosidase, subcellular fractionation was carried out using cells expressing the mannosidase-Myc-HDEL from the strong constitutive GAP promoter.

Briefly, 0.5 g of wet weight yeast cells were lysed using 4 x 1 min vortexing with 4.5 g glass beads in 1 ml lysis-buffer (50 mM Tris-HCL pH 7.5 containing 0.6 M sorbitol, 10 mM β-mercaptoethanol and 5 mM $MgCl_2$). Between vortexing periods, the mixture was placed on ice for 5 min. The supernatant was collected and the glass beads were washed once with lysis-buffer, and the supernatant of this washing step was added to the first supernatant. This lysate was subjected to a differential centrifugation procedure. The P10000 pellet was solubilized in 0.5 ml of a 60% sucrose solution in lysis buffer. This solution was placed at the bottom of an Ultraclear ultracentrifuge tube (Beckman) of 14 x 89 mm. Subsequently, 1.5 ml each of sucrose solutions of 55, 50, 45, 42.5, 40, and 37.5% were carefully layered over each other. The tube was filled to the edge with 35% sucrose. Isopycnic sucrose gradient centrifugation was performed for 14 h at 180,000g in a Beckman SW 41 rotor in a Beckman Model L8-70 preparative ultracentrifuge. After completion, 1ml fractions were collected from the top and partially dialysed from excess sucrose, evaporated to dryness in a vacuum centrifuge. After redissolving the pellet in Laemmli buffer, the samples were subjected to SDS-PAGE in triplicate and the Western blots were treated with anti-HA, anti-Myc or anti-PDI ("PDI" for Protein Disulfide Isomerase), respectively.

The results illustrated almost exact cosedimentation of the MFManMycHDEL protein with the Protein Disulfide Isomerase marker protein (which is also targeted with a HDEL signal sequence) (**Figure 5**). In the same assay, the HA-tagged OCH1 was distributed over the whole gradient, with the highest abundance in fractions having a density lower than that of the fractions containing the mannosidase and the PDI. This result indicated that the mannosidase was targeted to

the expected location (the ER-Golgi boundary) by the addition of an HDEL signal. In contrast, the mannosidase without HDEL, expressed from inducible alcohol oxidase I promoter (which was of comparable strength as the GAP promoter), was secreted at a high level of about 20 mg/l.

Immunofluorescence microscopy -- To confirm the correct targeting of the mannosidase-Myc-HDEL, an immunofluorescence microscopy experiment was performed.

Briefly, yeast cultures were grown to OD_{600} in YPD (for pGAPZMFManMycHDEL) or in YMP following a YPGlycerol growth phase for pPICZBMFManMycHDEL. Formaldehyde was added to the yeast cultures to a final concentration of 4% and incubated for 10 min at room temperature. Cells were pelleted and resuspended in 50mM potassium phosphate buffer pH 6.5 containing 1mM $MgCl_2$ and 4% formaldehyde and incubated for 2h at room temperature. After pelleting, the cells were resuspended to an OD_{600}=10 in 100 mM potassium phosphate buffer pH 7.5 containing 1mM $MgCl_2$ and EDTA-free Complete™ protease inhibitor cocktail (Boehringer Mannheim). To 100 µl of cell suspension, 0.6 µl of β-mercapto-ethanol and 20µl of 20,000 U/ml Zymolyase 100T (ICN) were added, followed by a 25 minute incubation with gentle shaking. The cells were washed twice in the incubation buffer and added to poly-lysine coated cover slips (these are prepared using adhesive rings normally in use for reinforcing perforations in paper). Excess liquid was blotted with a cotton swab and the cells were allowed to dry at 20°C. All blocking, antibody incubation and washing steps are performed in PBS containing 0.05% bovine serum albumin. Primary antibodies are used at 2µg/µl and secondary antibodies conjugated to flurophores (Molecular probes) were used at 5µg/µl. The nucleus was stained with the nucleic acid stain HOECHST 33258. After fixation and cell wall permeabilization, the integrity of the yeast cell morphology was checked in phase contrast microscopy and after immunostaining, the slides were examined under a Zeiss Axiophot fluroresensce microscope equipped with a Kodak digital camera.

Images were processed using Macprobe 4.0 software and prepared with Corel Photopaint 9.0.

The Golgi marker protein OCH1-HA gave the typical Golgi staining pattern described in the literature (speckle-like staining). Staining with the 9E10 monoclonal anti-Myc antibody, recognizing mannosidase-Myc-HDEL, gave a perinuclear staining pattern with some disparate staining in the cytoplasm, highly indicative for an ER targeting (**Figure 4**).

Based on the foregoing experiments, it is concluded that the *Trichoderma reesei* mannosidase-Myc-HDEL was targeted to the ER-Golgi boundary.

Example 2

Co-expression of Mannosidase-HDEL with Recombinant Glycoproteins

Co-expression of Mannosidase-HDEL with the *Trypanosoma cruzi trans*-Sialidase

The cloning of a *Trypanosoma cruzi trans*-sialidase gene coding for an active *trans*-sialidase member without the C-terminal repeat domain has been described by Laroy et al. (*Protein Expression and Purification* 20: 389, 2000) which is incorporated herein by reference. The sequence of this *Trypanosoma cruzi trans*-sialidase gene is available through NCBI Genbank under the Accession No. AJ276679. For expression in *P. pastoris*, the entire gene was cloned in pHILD2 (Invitrogen, San Diego, CA), creating pHILD2-TS. To allow better secretion, pPIC9-TS was created in which *trans*-sialidase was linked to the prepro secretion signal of the yeast α–mating factor. Plasmids pPIC9-TSE and pCAGGS-prepro-TSE were created where the epitope E-tag was added to the C-terminal of the *trans*-sialidase to allow easy detection and purification. The construction of pHILD2-TS, pPIC9-TSE and pCAGGS-prepro-TSE has been described by Laroy et al. (2000), incorporated herein by reference. The vectors used in the construction were made available through http://www.belspo.be/bccm/lmbp.htm#main for pCAGGS (No. LMBP 2453), Invitrogen, San Diego, CA for pHILD2 and pPIC9, and Pharmacia Biotech for pCANTAB-5E.

Plasmid pPIC9-TSE was linearized with SstI and was transformed into *P. pastoris* GS115 (*his4*) strain by electroporation according to the manufacturer's instructions (Invitrogen). One of the transformants was further transformed with plasmid pGAPZMFManHDEL, establishing a strain co-expressing Mannosidase-HDEL and the *Trypanosoma cruzi trans*-sialidase.

Fermentation and protein purification was according to the procedures described by Laroy et al. (2000).

Purified *trans*-sialidase was subject to carbohydrate analysis according to Callewaert et al., *Glycobiology* 11, 4, 275-281, 2001. Briefly, the glycoproteins were bound to the PVDF membrane in the wells of a 96-well plate, reduced, alkylated and submitted to peptide-N-glycosidase F deglycosylation. The glycans were derivatised with 8-amino-1,3,6-pyrenetrisulfonic acid by reductive amination. Subsequently, the excess free label was removed using Sephadex G10-packed spin columns and the glycans were analysed by electrophoresis on a 36 cm sequencing gel on an ABI 377A DNA-sequencer and detected using the built-in argon laser. Digests with 3 mU/ml purified *T. reesei* α-1,2-mannosidase (described by Maras et al., *J. Biotechnol.* 77, 255-63, 2000) were also performed in 20 mM sodium acetate pH=5.0. The glycans derived from 1 μg of the purifed recombinant glycoproteins were used as the substrate. 1U of the α-1,2-mannosidase is defined as the amount of enzyme that releases 1 μmol of mannose from baker's yeast mannan per minute at 37°C and pH=5.0.

As can be seen in **Figure 6**, panel B, the major N-glycan on *trans*-sialidase was $Man_8GlcNAc_2$ (Compare with panel F, representing an analysis of the N-glycans of bovine RNAseB. The one but last peak in this profile is $Man_8GlcNAc_2$, the first peak is $Man_5GlcNAc_2$). *In vitro,* this glycan was digestible to $Man_5GlcNAc_2$ with α-1,2-mannosidase (**Figure 6**, panel C). In the N-glycan profile of the *trans*-sialidase co-expressed with mannosidase-HDEL, the major peak corresponded to $Man_5GlcNAc_2$ (**Figure 6**, panel D).

Co-expression of Mannosidase-HDEL with the Influenza A virus haemagglutinin

The Influenza A virus haemagglutinin was known to be glycosylated in *Pichia pastoris* with high-mannose N-glycans containing 9-12 mannose residues (Saelens et al. *Eur. J. Biochem.* 260: 166-175, 1999). The effect of a co-expressed mannosidase on the N-glycans of the haemagglutinin was assessed in an N-glycan

profiling method described below. In addition, to compare the efficiency of the *Trichoderma* enzyme (having a temperature optimum of 60°C) with a mammalian mannosidase having a temperature optimum of 37°C, the catalytic domain of the mouse mannosidase IB from a mouse cDNA-library was cloned and tagged with a HDEL signal by PCR amplification. This ORF was cloned after the prepro-signal sequence of the *S. cerevisiae* α-mating factor under the control of the GAP promoter. Expression of the mannosidase-HDEL transgenes on the mRNA level was confirmed by qualitative Northern blotting.

The haemagglutinin was expressed and purified from a non-mannosidase expressing control strain and from a strains co-expressing the *Trichoderma reesei* mannosidase-HDEL or the mouse mannosidase IB-HDEL according to the procedure described by Kulakosky et al. *Glycobiology* 8: 741-745 (1998). The purified haemagglutin was subjected to PNGase F digestion as described by Saelens et al. *Eur. J. Biochem.* 260: 166-175, 1999. The proteins and glycans were precipitated with 3 volumes of ice-cold acetone and the glycans were extracted from the pellet with 60% methanol. Following vacuum evaporation, the glycans were labeled with 8-amino-1,3,6 pyrenetrisulfonic acid by adding 1 µl of a 1:1 mixture of 20 mM APTS in 1.2M citric acid and 1M N_aCNBH_3 in DMSO and incubating for 16h at 37°C at the bottom of a 250 µl PCR-tube. The reaction was stopped by the addition of 10 µl deionized water and the mixture was loaded on a 1.2 cm Sephadex G10 bed packed to dryness in a microspin-column by centrifugation in a swinging bucket rotor, which provided for a flat resin surface. After loading, 50 µl deionised water was carefully added to the resin bed and the spin column was briefly centrifuged for 5 seconds at 750g in a tabletop centrifuge. This elution process was repeated twice and all the eluates were pooled and evaporated to dryness in a Speedvac vacuum centrifuge (Savant). The labeled glycans were reconstituted in 1.5 µl gel loading buffer containing 50% formamide and 0.5 µl Genescan 500™, labeled with rhodamine (Perkin Elmer Bioscience), serving as an internal reference standard. This mixture was loaded on a

DNA-sequencing gel containing 10% of a 19:1 mixture of acrylamide:bisacrylamide (Biorad, Hercules, CA, USA) and made up in the standard DNA-sequencing buffer (89 mM Tris, 89 mM borate, 2.2 mM EDTA). Polymerization of the gel was catalyzed by the addition of 200 µl 10% ammononiumpersulfate solution in water and 20 µl TEMED. The gel was of the standard 36 cm well-to-read length and was run on an Applied Biosystems Model 373A DNA-sequencing apparatus. Prerunning of the gel was done at 1000 V for 15 min. and after loading, the gel was electrophoresed for 8h at 1250 V without heating. This methodology gives a limit of detection of 10 fmol per peak. The data were analysed with Genescan 3.0 software.

As shown in **Figure 7**, the *Trichoderma reesei* α-1,2-mannosidase provided the most complete reduction in the number of α-1,2-mannoses present on the N-glycans. The N-glycan processing by mouse mannosidase IB-HDEL was less efficient than by the *Trichoderma reesei* α-1,2-mannosidase.

Despite the efficient removal of α-1,2-mannoses from the N-glycans of haemagglutinin, no $Man_5GlcNAc_2$ was obtained. Even after digestion of the N-glycans with 3 mU of purified *Trichoderma reesei* α-1,2-mannosidase, only $Man_6GlcNAc_2$ was obtained as the smallest sugar chain. These results indicated that the remaining residues were possibly α-1,6-linked mannoses, originating from the initiating OCH1 α-1,6-mannosyltransferase enzymatic activities. OCH1 was observed to be localized to very early part of the Golgi apparatus and could act on the N-glycans of haemagglutinin before complete digestion of the $Man_8GlcNAc_2$ precursor to $Man_5GlcNAc_2$ by the mannosidases-HDEL. Thus, for proteins whose glycans are efficiently modified by the α-1,6-mannosyltransferase, an inactivation of the OCH1 gene coding for the transferase would be desirable in order to obtain proteins with $Man_5GlcNAc_2$.

Example 3

Inactivation of the *Pichia* OCH1 Gene:

A *Pichia pastoris* sequence was found in the GenBank under Accession No. E12456 and was described in Japanese Patent Application No. 07145005, incorporated herein by reference. This sequence shows all typical features of an α-1,6-mannosyltransferase and is most homologous to the *S. cerevisiae* OCH1, thus referred to herein as the *Pichia pastoris* OCH1 gene.

First, the full ORF of the *Pichia pastoris* OCH1 gene was PCR cloned in pUC18 to obtain plasmid pUC18pOCH1. pUC18pOCH1 was cut with HindIII, blunt-ended with T4 polymerase, then cut with XbaI, releasing a fragment containing the 5' part of the *Pichia pastoris* OCH1 gene. This fragment was ligated into the vector pBLURA IX (available from the Keck Graduate Institute, Dr. James Cregg, http://www.kgi.edu/html/noncore/faculty/cregg/cregg.htm), which had been cut with *Eco* RI, blunt-ended with T4 polymerase, and then cut with *Nhe* I. This ligation generated pBLURA5'PpPCH1, as shown in **Figure 8**.

Disruption of this *Pichia* OCH1 gene in the *Pichia* genome was achieved by single homologous recombination using pBLURA5'PpOCH1, as illustrated in **Figure 9**. As a result of the single homologous recombination, the OCH1 gene on the *Pichia* chromosome was replaced with two OCH1 sequences: one consisted only about the first one third of the full OCH1 ORF, the other had a full OCH1 ORF without a OCH1 promoter. Single homologous recombination was achieved as follows. Cells of the *Pichia* strain yGC4 were transformed by electroporation with pBLURA5'PpOCH1 which had been linearized with the single cutter *Bst* BI. About 500 transformants were obtained on minimal medium containing 1M sorbitol, biotin, arginine, adenine and histidine and incubation at 27°C. Thirty-two of these transformants were picked and re-selected under the same conditions. Twelve clones

were further analyzed for correct genomic integration of the cassette by PCR. Seven of the twelve URA prototrophic clones contained the cassette in the correct locus.

One of the OCH1-inactivated clones was also further transformed with pGAPZMFManHDEL to produce "supertransformants". Both the OCH1-inactivated clone and three supertransformants also expressing the ManHDEL were evaluated in cell wall glycan analysis as follows. Yeast cells were grown in 10 ml YPD to an OD_{600}=2 and mannoproteins were prepared by autoclaving the yeast cells in 20 mM sodium citrate buffer pH7 for 90 min at 120 °C and recovery of the supernatant after centrifugation. Proteins were precipitated from this supernatant with 3 volumes of cold methanol. The protein preparation obtained in this way was used for N-glycan analysis using DSA-FACE as described by Çallewaert et al. (2001) *Glycobiology* 11, 275-281. As shown in **Figure 10**, there was an increased amount of $Man_8GlcNAc_2$ glycan in the OCH1-inactivanted clone as compared to parent strain yGC4, indicative of a reduced activity of the OCH1 enzyme. In all three supertransformants which also expressed the HDEL-tagged α–1,2 mannosidase, the production of $Man_5GlcNAc_2$ was observed. Furthermore, upon digestion of the same glycan mixtures with 3 mU/ml purified recombinant *Trichoderma reesei* α-1,2-mannosidase, more $Man_5GlcNAc_2$ was formed in the strain transformed with pBLURA5'PpOCH1 than in the parent strain (**Figure 11**, compare panel 2 and 3).

These results indicated that the production of glycoproteins with Man_5 glycans could be facilitated by the inactivation of the OCH1 gene.

Mutagenesis PCR was performed on plasmid pPICZB5'PpOCH1Mut (plasmid construction described in Example 8) using the forward primer Kai13: 5'-ATCTAAgCTATATTCgCCgTTTCTgTCATT-3' (SEQ ID NO: 43) and the reverse primer Kai 14: 5'- CgATTATggATgTTAgATCTgATCTCATgA-3' (SEQ ID NO: 44). This PCR resulted in a product that lacked the entire AOX1 promoter and the coding sequence for the first 24 amino acids of the OCH1 fragment in

pPICZB5'PpOCH1Mut. This product was self-ligated and the resulting plasmid was named pZ5'PpOCH1Trunc (SEQ ID NO: 45).

This plasmid was linearized with BstBI and the product was transformed to *Pichia pastoris* GS115 and yGC4. Transformants were selected on YPDS-Zeo and the genomic integration was tested using 2 PCR reactions: one with the primer couple Kai10 (SEQ ID NO: 41)/Kai7 (SEQ ID NO: 40) and one with the couple Kai5 (SEQ ID NO: 39)/Kai 11 (SEQ ID NO: 42). The primer hybridisation sites can be seen on **Figure 32.** In **Figure 32,** the desirable genomic situation at the OCH1 locus after transformation is shown.

24 clones were analysed using these PCR reactions and the result for 8 clones is shown in **Figure 34.** We found that at least 8/24 clones had the plasmid correctly integrated in the OCH1 locus.

All clones with a correct integration were vanadate resistant (YPD plates containing 5 mM sodium vanadate, growth at 30 °C), a typical hallmark for N-glycosylation defective yeast cells (Proc Natl Acad Sci U S A 1991 Apr 15;88(8):3209-12).

Example 4

Expression of Glucosidase II in *Pichia pastoris*

4.1 Amplification of the GLSII alpha subunit ORF from *S. cerevisiae*.

Genomic DNA was prepared from the *S. cerevisiae* strain INVS (α, leu2-3, 112 his3∈1, trpl-289, ura3-52), using the Nucleon kit (Amersham). A touch-down PCR reaction was performed using this genomic DNA as template and the LA TaKaRa polymerase (ImTec Diagnostics). The sequence of the PCR primers was based on the known sequence of the *S. cerevisiae* GLSII ORF:

Sense primer: 5' CCG CTC GAG (Xho I) ATG GTC CTT TTG AAA TGG CTC 3'

(SEQ ID NO:12)

Antisense primer: 5' CCG GGC CCA (Apa I) AAA ATA ACT TCC CAA TCT TCA G 3' (SEQ ID NO:13)

4.2 Cloning of the *S. cerevisiae* glucosidase II ORF into *Pichia pastoris* expression vectors.

Construction of the glucosidase II expression vectors -- The PCR fragment was digested with *Xho I*/*Apa* I and ligated into the pGAPZA vector (Invitrogen), thereby placing the ORF under the transcriptional control of the GAP promoter. Using this strategy, the myc and the His6 tag were placed in frame to the C-terminus of Glucosidase II, creating pGAPZAGLSII. The complete ORF of pGAPZAGLSII was then sequenced to ensure that no mutations were generated in the PCR reaction. The sequence of the vector pGAPZAGLSII was set forth in SEQ ID NO: 18. The GLSII ORF from the pGAPZAGLSII vector was cloned into vector pPICZA (Invitrogen) to create pPICZAGLSII, thereby placing the ORF under the transcriptional control of the AOXI promoter. The GLSII ORF from the

pGAPZAGLSII vector was cloned into vector pAOX2ZA, thereby placing the ORF under the transcriptional control of the AOX2 promoter. This vector was created by replacing the multi cloning site of vector pAOX2ZB with the multi cloning site of pPICZA. Vector pAOX2ZB was generated by replacing the AOX1 promotor of pPICZB by the AOX2 promotor region of the AOX2 gene (Martinet et al., Biotechnology Letters 21). The AOX2 promotor region was generated by PCR on *Pichia* genomic DNA with the sense primer 5'GACGAGATCTTTTTTTCAGACCATATGACCGG 3' (SEQ ID NO: 26) and the antisense primer 5'GCGGAATTCTTTTCTCAGTTGATTTGTTTGT 3' (SEQ ID NO: 27). The GLSII ORF from the pGAPZGLSII vector was cloned into vector pYPT1ZA to create pYPTIZAGLSII, thereby placing the ORF under the transcriptional control of the YPT1 promoter. Vector pYPTZA was created by replacing the AOX1 promoter of pPICZA by the YPT1 promoter present on the plasmid pIB3 (GenBank accession number AF027960)(Sears et al., Yeast 14, pg 783-790, 1998). All constructs contain the phleomycin resistance gene. The resulting final expression vectors (pGAPZAGLSII, pAOX2ZAGLSII, pPICZAGLSII and pYPT1ZAGLSII) are depicted in **Figures 12-15.**

Similar expression vectors were constructed, carrying the Ampicillin resistance marker and the *Pichia* ADE1 selection marker. In principle, the Zeocin resistance expression cassette of the plasmids pAOX2ZAGLSII, pGAPZAGLSII and pYPT1ZAGLSII was replaced by the Ampicillin and *Pichia* ADE1 cassette of the vector pBLADE IX (Cregg, J.M.) to result in the vectors pAOX2ADE1glsII, pGAPADE1glsII and pYPT1ADE1glsII. Vector pPICADE1glsII was obtained by inserting the glucosidase II open reading frame into the multiple cloning site of the vector pBLADE IX (Cregg, J.M.). The resulting final expression vectors (pGAPADE1glsII, pAOX2ADE1glsII, pPICADE1glsII and pYPT1ADE1glsII) are depicted in **Figures 16-20.**

Adding the ER retention tag HDEL to Glucosidase II expression vectors – The following primers were used to generate an HDEL-containing PCR fragment:

Primer 1: 5'GCG GGT CGA *C/CA C/GA C/GA A/CT G/TG A/*GT TTT AGC CTT
Sal I H D E L stop
AGA CAT GAC 3' (SEQ ID NO:28)
Primer 2: 5'CAG GAG CAAA GCT CGT ACG AG 3' (SEQ ID NO:29)
Spl I

PCR was performed on pGAPZAGLSII with Taq pol., at 60°C. The PCR fragment of 225 bp was cut with *Sal I/Spl* I and ligated into the *Sal I/Spl* I opened pGAPZAGLSII vector, creating plasmid pGAPZAglsIIHDEL. The sequence of plasmid pGAPZAglsIIHDEL is set forth in SEQ ID NO: 24. The construction strategy and the resulting final expression vectors (pGAPADE1glsIIHDEL and pGAPZAglsIIHDEL) are depicted in **Figures 20-21.**

4.3 Transformation of a *Pichia pastoris* strain.

Transformation was performed using the conventional electroporation techniques, as described by Invitrogen. Cells of the *Pichia pastoris* strain PPY12-OH were transformed with pGAPZGLSII which had been cut with the single cutter *Avr* II. Transformants were selected based on their resistance to zeocin.

Genomic analysis of the transformants -- Genomic DNA was prepared from some zeocin resistant *Pichia* transformants. A PCR reaction was performed on the genomic DNA in order to determine whether or not the glucosidase II gene was integrated into the yeast genome. PCR was performed using Taq DNA polymerase (Boehinger) (2.5 mM $MgCl_2$, 55°C for annealing). The primers were the same as the ones we used for the amplification of the ORF on *S. cerevisiae* genomic DNA. pGAPZAGLSII transformants were confirmed by the presence of a specific PCR product indicative of the glucosidase II ORF.

4.4 Expression and secretion of the *S. cerevisiae* glucosidase II alpha subunit in *Pichia pastoris*

Analysis at the transcriptional level -- RNA was prepared from the transformants which scored positive after the genomic analysis. RNA was prepared using acid phenol. From each sample, 15 µg of RNA was loaded on a formaldehyde agarose gel. After electrophoresis the RNA was blotted on a Hybond N membrane. The membrane was hybridized using a radioactive probe, which consists of a 344 bp glucosidase II specific fragment, corresponding to the 3' region of the glucosidase II ORF. No signals were detected with non-transformed control strains, whereas clear signals were observed with transformants.

Analysis at the protein level using a double membrane assay -- A nitrocellulose membrane was placed on a buffered dextrose medium (BMDY). On top of that nitrocellulose membrane, a cellulose acetate membrane was placed. *Pichia* transformants of pGAPZAGLSII were streaked on the cellulose acetate and grown for a few days. The yeast cells remained on the cellulose acetate, while the secreted proteins crossed this membrane. As such the secreted protein was captured onto the nitrocellulose membrane. After a few days the cellulose acetate, containing the yeast colonies, was removed. The nitrocellulose membrane was analyzed for the presence of glucosidase II using anti-myc antibody. Most of the transformants gave a clear signal as compared to a faint, hardly visible signal with the WT, non-transformed strain.

Extracellular expression – PPY12-OH transformants of the construct pGAPZAGLSII(mychis6) (strains 12, 14 and 18) and transformants of the construct pGAPZAGLSII(myc)HDEL (strains H1, H2 and H3) were grown for 2 days on 2x10 ml BMDY medium. These 6 transformants earlier scored positive both on the genomic level (PCR on gDNA) and on the RNA level (Northern blot). The culture medium was collected by centrifugation and concentrated with Vivaspin columns to about 1 ml. Proteins from this concentrate were precipitated with TCA, resuspended

in Laemmli buffer and loaded for SDS-PAGE analysis. Proteins were blotted to nitrocellulose membrane. The blot was incubated overnight with anti-myc Ab. The secondary Ab was linked to peroxidase. Using the Renaissance luminiscence detection kit (NEN) and a light sensitive film (Kodak), a strong band at about 110 kDa was observed for the transformants 12, 14 and 18, indicating that GLSII was expressed and secreted from these transformants. No signal was obtained for the transformants H1-3, which indicate that the HDEL tag, which was added C-terminally to the GLSII ORF, resulted in an ER localization of the protein, preventing GLSII to be secreted into the growth medium.

Intracellular expression – The 6 transformants and the WT strain were grown for 2 days in 500 ml BMDY. The cells were collected by centrifugation, washed, resuspended into a minimal volume (50 mM Tris.HCl pH 7.5, 5% glycerol) and broken using glass beads. The cell debris was removed through several centrifugation steps (low speed centrifugation (2000-3000g)). Membranes were obtained from the supernatant through ultracentrifugation. The pellets were resuspended in Laemmli buffer and loaded for SDS-PAGE analysis. The proteins were blotted on a nitrocellulose membrane. The intracellular GLSII expression was checked using anti-myc Ab and peroxidase conjugated secondary Ab. Following the luminescence detection, a band of about 110 kDA was observed with the GLSIIHDEL transformants (H1 and H3 and faint signal for H2), but not with the WT and the GLSII expression strains. These results, together with the results obtained for the extracellular expression, clearly indicated that the heterologous glucosidase II was retained within the ER when HDEL tagged.

Since this way of preparing intracellular proteins resulted in a rather big background signal, a new method was used. The WT strain (PPY12-OH) and two transformed strains 18 (GLSII) and H3 (GLSIIHDEL) were grown in 200 ml BMGY for two days at 30 degrees celcius, until an OD600 of about 12 was reached. Cells were washed with 50 mM Tris.HCl pH 7.5 and resuspended in 10 ml of that same

buffer. Cells lysis was performed through French press (3 times). Cell debris was precipitated at 3000g for 10 minutes. The resulting supernatant was ultracentrifuged for 1 hr at 35,000 rpm in a Beckman SW41.1 rotor to precipitate the intracellular membranes. The membrane pellet was resuspended in membrane suspension buffer (50 mM Tris.HCl pH 7.4 – 150 mM NaCl – 1mM EDTA pH 8.0 – 25x complete protease inhibitor (Roche)). To this suspension Triton-X100 was added to 1% final volume. The suspension was incubated overnight at 4 degrees Celcius on a rotating wheel. Next, the membranes were precipitated through ultracentrifugation (1hr at 40,000 rpm in a Beckman SW50.1 rotor). The supernatant contains the overnight extracted membrane and luminal proteins. Protein concentration was measured and equal amounts (500 ug) of the three strains were loaded for SDS-PAGE analysis. After electrophoresis, proteins were transferred to a nitrocellulose membrane. Membranes were screened with antibodies as described earlier. This time, the intracellular presence of glucosidase II could not only be established in transformant H3 (in which a HDEL tagged GLSII ORF was transformed), but also in transformant 18 (the one in which a non-HDEL tagged GLSII ORF was transformed). This could be the result from the interaction of the heterologous glucosidase alpha subunit with the endogenous *Pichia* beta subunit. It was believed to be the result of the slow trafficking of the heterologous glucosidase II within the secretion pathway, due to folding problems. Since only low amounts of the protein were discovered in the secretion medium of the GLSII expression strains (i.e. strain 18), this certainly was the case.

4.5 Purification and activity assays of the recombinant glucosidase II alpha submit

Glucosidase II assay on extracellular medium:

A GLSII assay was set up as follows and was tested using a commercially available yeast alpha-glucosidase (Sigma) as a positive control.

Composition: 70 µl 80 mM phosphate-citrate buffer pH 6.8, 7 µl 250 mM mannose, 3.5 µl 250 mM 2-deoxy-D-glucose, 0.8 µl 4-MeUmbelliferyl-alpha-D-glucopyranoside (1 µM). Three assays were performed: one with 1 unit commercial enzyme, one without the enzyme and one with the enzyme but without the substrate. The assay mixture was incubated overnight at 30°C. When illuminated with UV, only the reaction mixture with both the enzyme and the substrate showed fluorescence (**Figure 22**). This indicates that the assay was very specific in detecting the activity of the alpha-glucosidase.

WT PPY12-OH, strain 18 and strain H3 were grown during 2 days in 2x10 ml growth medium. Cells were spun down and medium was adjusted to 300 mM NaCl and 10 mM imidazol and concentrated with Vivaspin columns to 0.5-1ml. Medium was loaded onto a Ni-NTA spin column (Qiagen) and the purification was performed according to the manufactures recommendations. Protein was eluted from the column in 2x100 µl elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazol pH 8.0). From each eluate, 20 µl was assayed for its glucosidase II activity. 0.33 units of the commercial enzyme diluted in 20 µl of the elution buffer was used as a positive control. Fluorescence was observed with the positive control and the elute of strain 18, the strain which secreted the enzyme into the growth medium. These results indicate that the recombinant *S. cerevisiae* GLSII alpha subunit, secreted by *Pichia pastoris*, was a functionally active enzyme. The activity was not seen in the WT (untransformed) strain, nor in strain H3 as the GLSII was expressed intracellularly (**Figure 23**). These results also indicate that the beta subunit is not necessary for the functionality of the alpha part of the protein.

Glucosidase II activity on intracellular total protein:

Membrane and luminal proteins from the WT strain, transformant 18 and transformant H3, were prepared by breaking the cells through a French press, as described in section 4.4 (intracellular expression). Equal amounts of total intracellular protein were

assayed through the GLSII assay as described above. Due to the substrate used, total intracellular alpha-glucosidase activity (glucosidase I and II, lysosomal alpha-glucosidases, etc.) was assayed, and not only the glucosidase II activity. However, when performing the assay on both WT and transformants 18 and H3, a twofold higher alpha-glucosidase activity was found in both transformed strains. Since these strains only differ from the WT strain in the expression of the *S. cerevisiae* glucosidase II protein (with or without HDEL), it is believed that this twofold increase in activity results from the activity of the recombinant protein. The whole procedure of protein preparation and glucosidase II assay was repeated a second time; again a twofold increase in total intracellular alpha-glucosidase activity was found for the two transformants, when compared to the WT untransformed strain.

Example 5

Creating *Pichia* Strains Expressing both Glucosidase II and Mannosidase

Strain GS115 was transformed with pGAPZGLSII and pGAPZglsIIHDEL. Transformants were selected on YPDSzeo.

Strain YGC4 was transformed with the following constructs, respectively:

(1) pGAPADEglsII and pGAPADEglsIIHDEL, selection on synthetic sorbitol medium without adenine;

(2) pGAPZMFManHDEL: selection on YPDSzeo; and

(3) pGAPZMFManHDEL/pGAPADEglsIIHDEL: selection on synthetic sorbitol medium without adenine and with zeocin.

Strain YGC4 with pBLURA5'PpOCH1 and expressing MFmannosidaseHDEL was transformed with pGAPADEglsII and pGAPADEglsIIHDEL. Selection of transformants was done on synthetic sorbitol medium without adenine and uracil.

For all transformations, colonies were obtained. Transformants with the expression vector(s) integrated into the genome, determined by PCR, were obtained. Expression of GLSII from some of these transformants was observed.

Example 6

Introduction of GlcNAc-Transferase I in *Pichia pastoris*

6.1 Vector Construction

The human GnTI gene was isolated and described by Kumar et al (*Proc Natl Acad Sci U S A.* 87(24):9948-52, 1990). The sequence of this gene is available at NCBI GenBank under Accession No. M61829, as set in forth by SEQ ID NO: 34. pUChGNTI (Maras et al. *Eur J Biochem.*249(3):701-7, 1997), which contains the full length human GnTI, was digested by SgrA I and Xba I and blunted by Klenow enzyme. The resulting vector contained the catalytic domain of GnTI. This fragment was then ligated with an insert fragment from plasmid Yep352Kre2 digested by Sac I and Pvu I, then made blunt by T4 DNA Polymerase. The insert fragment encodes a peptide derived from the N-terminal part of *S. cerevisiae* Kre2 which is responsible for the localization of expressed proteins in Golgi apparatus of *Pichia pastoris*. The resulting plasmid pUCKrecohGnTI was then cut by EcoR V and Hind III, blunted by Klenow enzyme. The fragment containing the Kre2 and hGnTI fusion was then inserted into vector pPIC6A which was digested by Sal I and EcoR I, blunted by Klenow enzyme. The final plasmid was named pPIC6AKrecoGnTI and the construction strategy is graphically depicted in **Figure 24**. pPIC6AKrecoGnTI (SEQ ID NO: 36, Open Reading Frame: 2699bp - 4027bp) contains the fusion construct KrecoGnTI under promoter AOX1 and blasticidin resistance marker.

6.2 Yeast transformation and genomic integration

All transformations to *Pichia pastoris* were performed with electroporation according to the methods in Invitrogen. The strain used was GS 115 containing α–1,2–mannosidase-HDEL and Trans-sialidase (TS). Plasmid pPIC6AKrecoGntI was

linearised by Nsi I digestion in the 5' AOX region.Transformants carrying blasticidin resistance gene were selected on YPD containing 500 ug/ml blasticidin and 1M sorbitol. Genomic intergration of expression cassettes was verified using PCR with genomic DNA purified from *Pichia* strains according the Yeast Miniprep method (Nucleon). The 5' (SEQ ID NO: 11) and 3' AOX1 (SEQ ID NO: 9) primers were used to verify the presence of the construct.

The results indicated that the correct insert of 1620 bp was obtained from all 4 blasticidin resistant colonies. The control colony (a wild type or an untransformed strain) did not have this PCR fragment.

6.3 Glycan analyses

The transformants and the untransformed organism (WT: GS115 containing TS and ManHDEL) were grown overnight in BMGY. The AOX1 promoter was induced during 24 h with 1% MetOH. Extra cellular medium was collected and dialysed against PBS. Oligosaccharides were isolated and separated by DSA-FACE. The results are shown in **Figure 25**. A peak where $GlcNAcMan_5GlcNAc_2$ was observed (indicated by an arrow in **Figure 25**). This peak was absent in the WT strain. Additionally an increase of the peak indicated with an asterix was also observed. Both, the additional peak and the increased peak returned to the WT situation after digestion with Jack Bean β-*N*-acetylhexosaminidase, a glycosidase lysing β-*N*-acetylhexosaminyl residues (presented in Figure 26).

In order to evaluate the transfer of GlcNAc to glycans of a heterologous protein or the percentage of $Man_5GlcNAc_2$ that was converted to $GlcNAcMan_5GlcNAc_2$, Transsialidase (TS) was purified. Yeast strains were cultured in BMMY during 30 h. Medium was collected and loaded on an anti E-tag column. Glycans were then analysed.

Transsialidase (TS) produced in three different strains (*i.e.* GS115, GS115 ManHDEL and GS115 ManHDEL Kre2GnTI) was purified by affinity chromatography (anti E-tag column). The N-glycans were isolated by digestion with PNGase F and analyzed by DSA-FACE. As shown in **Figure 27** shows that, in glycans from TS produced in GS115 containing MannosidaseHDEL and Kre2-GlcNAc-transferase, almost all $Man_5GlcNAc_2$ was converted to GlcNAcMan5GlcNAc2.

Alternatively TS from the three different strains was separated by SDS-PAGE and blotted on a nitrocellulose membrane. First the proteins were visualized by Ponceau red staining to show that equal amounts of protein were loaded. Then they were screened by biotin conjugated lectin GSII, a lectin specific against terminal GlcNAc. The lectin was visualized by POD conjugated streptavidin and a luminescent substrate. The two blots are presented in **Figure 28**, showing that only the band in lane 3 was stained by the lectin. Lane 3 represents TS produced from strain GS115 containing ManHDEL Kre2GnTI. These results indicate that terminal GlcNAc is present only in strains with ManHDEL Kre2GnTI.

An additional proof of the presence of $GlcNAcMan_5GlcNAc_2$ was obtained from digestions with exoglycosidases on the N-glycans of TS produced in the strain GS115 ManHDEL Kre2GnTI. TS was purified, and its N-glycans were isolated and analyzed by DSA-FACE. Digestions with Jack Bean β-N-acetylhexosaminidase were performed. The result is shown in **Figure 29**.

As can be seen in **Figure 29**, panel 2, almost no $Man_5GlcNAc_2$ was present in undigested N-glycans from TS produced in strain GS115 ManHDEL Kre2GnTI. Both the $Man_5GlcNAc_2$ and the $Man_6GlcNAc_2$ peaks were increased if the same glycans as in panel 2 were digested with β-N-acetylhexosaminidase, as seen in panel 3. This indicates the presence of GlcNAc on both structures. Panel 1 is oligomaltose reference, while panel 4 is the result of N-glycans from RNaseB.

Example 7
Mannosidase II expression in Pichia

7.1 Vector construction

Mammalian mannosidase II is located in the same Golgi region as is GnTI. Therefore the N-terminal part of mannosidase II that is responsible for the localization of the protein in the Golgi was replaced by that of ScKre2, as was done to localize GnTI in the Golgi apparatus of *Pichia pastoris*. A construction strategy was set up that ultimately led to a plasmid containing Kre2ManII ORF under the control of the gap promoter with the ARG4 marker for selection in *Pichia pastoris*. The plasmid containing the human mannosidase II gene was pCDNA3ManII. The sequence of this gene is available at EMBL under Accession No. U31520 (SEQ ID NO: 31). Plasmid pBLARGIX was as pulished in Cereghino et al. *Gene* 263(1-2):159-69, 2001.

The mannosidase cDNA fragment was puried by PCR. The primers used were the sense primer: 5'CGCGGATCCATGGCCAAAAAAGTTCACAAAGCAATTTC3' (SEQ ID NO: 32) and antisense primer: 5'GTGTCCCATTCATTAGATTC3' (SEQ ID NO: 33). The PCR fragment was then digested by BamH I and PflM I and inserted into vector pCDNA3manII(2) opened with same enzymes. The product was named pCDNA3manII(3). pCDNA3manII(2) was made by digesting pCDNA3manII with Msc I and Stu I, followed by self ligation. A plasmid YEp352Kre2BglII was made through inserting a Bgl II linker into T4 DNA Polymerase blunted plasmid YEp352Kre2, which had been digested by Sac I. YEp352Kre2BglII was cut by Pvu I and blunted by T4 DNA Polymerase to generate the coding sequence for the N-terminal part of Kre2 and subsequently digested by Bgl II. The coding fragment of the N-terminal part of Kre2 was purified and ligated to the vector of pCDNA3manII(3) digested by Msc I and Bam HI. The resulting plasmid was named pCDNA3KreManII. The plasmid carrying a GAP promoter (pGAPZA) was cut by Nsi I and BamH I, made blunt by T4 DNA Polymerase. The fragment containing the GAP promoter was then

inserted into the vector pBLARG IX, which carried the selection marker ARG4, opened by Spe I and Afl II digestion, blunted by Klenow enzyme. The intermediate plasmid was named pGAPAArg. pGAPAArg was further digested by Sfi I, blunted by Klenow, followed by Xho I digestion. The resulting vector was then ligated with the insert made from plasmid pCDNA3KreManII digested by EcoR V and Xho I. The final plasmid was named pGAPKreManII. The cloning strategy is depicted in **Figure** 30A and 30B.

7.2 Yeast transformation and genomic integration

MannosidaseHDEL (ManHDEL) was introduced (transformation with pGAPZMFManHDEL) to strain YGC-4. Genomic characterization of the transformed cells was done by PCR using a Gap primer (SEQ ID NO: 10) and a 3'AOX primer (SEQ ID NOs: 9). A band of 2010 bp indicated the presence of the construct in *Pichia pastoris* transformants. N-glycans derived from secreted and cell wall glycoproteins were analyzed by means of DSA FACE. Similar to the results obtained in strain GS115, glycans were trimmed and the presence of $Man_5GlcNAc_2$ was observed. Cell wall glycoproteins (mannoproteins) were prepared in a miniaturized way as described by Peat et al. (*J. Chem. Soc.* 29, 1961).

From the resulting YGC-4 ManHDEL clones, Clone 1 (M1) was transformed with the plasmid pPIC6KrecoGnTI. Genomic DNA was prepared and analyzed by PCR for integration of the construct into the genome using 5'AOX and 3'AOX primers (SEQ ID NOs: 11 and 9). Analysis of mannoproteins revealed the presence of $GlcNAcMan_5GlcNAc_2$ compearable to the situation in the GS115 strain.

Clone 4 (YGC-4 M1 GnTI4) was transformed with plasmid pGAPKreManII by electroporation after linerisation of the plasmid in the 3' AOX I region by a digestion with Pvu II. Transformants having the ARG4 gene were selected on medium containing no arginin. DNA was prepeared and analysed for the presence of the

construct by PCR using primer SEQ ID NO: 33 and SEQ ID NO: 10. An amplified fragment of 617 bp was seen indicating the presence of the construct.

Example 8
Mannosidase-HDEL expression and OCH1 inactivation

8.1 Vector construction

The following experiments were performed to generate a plasmid that would carry both a mannosidase-HDEL expression unit and an inactivated OCH1 fragment. A point mutation in the very 5' part of the OCH1 CDS was introduced, changing the codon for amino acid 12 to a stop codon, using QuickChange® kit from Stratagene, so that mRNA produced by a potential cryptic promoter activity would translate into an inactive protein. The template used was pUC18PpOCH1. The presence of the stop codon was confirmed by sequencing.

Cloning in pPICZB to pick up the AOX1 transcription terminator was done by ligating the blunted HindIII fragment of the mutated pUC18PpOCH1 variant in the blunted, XhoI/EcoRI opened pPICZB. The resulting plasmid was called pPICZB5'PpOCH1Mut. The Bst BI site in the GAP promoter of pGAPZMFManHDEL was removed to allow linearization of the final vectors in the 5'PpOCH1 sequence for integration. Hereto, pGAPZMFManHDEL was digested with Bst BI and blunted with T4 polymerase and dXTPs. Subsequently, the linearized and blunted plasmid was purified and digested with BamHI. The vector and the released fragment (containing the ManHDEL open reading frame and the 3'AOXI transcription terminator were separated by agarose gel electrophoresis and the DNA fragments were eluted from the gel. Fragment and vector were religated in a 10:1 ratio using T4 DNA ligase, overnight at 16 °C.

The mutated 5'PpOCH1-3'AOX1 fragment was released by PstI/BamHI digestion, blunted and ligated in pGAPZMFManHDEL-Bst BI that had been opened with BglII and blunted. The resulting plasmid was called

pZMFManHDEL5'PpOCH1Mut (SEQ ID NO:46). The map of this plasmid is shown on **Figure 33**.

CLAIMS

1. A vector expressing an α-1,2-mannosidase or a functional part thereof in a methylotrophic yeast, wherein said vector comprises a nucleotide sequence coding for said α-1,2-mannosidase or said functional part thereof, and wherein said vector further comprises a nucleotide sequence which inhibits the expression of the functional OCH1 protein in said methylotrophic yeast.

2. The vector of claim 1, wherein said methylotrophic yeast is *Pichia*.

3. The vector of claim 2, wherein said *Pichia* is *Pichia pastoris*.

4. The vector of claim 1, wherein said α-1,2-mannosidase or said functional part thereof is a protein from a fungal species.

5. The vector of claim 4, wherein said fungus is *Trichoderma reesei*.

6. The vector of claim 1, wherein said α-1,2-mannosidase or said functional part thereof is a protein from a mammalian species.

7. The vector of claim 6, wherein said α-1,2-mannosidase or said functional part thereof is murine α-1,2-mannosidase IA or IB.

8. The vector of claim 1, wherein said α-1,2-mannosidase or said functional part thereof is tagged with an ER-retention signal.

9. The vector of claim 8, wherein said ER-retention signal comprises peptide HDEL.

10. The vector of claim 1, wherein said OCH1 is a protein from a fungal species.

11. The vector of claim 10, wherein said fungus is *Pichia pastoris.*

12. The vector of claim 1, wherein the nucleotide coding sequences are operably linked to a promoter and a 3' termination sequence.

13. The vector of claim 12, wherein said promoter is the promoter of a gene selected from the group consisting of AOXI, AOXII, GAP, YPT1, and FLD.

14. A vector designated as pZMFManHDEL5'PpOCH1Mut.

15. A vector expressing a GlcNAc-transferase I or a functional part thereof in a methylotrophic yeast strain, comprising a nucleotide sequence coding for said GlcNAc-transferase I or said functional part.

16. The vector of claim 15, where in the said methylotrophic strain is *Pichia.*

17. The vector of claim 16, where in the said *Pichia* is *Pichia pastoris.*

18. The vector of claim 15, wherein said GlcNAc-transferase I is a protein from a mammalian species.

19. The vector of claim 15, wherein the said GlcNAc-transferase I is a protein from a group of species consisting of plants, insects, nematodes and protozoa.

20. The vector of claim 19, wherein the said protozoa is *Leishmania tarentolae*.

21. The vector of claim 15, wherein the said GlcNAc-transferase I is a protein obtained through protein engineering experiments.

22. The vector of claim 18, wherein said mammalian species is human.

23. The vector of claim 15, wherein said GlcNAc-transferase I or said functional part is operably linked to a Golgi localization signal.

24. The vector of claim 23, wherein said Golgi localization signal comprises a peptide derived from the N-terminal part of *Saccharomyces cerevisiae* Kre2.

25. The vector of claim 15, wherein the nucleotide sequence coding for said GlcNAc-transferase I or said functional part is operably linked to a promoter and a 3' termination sequence.

26. The vector of claim 25, wherein said promoter is the promoter of a gene selected from the group consisting of AOXI, AOXII, GAP, YPT1, and FLD.

27. The vector of claim 15, further comprising a nucleotide sequence coding for an α-1,2-mannosidase or a functional part thereof.

28. A vector having the designation of pPIC6AKrecoGnTI.

29. A vector expressing a mannosidase II or a functional part thereof in a methylotrophic yeast, comprising a nucleotide sequence coding for said mannosidase II or said functional part.

30. The vector of claim 29, wherein said mannosidase II is a protein from a mammalian species.

31. The vector of claim 30, wherein said mammalian species is human.

32. The vector of claim 29, wherein said mannosidase II or said functional part is operably linked to a Golgi localization signal.

33. The vector of claim 32, wherein said Golgi localization signal comprises a peptide derived from the N-terminal part of of *Saccharomyces cerevisiae* Kre2.

34. The vector of claim 29, wherein the nucleotide sequence coding for said mannosidase II or said functional part is operably linked to a promoter and a 3' termination sequence.

35. The vector of claim 34, wherein said promoter is the promoter of a gene selected from the group consisting of AOXI, AOXII, GAP, YPT1, and FLD.

36. The vector of claim 29 further comprising a nucleotide sequence coding for a GlcNAc-transferase I.

37. The vector of claim 36 further comprising a nucleotide sequence coding for α-1,2-mannosidase.

38. A vector having the designation pGAPKreManII.

39. A vector comprising at least one of (a) an α-1,2-mannosidase expression unit functional in methylotrophic yeasts; (b) a glucosidase II expression unit functional in methylotrophic yeasts; (c) a GlcNAc-transferase I expression unit functional in methylotrophic yeasts; (d) a mannosidase II expression unit functional in methylotrophic yeasts; or (e) a nucleotide sequence inhibiting the expression of the functional OCH1 protein in methylotrophic yeasts.

40. A method of reducing or modifying the glycosylation on proteins produced from a methylotrophic yeast, comprising transforming said yeast with any one of the vectors of claims 1, 14, 15, 28, 29, 38 or 39.

41. The method of claim 40, wherein said yeast is *Pichia pastoris*.

42. The method of claim 40, wherein said yeast is a *Pichia pastoris* strain selected from GS115 (NRRL Y-15851), GS190 (NRRL Y-18014), PPF1 (NRRL Y-18017), PPY12-OH, YGC4, or derivatives thereof.

43. The method of claim 40, wherein said yeast expresses a heterologous protein.

44. A strain of a methylotrophic yeast, wherein said strain is transformed with at least one of the vectors of claims 1, 14, 15, 28, 29, 38 or 39.

45. A methylotrophic yeast transformed with a vector which comprises GlcNAc-transferase I and a vector which comprises an α-1,2-mannosidase.

46. A methylotrophic yeast transformed with a vector which comprises GlcNAc-transferase I -α-1,2-mannosidase.

47. A method of reducing or modifying the glycosylation of a heterologous glycoprotein expressed from a methylotrophic yeast, comprising transforming cells of said methylotrophic yeast with at least one of the vectors of claims 1, 14, 15, 28, 29, 38 or 39, and producing said glycoprotein from the transformed cells.

48. A method of producing a glycoprotein with reduced or modified glycosylation in a methylotrophic yeast, comprising transforming said methylotrophic yeast with at least one of the vectors of claims 1, 14, 15, 28, 29, 38 or 39 and with a nucleotide sequence expressing said glycoprotein in said yeast, and producing said glycoprotein from the transformed cells.

49. A glycoprotein produced by the method of claim 48.

50. The glycoprotein of claim 49, wherein said glycoprotein has reduced immunogenicity relative to the glycoprotein produced from a wild type strain of said methylotrophic yeast.

51. The glycoprotein of claim 49, wherein said glycoprotein is part of pharmaceutical compositions.

52. A kit comprising any of the vectors of claims 1, 14, 15, 28, 29, 38 or 39.

53. The kit of claim 52, further comprising a methylotrophic yeast.

54. The kit comprising the methylotrophic yeast of claim 45 or 46.

55. The kit of claim 53 or 54, wherein the said methylotrophic yeast is *Pichia*.

56. The kit of claim 55, wherein the said *Pichia* is *Pichia pastoris*.

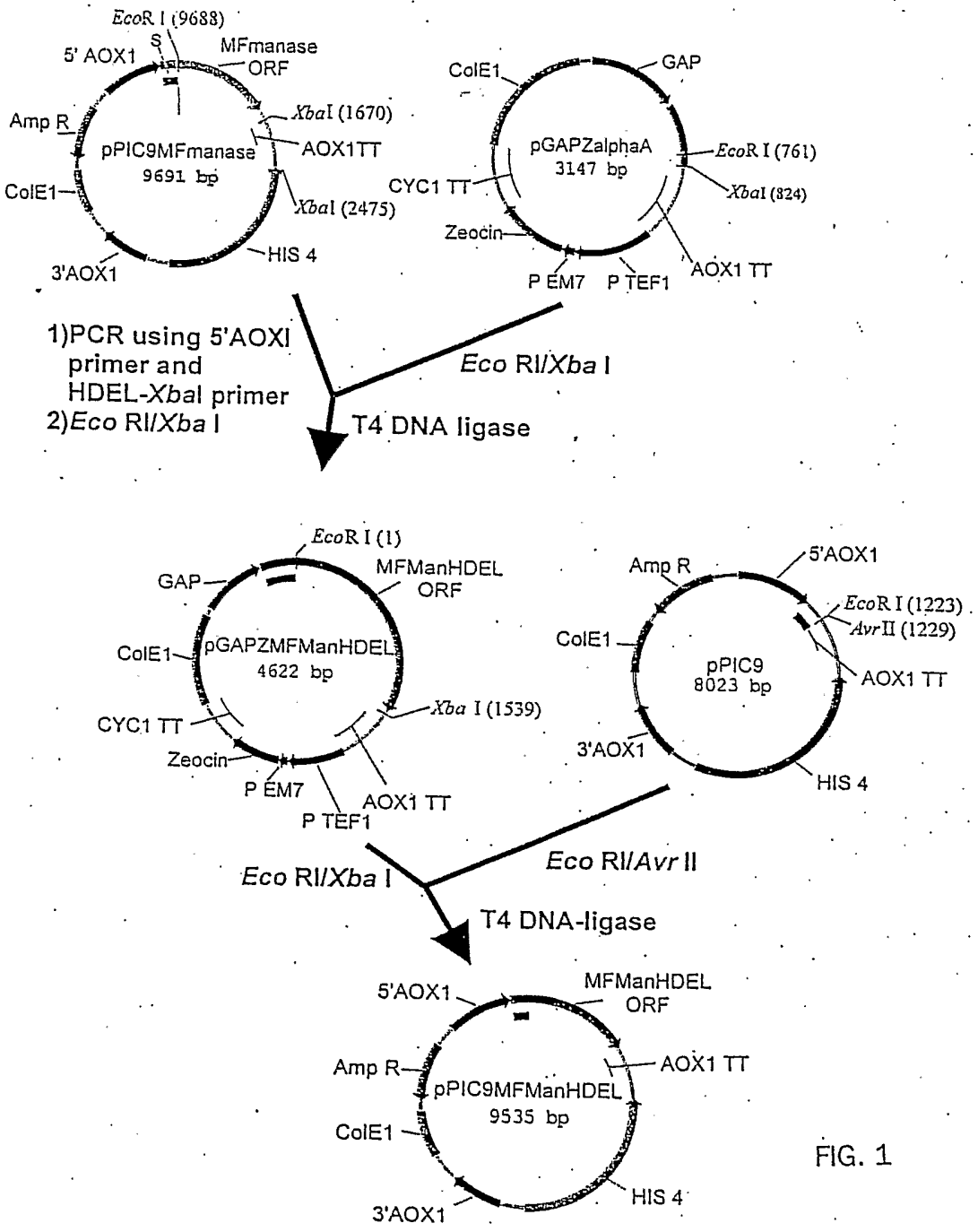

EcoR I (9688)
S
5' AOX1
MFmanase
ORF
Amp R
XbaI (1670)
pPIC9MFmanase
9691 bp
AOX1TT
ColE1
XbaI (2475)
3'AOX1
HIS 4
ColE1
GAP
pGAPZalphaA
3147 bp
EcoR I (761)
XbaI (824)
CYC1 TT
Zeocin
P EM7
P TEF1
AOX1 TT
1)PCR using 5'AOXI primer and HDEL-XbaI primer
2)Eco RI/Xba I
Eco RI/Xba I
T4 DNA ligase
EcoR I (1)
GAP
MFManHDEL
ORF
ColE1
pGAPZMFManHDEL
4622 bp
Xba I (1539)
CYC1 TT
Zeocin
P EM7
P TEF1
AOX1 TT
Amp R
5'AOX1
EcoR I (1223)
AvrII (1229)
ColE1
pPIC9
8023 bp
AOX1 TT
3'AOX1
HIS 4
Eco RI/Xba I
Eco RI/Avr II
T4 DNA-ligase
5'AOX1
MFManHDEL
ORF
Amp R
AOX1 TT
pPIC9MFManHDEL
9535 bp
ColE1
3'AOX1
HIS 4

FIG. 1

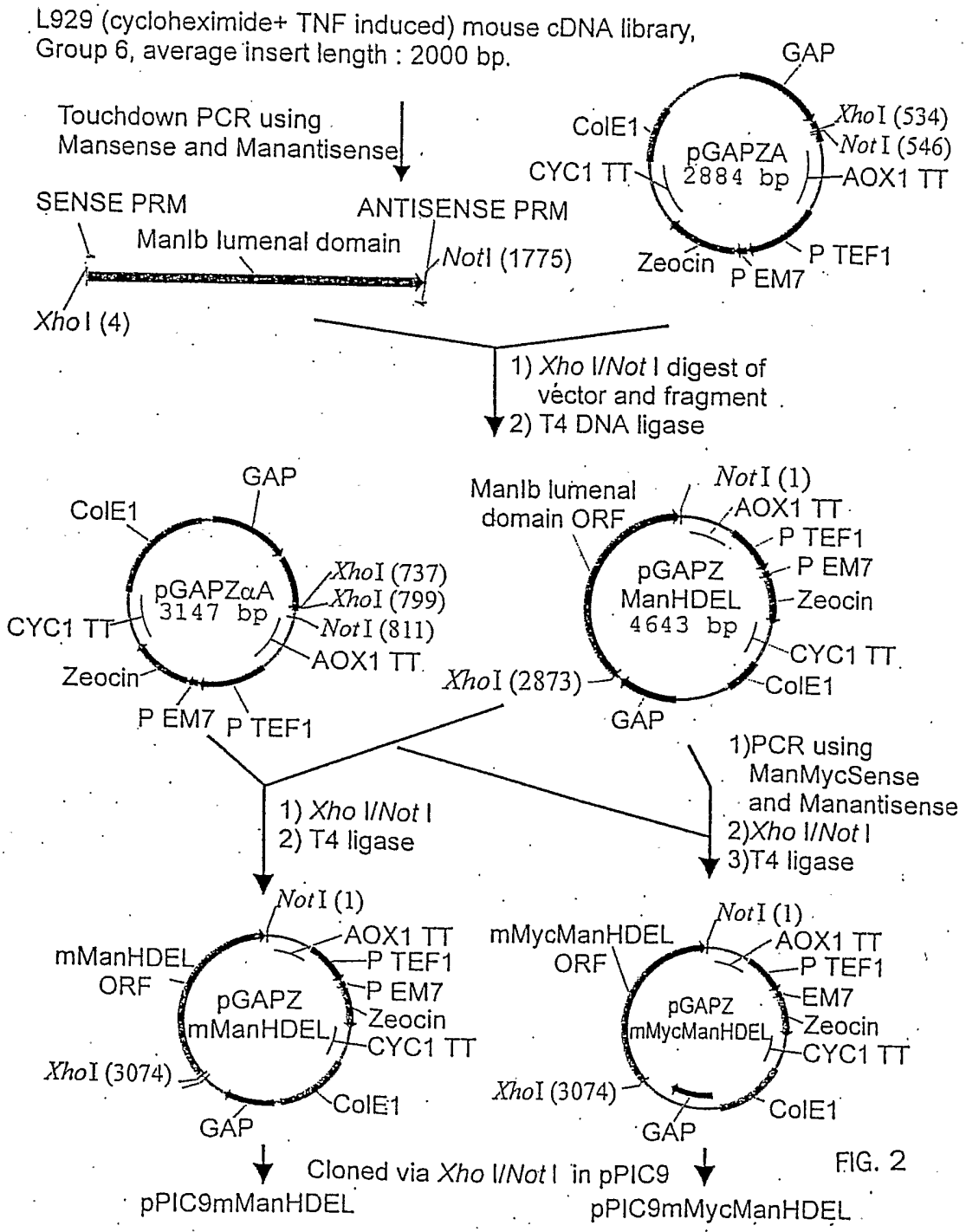

L929 (cycloheximide+ TNF induced) mouse cDNA library,
Group 6, average insert length : 2000 bp.
Touchdown PCR using
Mansense and Manantisense
SENSE PRM
ANTISENSE PRM
ManIb lumenal domain
NotI (1775)
XhoI (4)
GAP
ColE1
XhoI (534)
NotI (546)
pGAPZA
2884 bp
CYC1 TT
AOX1 TT
Zeocin
P TEF1
P EM7
1) Xho I/Not I digest of
vector and fragment
2) T4 DNA ligase
GAP
ColE1
pGAPZαA
3147 bp
XhoI (737)
XhoI (799)
NotI (811)
CYC1 TT
AOX1 TT
Zeocin
P EM7
P TEF1
ManIb lumenal
domain ORF
NotI (1)
AOX1 TT
P TEF1
P EM7
pGAPZ
ManHDEL
4643 bp
Zeocin
CYC1 TT
ColE1
XhoI (2873)
GAP
1) Xho I/Not I
2) T4 ligase
1)PCR using
ManMycSense
and Manantisense
2)Xho I/Not I
3)T4 ligase
NotI (1)
AOX1 TT
P TEF1
P EM7
Zeocin
CYC1 TT
mManHDEL
ORF
pGAPZ
mManHDEL
XhoI (3074)
ColE1
GAP
mMycManHDEL
ORF
NotI (1)
AOX1 TT
P TEF1
EM7
Zeocin
CYC1 TT
pGAPZ
mMycManHDEL
XhoI (3074)
ColE1
GAP
Cloned via Xho I/Not I in pPIC9
pPIC9mManHDEL
pPIC9mMycManHDEL

FIG. 2

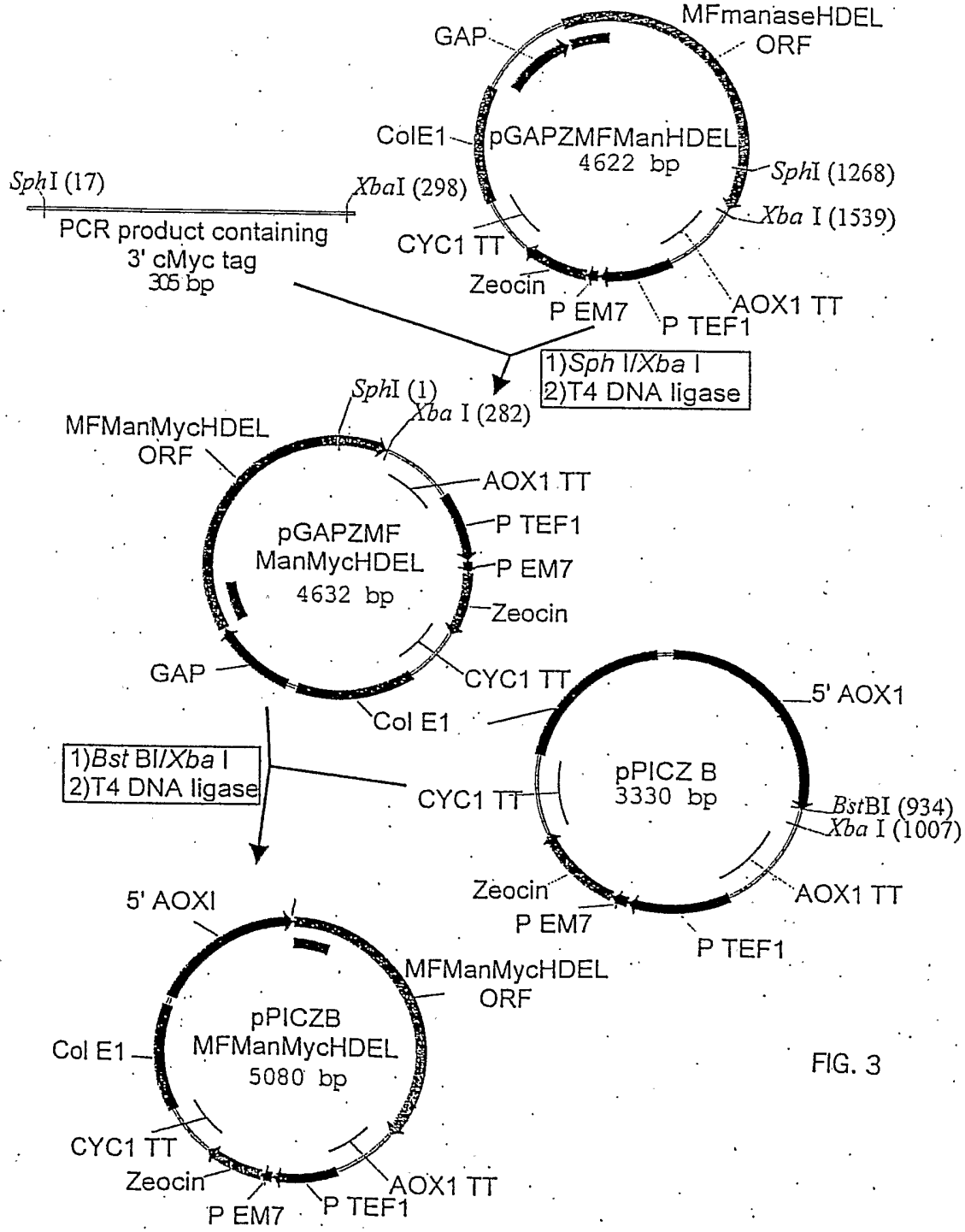

MFmanaseHDEL
ORF
GAP
ColE1
pGAPZMFManHDEL
4622 bp
SphI (1268)
Xba I (1539)
SphI (17)
XbaI (298)
PCR product containing
3' cMyc tag
305 bp
CYC1 TT
Zeocin
P EM7
P TEF1
AOX1 TT
1)Sph I/Xba I
2)T4 DNA ligase
SphI (1)
Xba I (282)
MFManMycHDEL
ORF
AOX1 TT
P TEF1
pGAPZMF
ManMycHDEL
4632 bp
P EM7
Zeocin
GAP
CYC1 TT
Col E1
1)Bst BI/Xba I
2)T4 DNA ligase
CYC1 TT
5' AOX1
pPICZ B
3330 bp
BstBI (934)
Xba I (1007)
Zeocin
P EM7
AOX1 TT
P TEF1
5' AOXI
MFManMycHDEL
ORF
pPICZB
MFManMycHDEL
5080 bp
Col E1
CYC1 TT
Zeocin
P EM7
P TEF1
AOX1 TT

FIG. 3

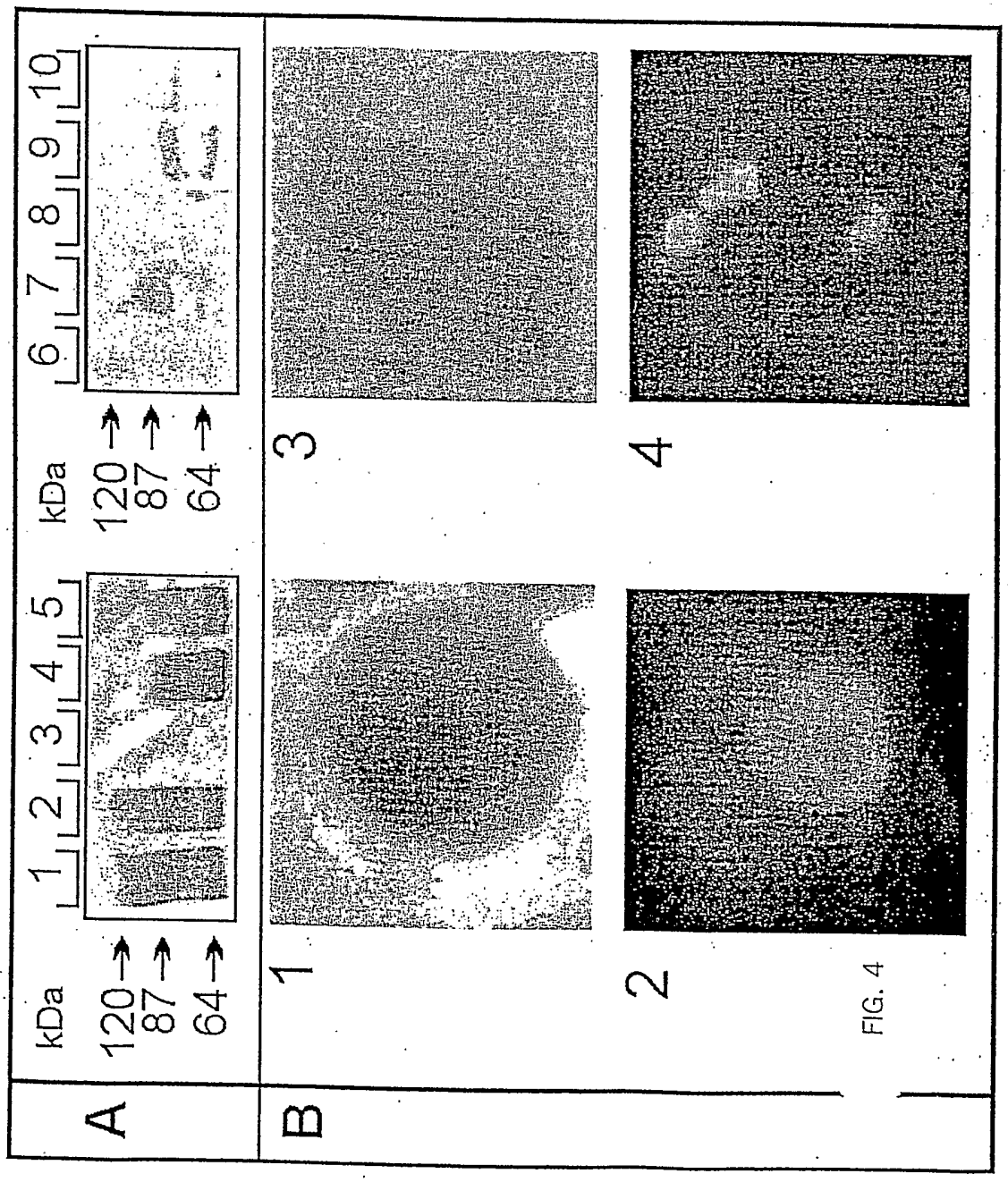


FIG. 4

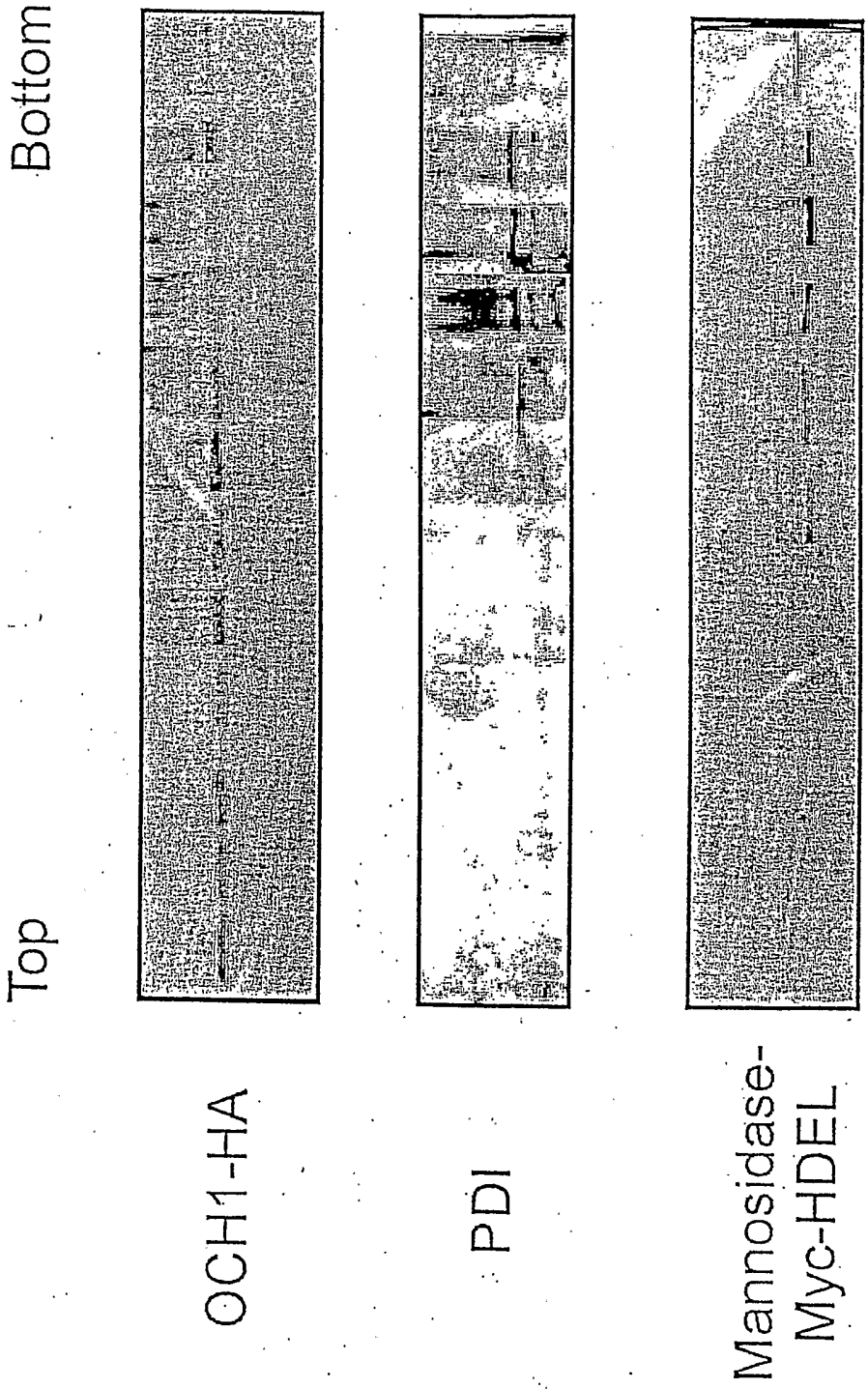


FIG. 5

FIG. 6

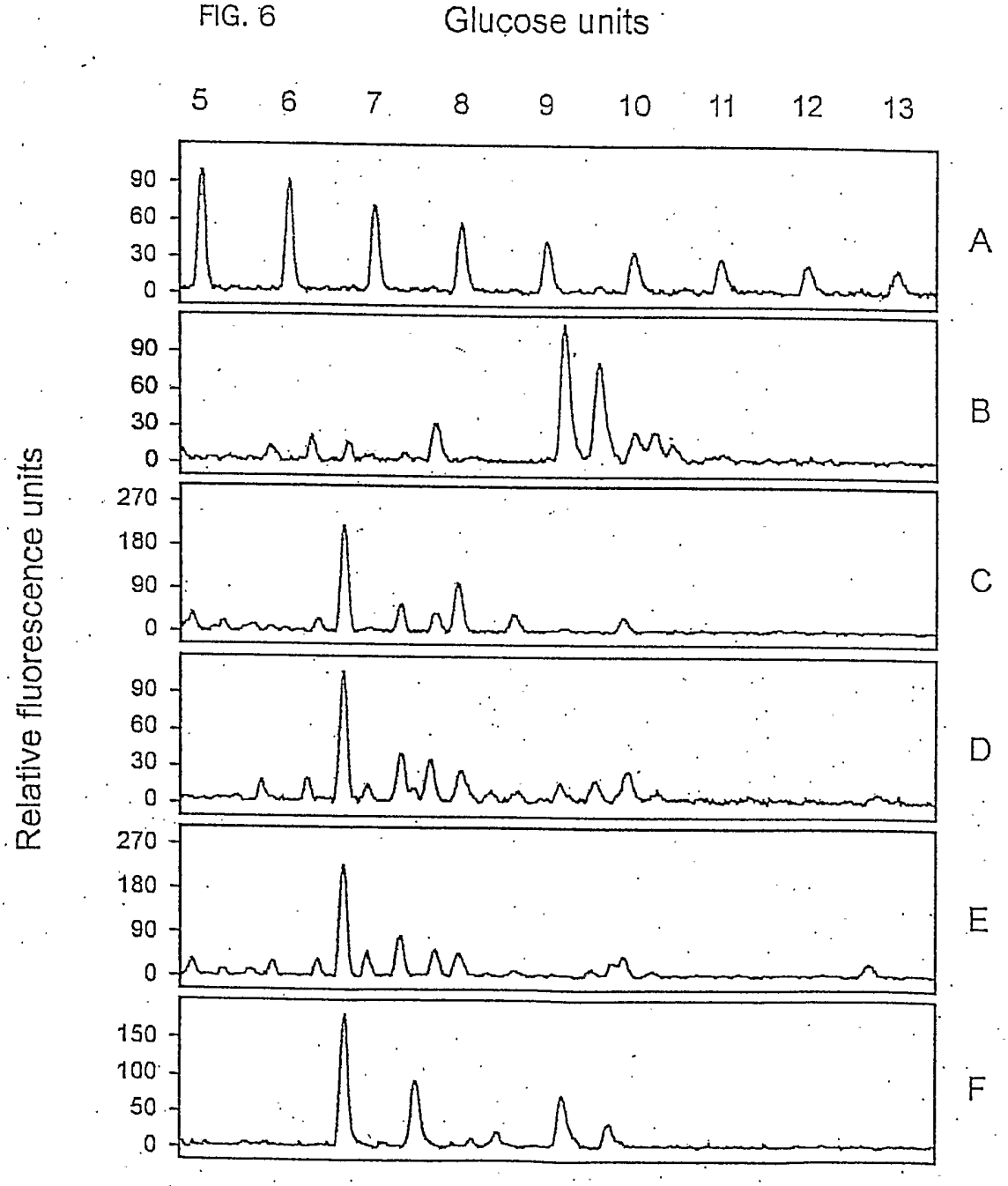
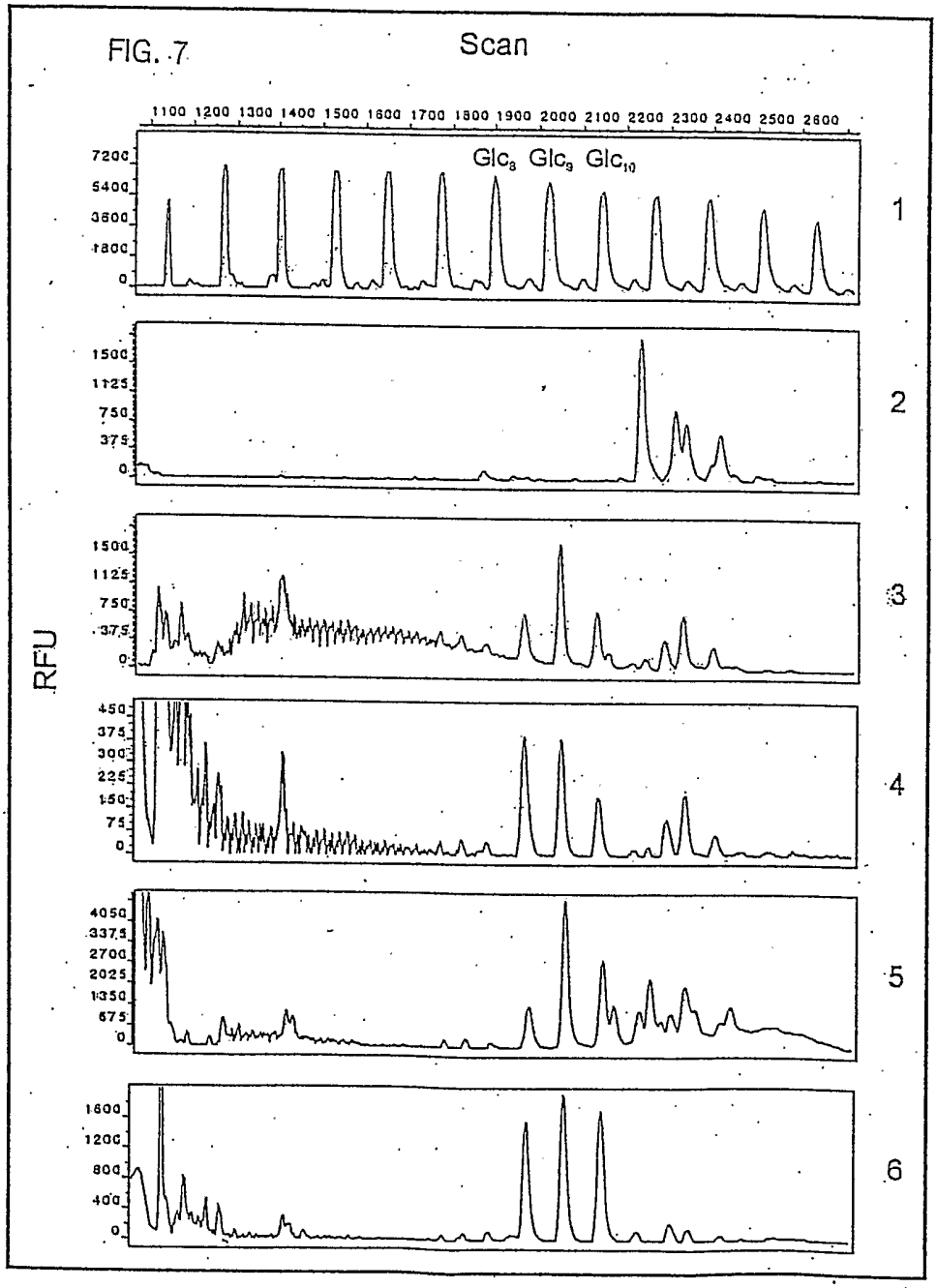

Glucose units
5
6
7
8
9
10
11
12
13
Relative fluorescence units
90
60
30
0
A
90
60
30
0
B
270
180
90
0
C
90
60
30
0
D
270
180
90
0
E
150
100
50
0
F

FIG. 7


Scan
1100 1200 1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 2400 2500 2600
RFU
Glc_8 Glc_9 Glc_{10}
7200
5400
3600
1800
0
1
1500
1125
750
375
0
2
1500
1125
750
375
0
3
450
375
300
225
150
75
0
4
4050
3375
2700
2025
1350
675
0
5
1600
1200
800
400
0
6

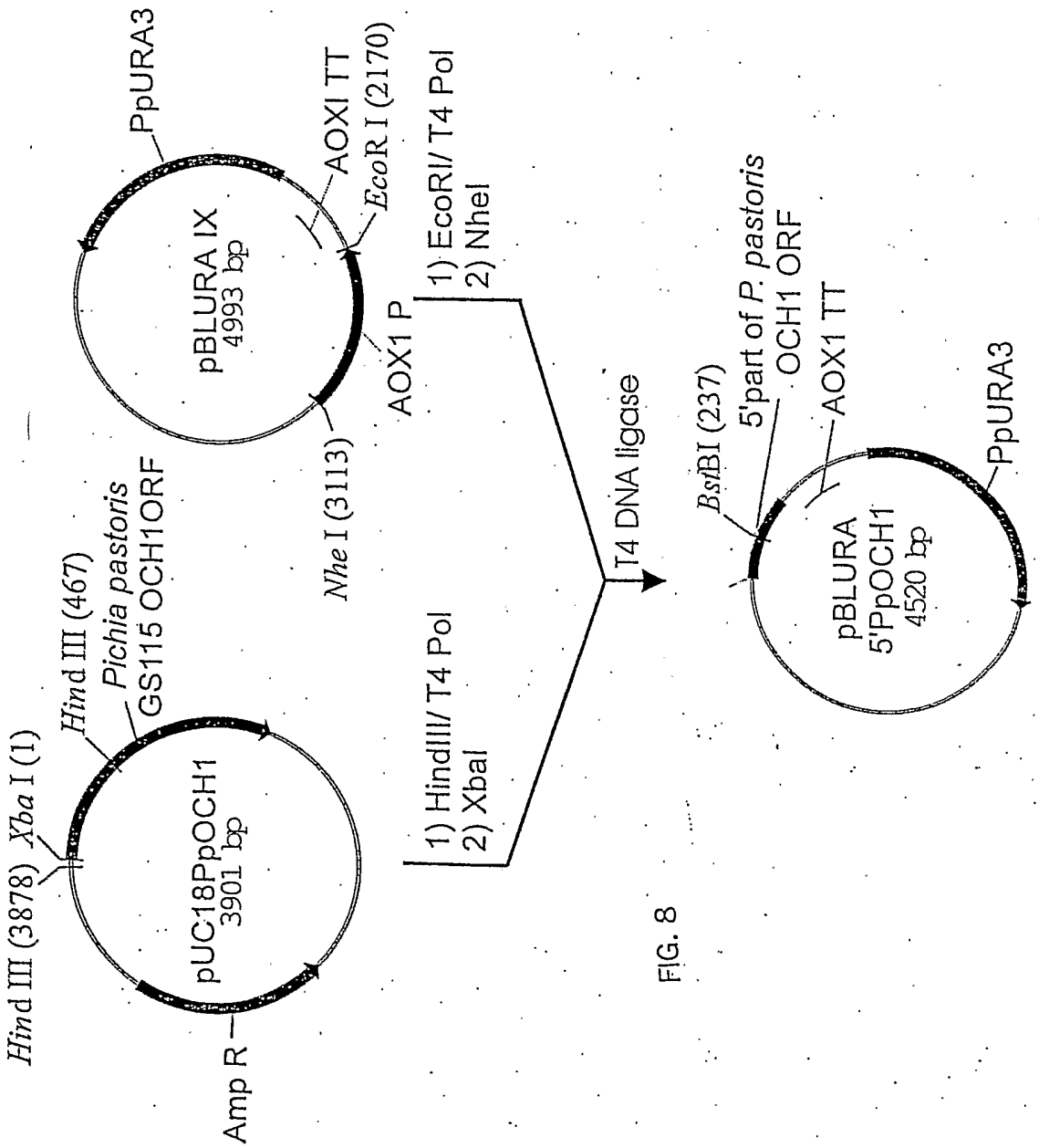

Hind III (3878) Xba I (1)
Hind III (467)
Pichia pastoris GS115 OCH1ORF
pUC18PpOCH1 3901 bp
Amp R
PpURA3
pBLURA IX 4993 bp
AOXI TT
EcoR I (2170)
Nhe I (3113)
AOX1 P
1) HindIII/ T4 Pol
2) XbaI
1) EcoRI/ T4 Pol
2) NheI
T4 DNA ligase
BstBI (237)
5'part of P. pastoris OCH1 ORF
AOX1 TT
PpURA3
pBLURA 5'PpOCH1 4520 bp

FIG. 8

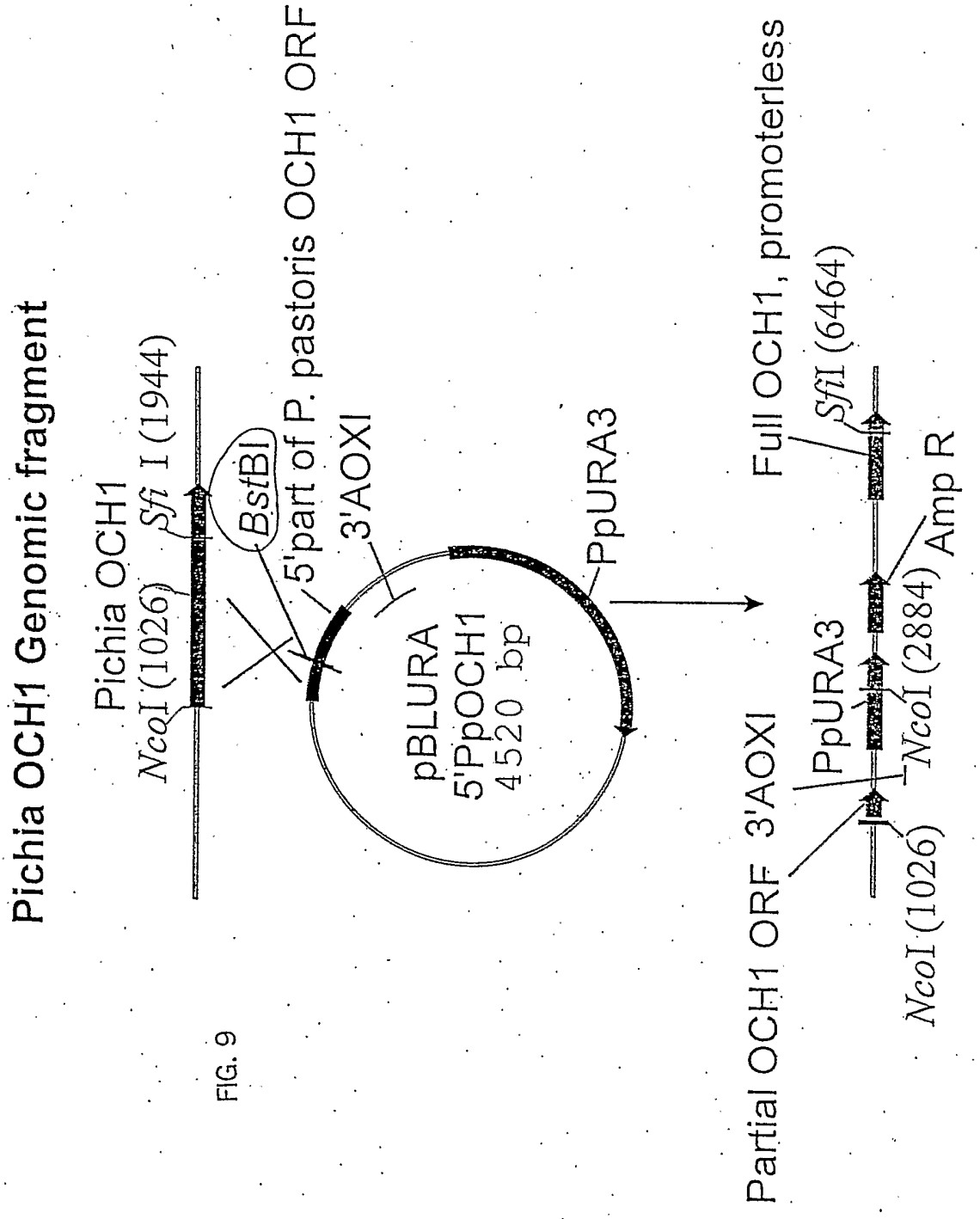

Pichia OCH1 Genomic fragment
Pichia OCH1
NcoI (1026)
Sfi I (1944)
BstBI
5'part of P. pastoris OCH1 ORF
3'AOXI
pBLURA
5'PpOCH1
4520 bp
PpURA3
Full OCH1, promoterless
SfiI (6464)
Amp R
Partial OCH1 ORF
3'AOXI
PpURA3
NcoI (2884)
NcoI (1026)

FIG. 9

FIG. 10

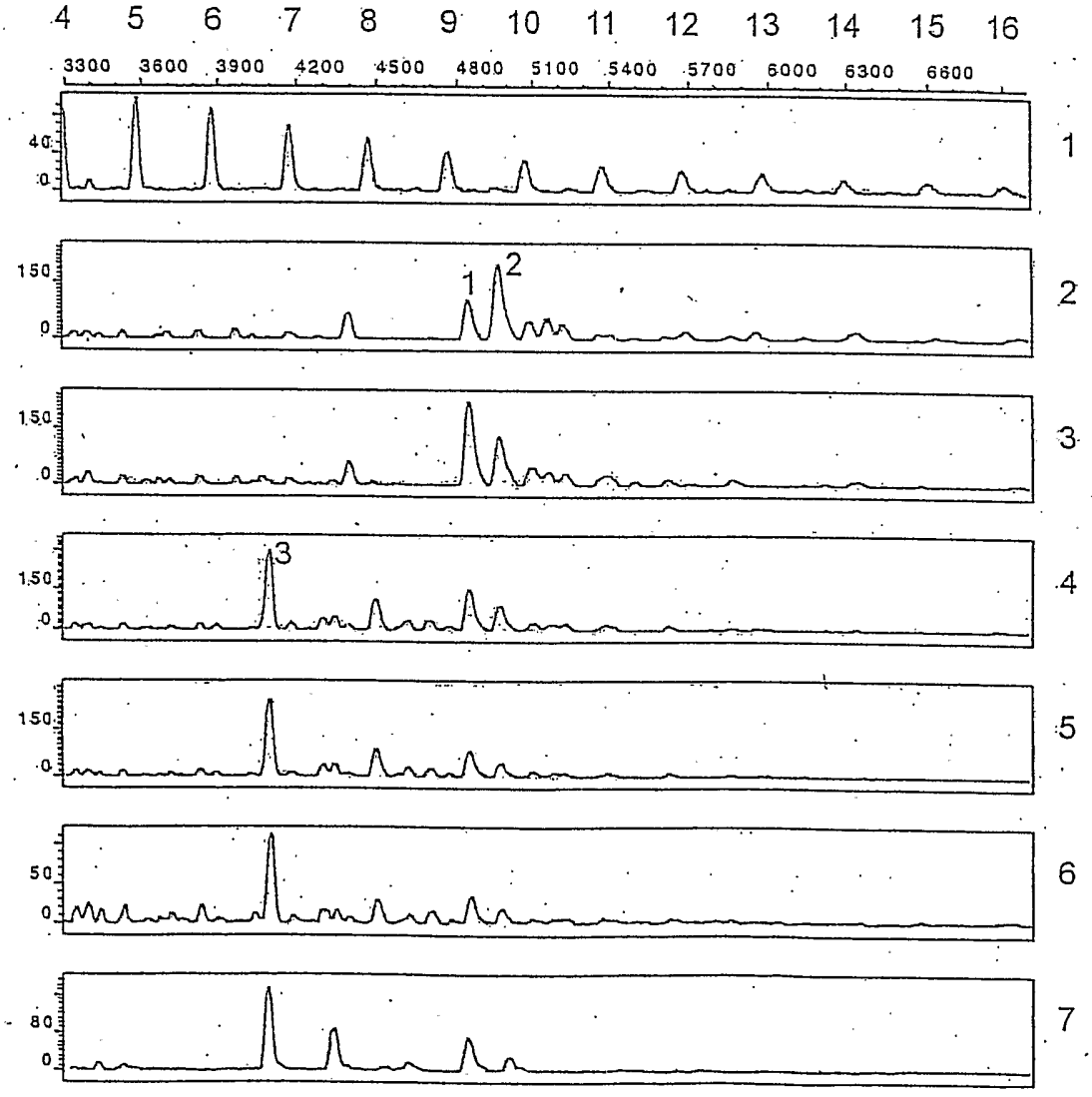
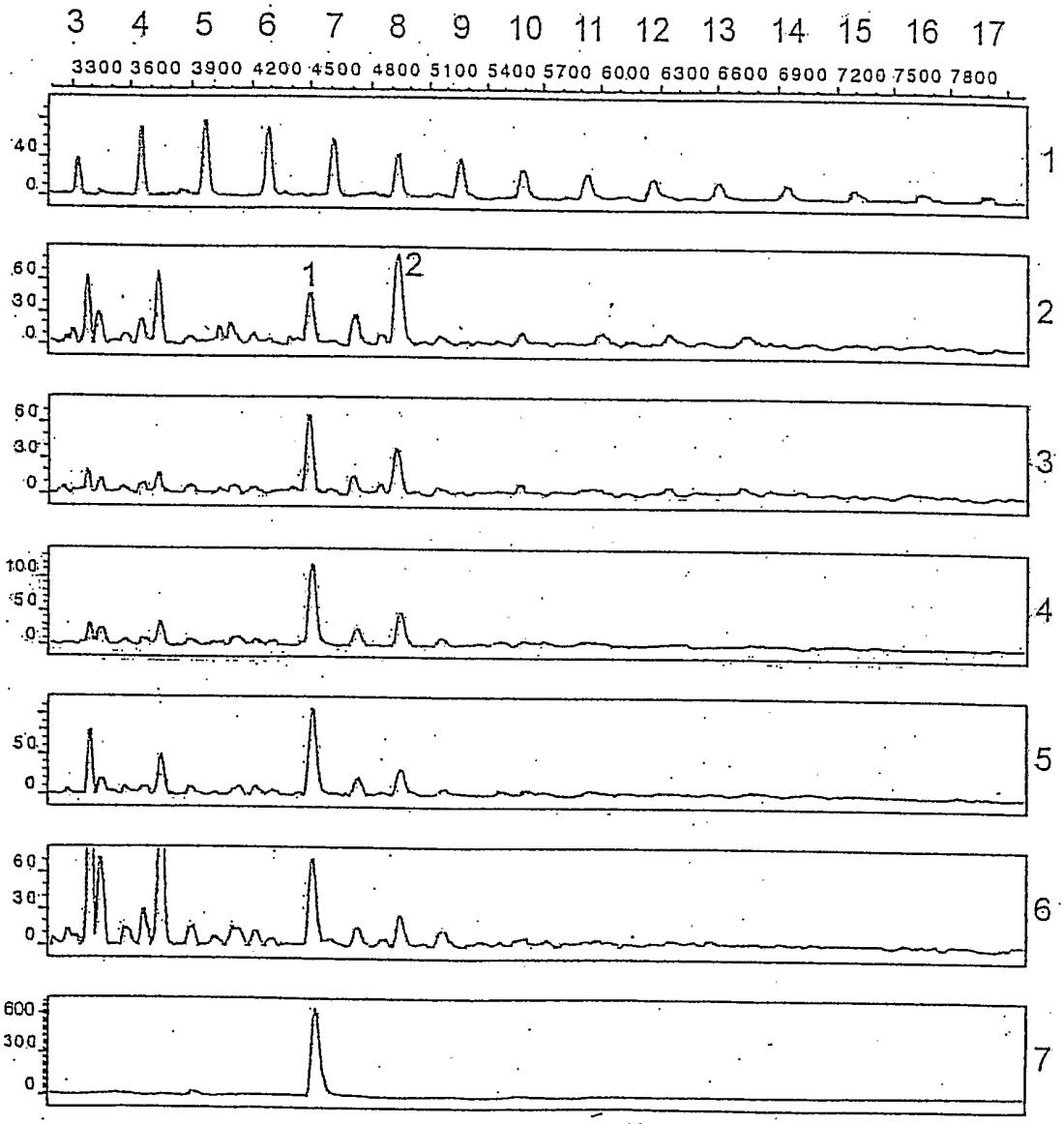

4 5 6 7 8 9 10 11 12 13 14 15 16
3300 3600 3900 4200 4500 4800 5100 5400 5700 6000 6300 6600
40
0
1
150
0
1 2
2
150
0
3
150
0
3
4
150
0
5
50
0
6
80
0
7

FIG. 11


3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
3300 3600 3900 4200 4500 4800 5100 5400 5700 6000 6300 6600 6900 7200 7500 7800
40
0
1
60
30
0
1
2
2
60
30
0
3
100
50
0
4
50
0
5
60
30
0
6
600
300
0
7

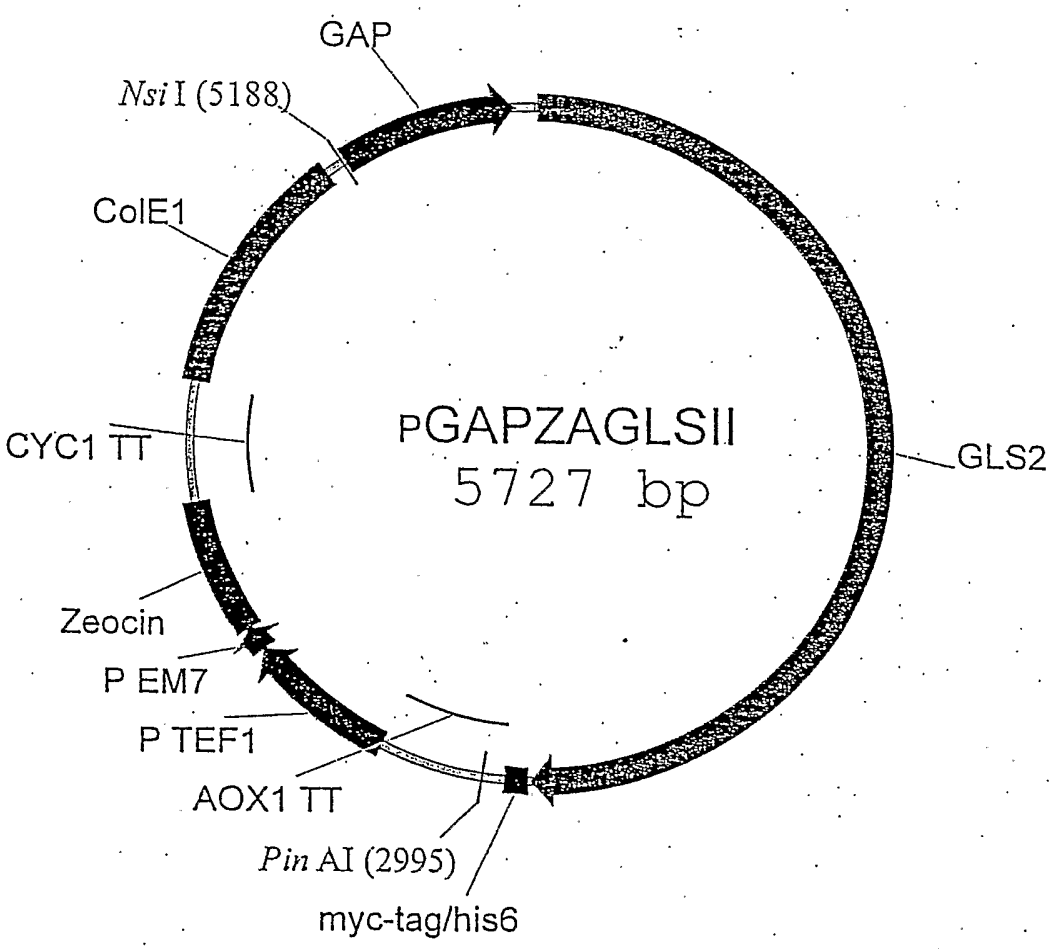

GAP
NsiI (5188)
ColE1
pGAPZAGLSII
5727 bp
CYC1 TT
GLS2
Zeocin
P EM7
P TEF1
AOX1 TT
PinAI (2995)
myc-tag/his6

FIG. 12

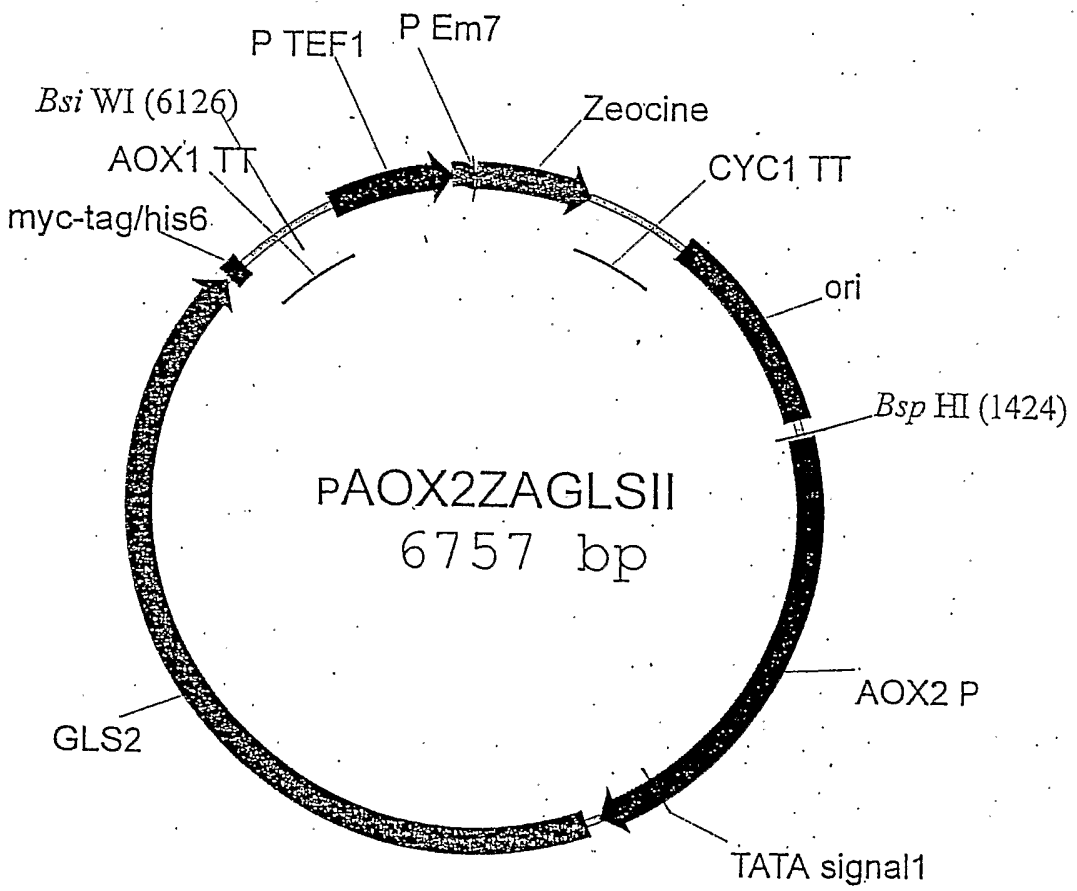

P TEF1
P Em7
Bsi WI (6126)
Zeocine
AOX1 TT
CYC1 TT
myc-tag/his6
ori
Bsp HI (1424)
pAOX2ZAGLSII
6757 bp
AOX2 P
GLS2
TATA signal1

FIG.13

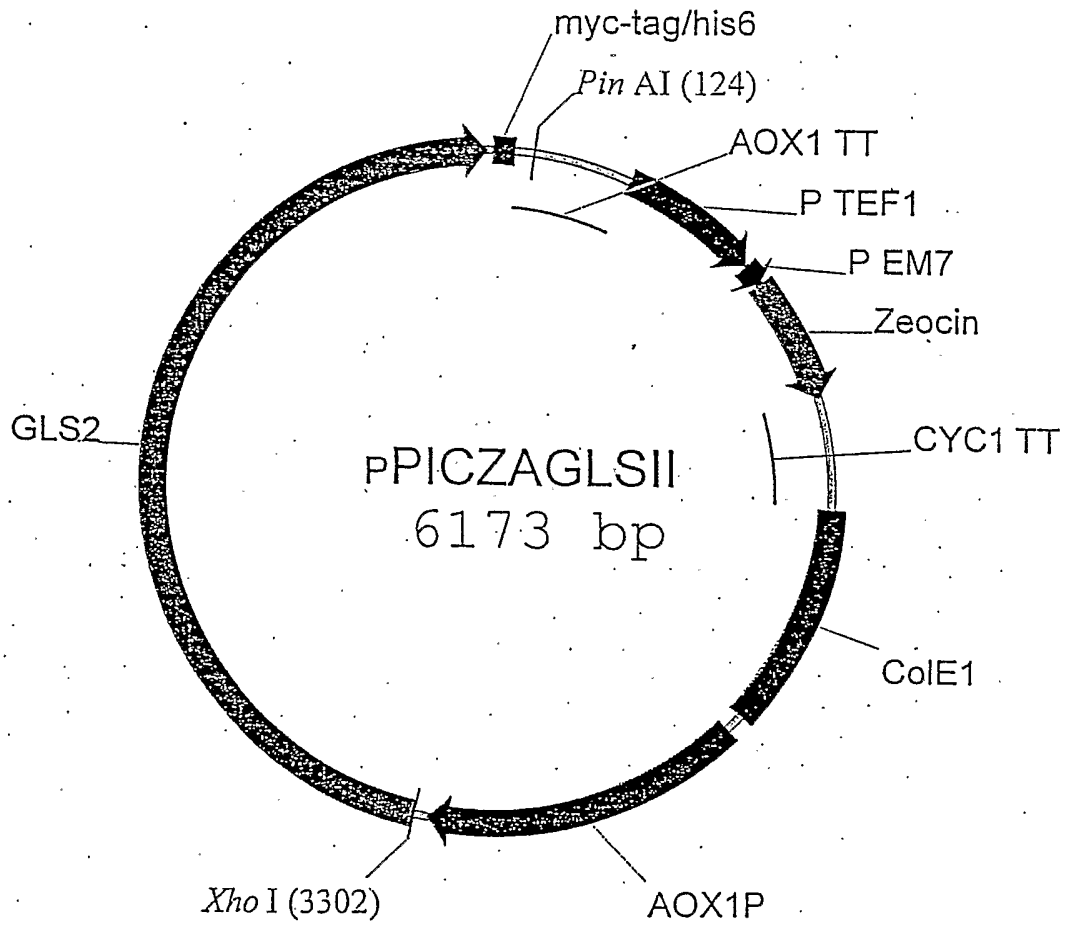

myc-tag/his6
Pin AI (124)
AOX1 TT
P TEF1
P EM7
Zeocin
CYC1 TT
ColE1
AOX1P
Xho I (3302)
GLS2
pPICZAGLSII
6173 bp

FIG. 14

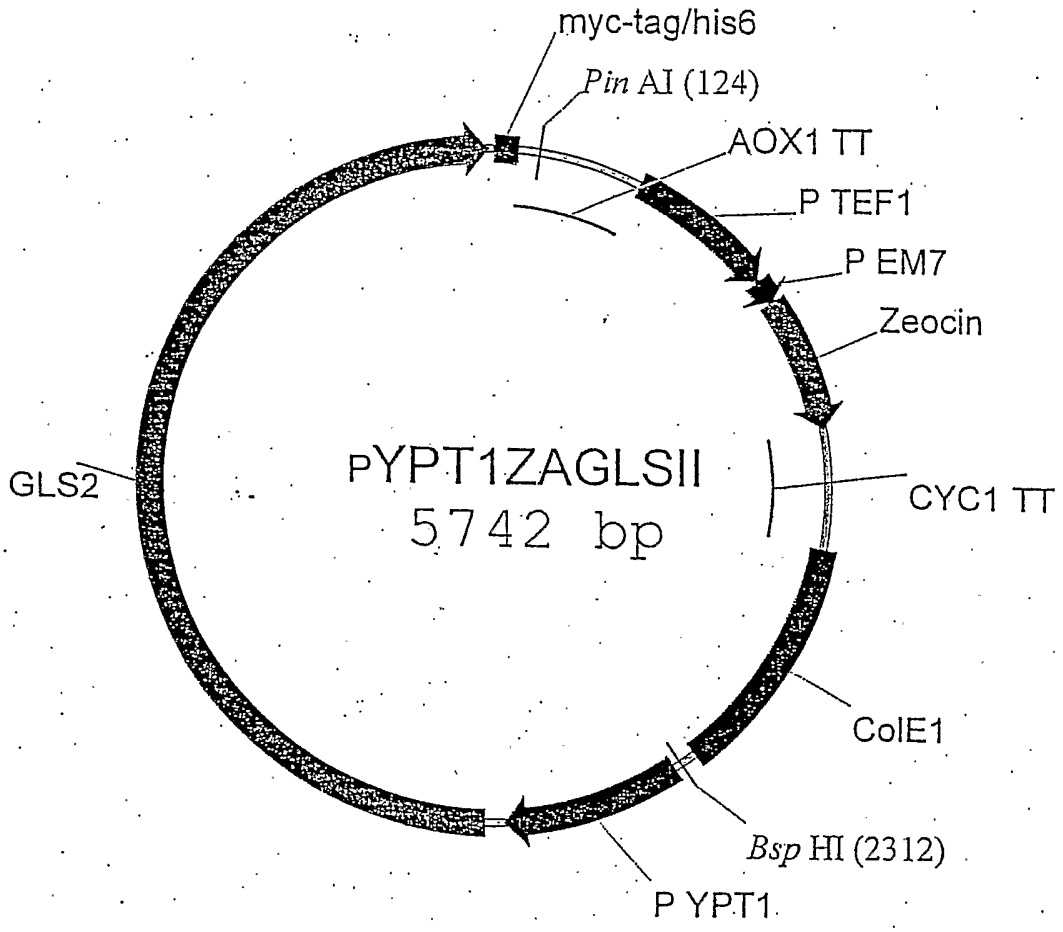

myc-tag/his6
Pin AI (124)
AOX1 TT
P TEF1
P EM7
Zeocin
CYC1 TT
ColE1
Bsp HI (2312)
P YPT1
GLS2
pYPT1ZAGLSII
5742 bp

FIG. 15

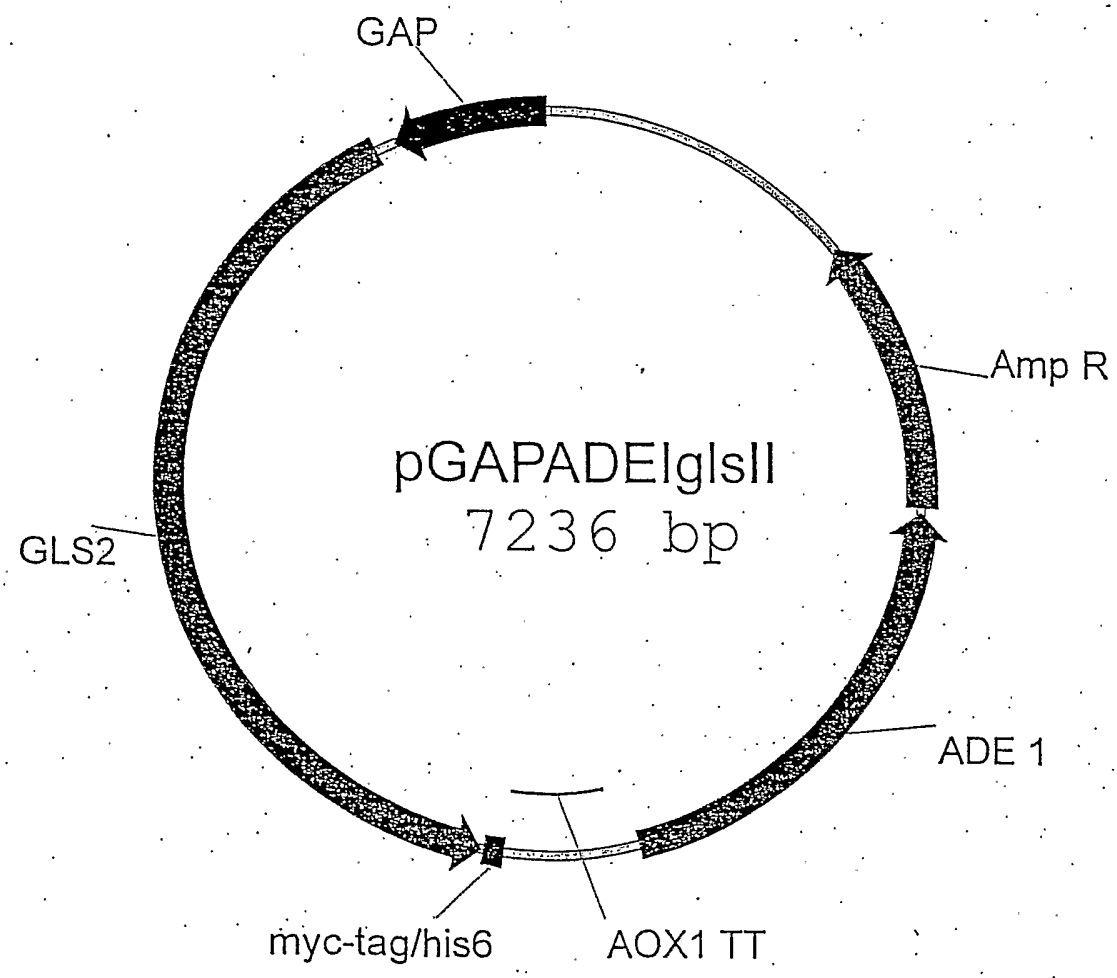

GAP
Amp R
pGAPADEIglsII
7236 bp
GLS2
ADE 1
myc-tag/his6
AOX1 TT

FIG. 16

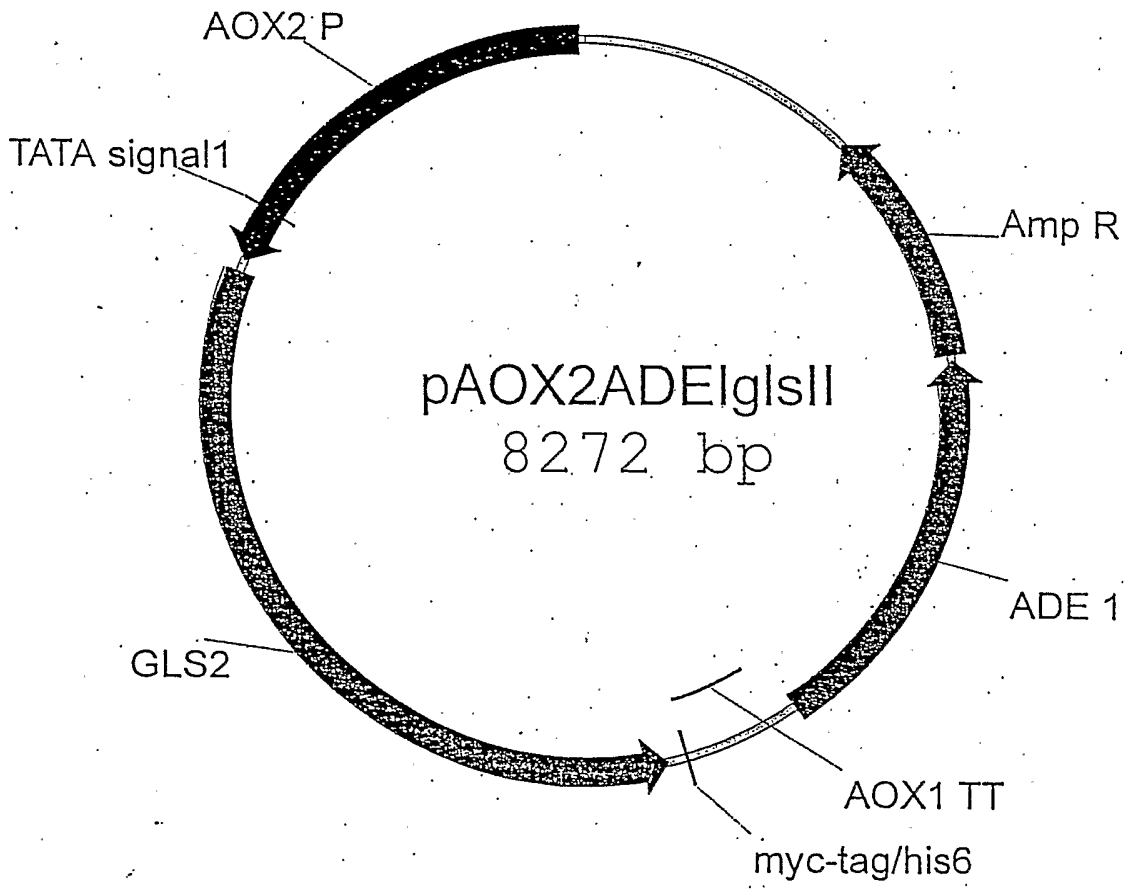

AOX2 P
TATA signal1
Amp R
pAOX2ADEIgIsII
8272 bp
ADE 1
GLS2
AOX1 TT
myc-tag/his6

FIG. 17

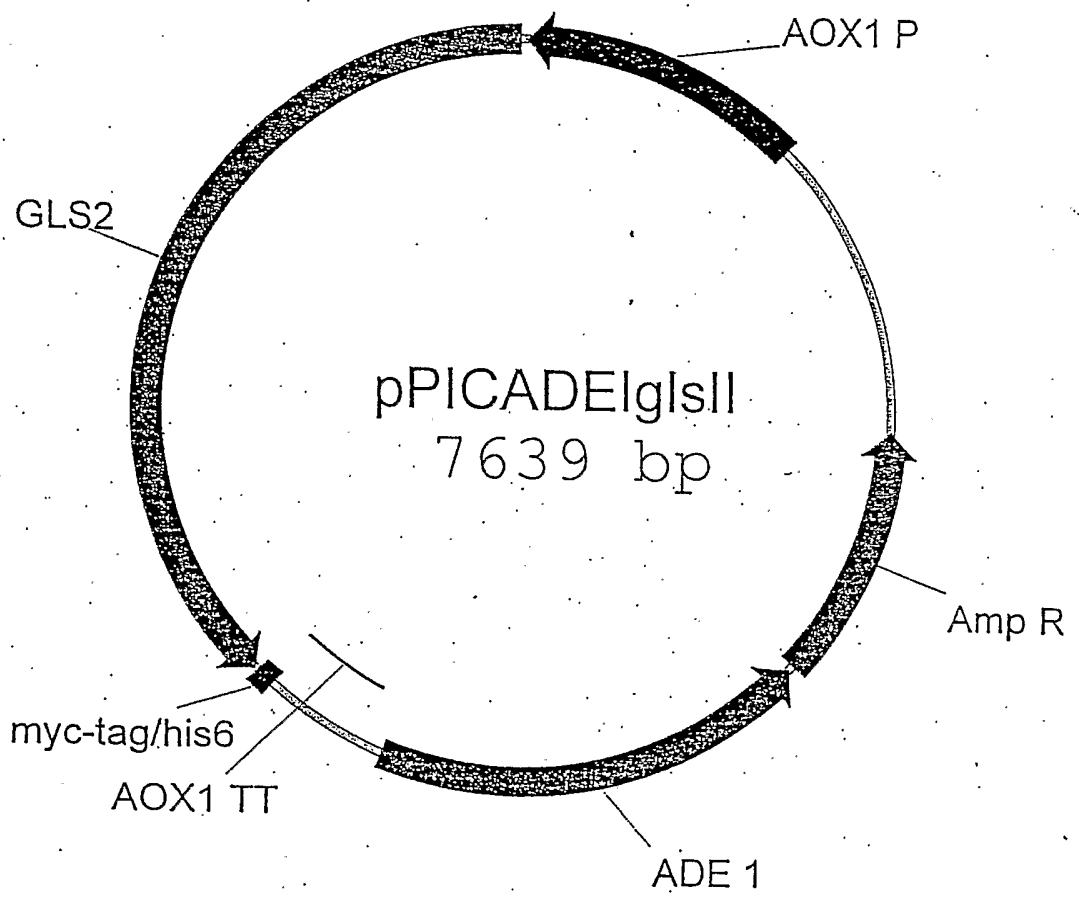

AOX1 P
GLS2
pPICADEIgIsII
7639 bp
Amp R
myc-tag/his6
AOX1 TT
ADE 1

FIG. 18

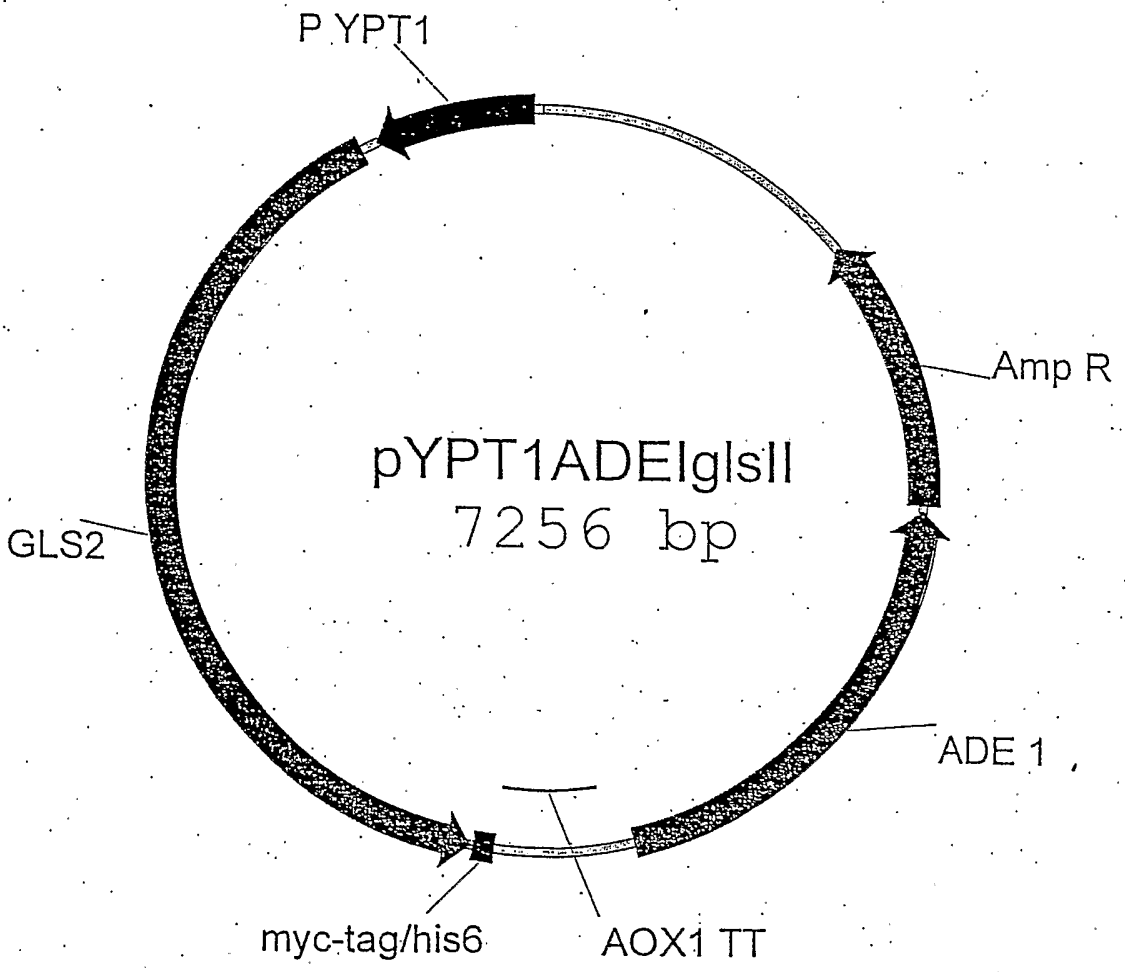

P YPT1
Amp R
pYPT1ADEIglsII
7256 bp
GLS2
ADE 1
myc-tag/his6
AOX1 TT

FIG. 19

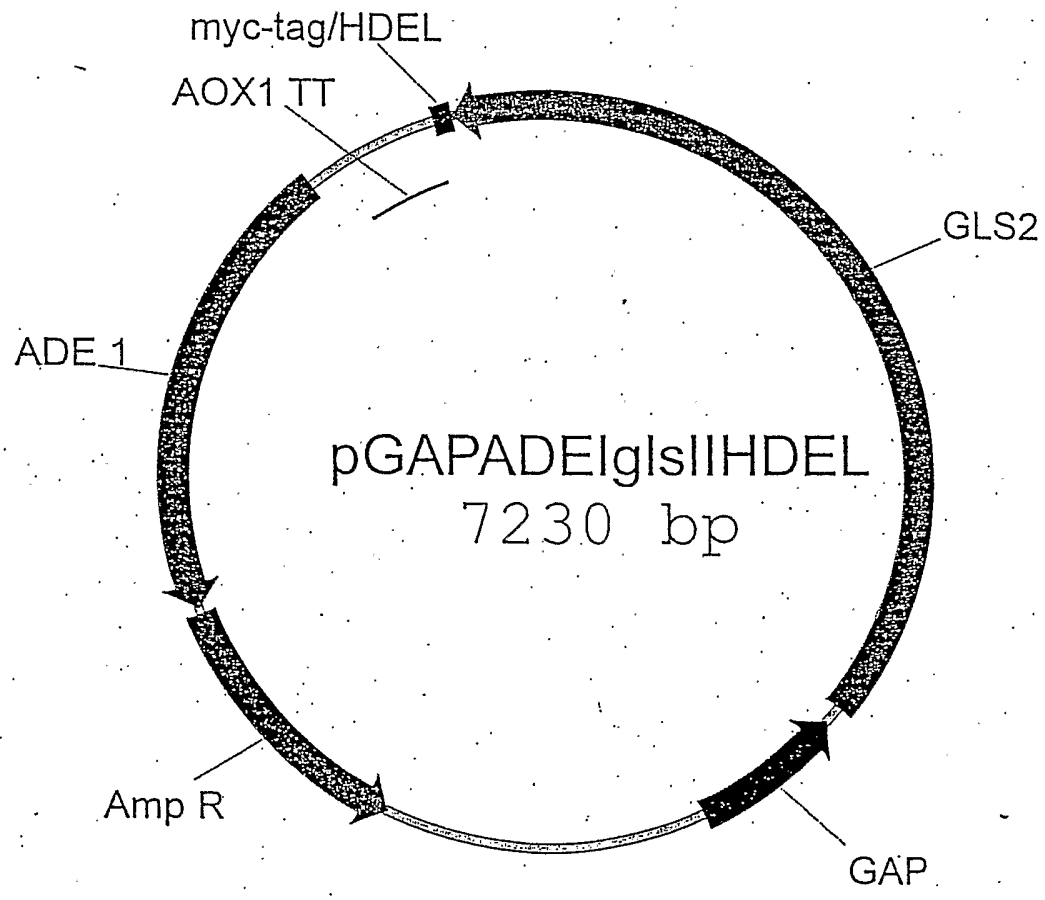

myc-tag/HDEL
AOX1 TT
GLS2
ADE 1
pGAPADEIglsIIHDEL
7230 bp
Amp R
GAP

FIG. 20

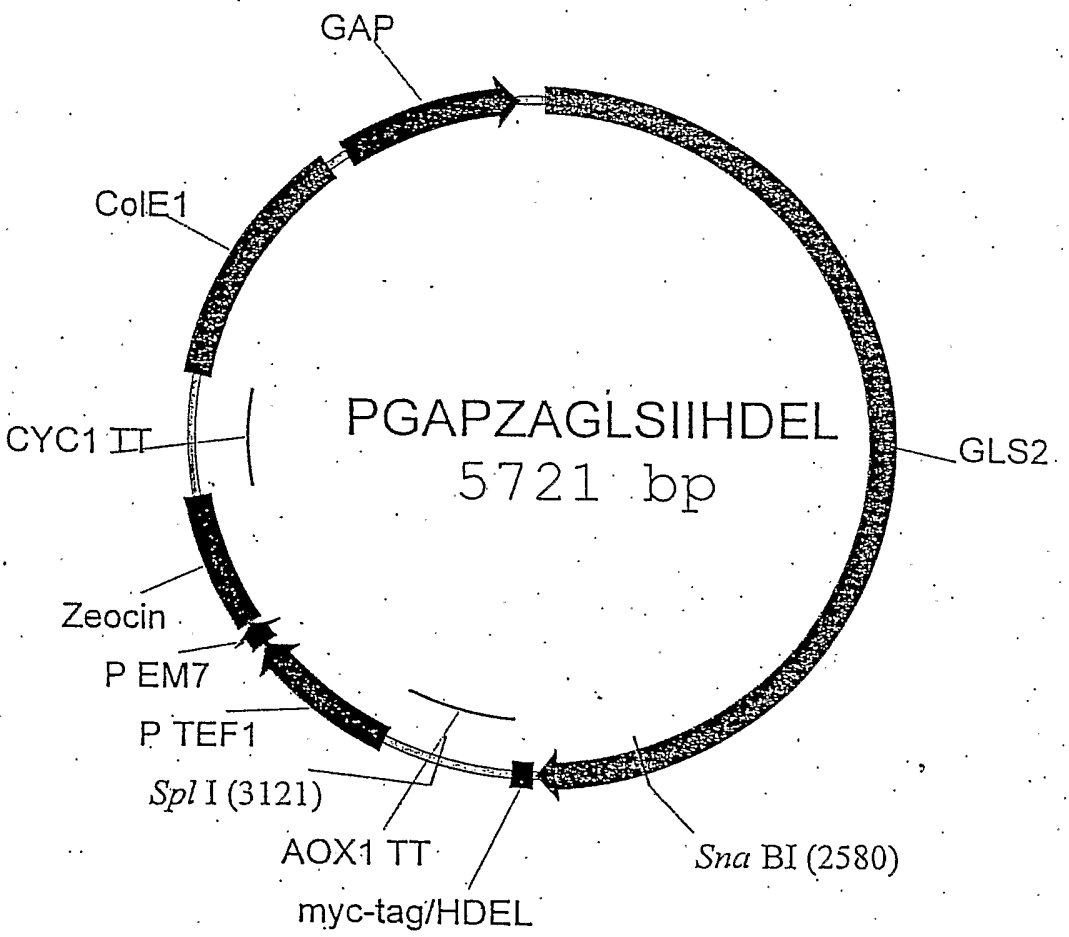

GAP
ColE1
CYC1 TT
PGAPZAGLSIIHDEL
5721 bp
GLS2
Zeocin
P EM7
P TEF1
Spl I (3121)
AOX1 TT
myc-tag/HDEL
Sna BI (2580)

FIG. 21

Glucosidase II assay on commercially available alpha-glucosidase

FIG. 22

Glucosidase II assay on heterologously expressed Pichia protein

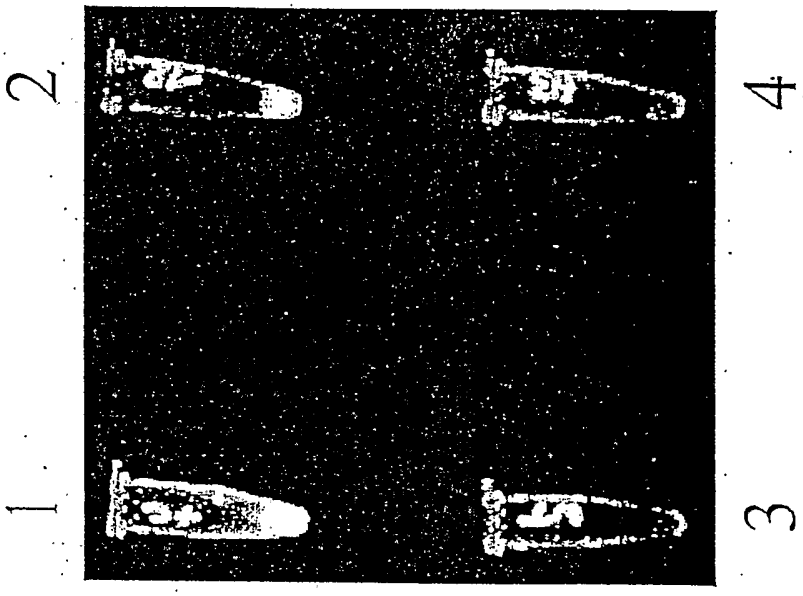


FIG. 23

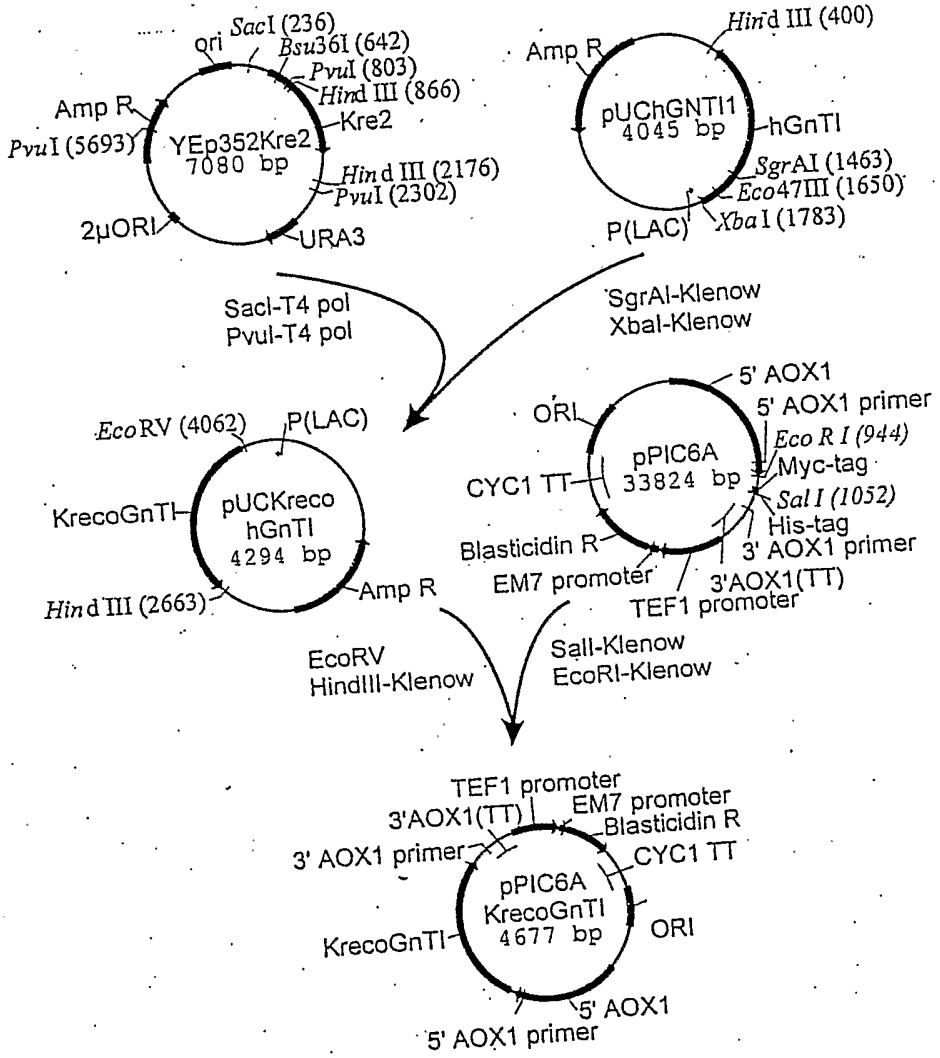

ori
SacI (236)
Bsu36I (642)
PvuI (803)
HindIII (866)
Amp R
Kre2
PvuI (5693)
YEp352Kre2
7080 bp
HindIII (2176)
PvuI (2302)
2µORI
URA3
Amp R
HindIII (400)
pUChGNTI1
4045 bp
hGnTI
SgrAI (1463)
Eco47III (1650)
P(LAC)
XbaI (1783)
SacI-T4 pol
PvuI-T4 pol
SgrAI-Klenow
XbaI-Klenow
EcoRV (4062)
P(LAC)
KrecoGnTI
pUCKreco
hGnTI
4294 bp
HindIII (2663)
Amp R
ORI
5' AOX1
5' AOX1 primer
Eco R I (944)
Myc-tag
CYC1 TT
pPIC6A
33824 bp
SalI (1052)
His-tag
Blasticidin R
3' AOX1 primer
EM7 promoter
3'AOX1(TT)
TEF1 promoter
EcoRV
HindIII-Klenow
SalI-Klenow
EcoRI-Klenow
TEF1 promoter
3'AOX1(TT)
EM7 promoter
Blasticidin R
3' AOX1 primer
CYC1 TT
pPIC6A
KrecoGnTI
4677 bp
ORI
KrecoGnTI
5' AOX1
5' AOX1 primer

FIG. 24

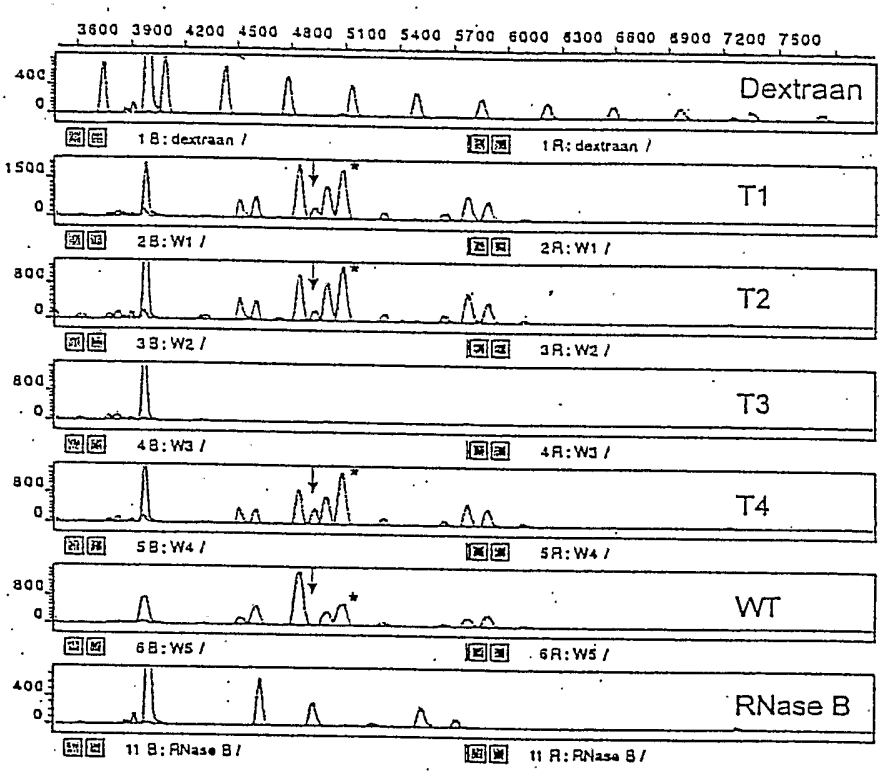

3600 3900 4200 4500 4800 5100 5400 5700 6000 6300 6600 6900 7200 7500
400
0
Dextraan
1B: dextraan /
1R: dextraan /
1500
0
T1
2B: W1 /
2R: W1 /
800
0
T2
3B: W2 /
3R: W2 /
800
0
T3
4B: W3 /
4R: W3 /
800
0
T4
5B: W4 /
5R: W4 /
800
0
WT
6B: W5 /
6R: W5 /
400
0
RNase B
11 B: RNase B /
11 R: RNase B /

FIG. 25

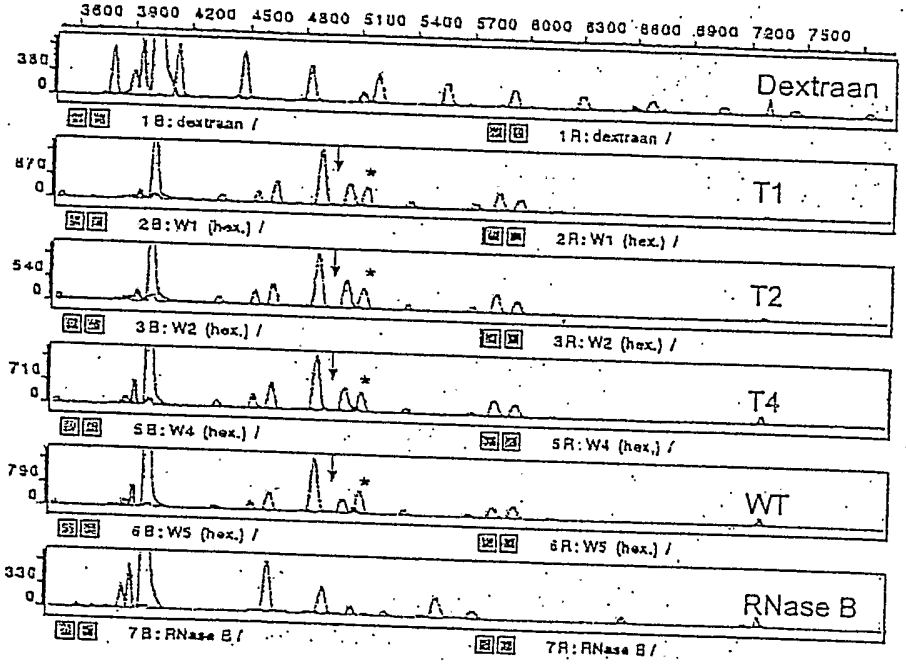

3600 3900 4200 4500 4800 5100 5400 5700 6000 6300 6600 6900 7200 7500
330
0
Dextraan
1B: dextraan /
1R: dextraan /
870
0
T1
2B: W1 (hex.) /
2R: W1 (hex.) /
540
0
T2
3B: W2 (hex.) /
3R: W2 (hex.) /
710
0
T4
5B: W4 (hex.) /
5R: W4 (hex.) /
790
0
WT
6B: W5 (hex.) /
6R: W5 (hex.) /
330
0
RNase B
7B: RNase B /
7R: RNase B /

FIG. 26

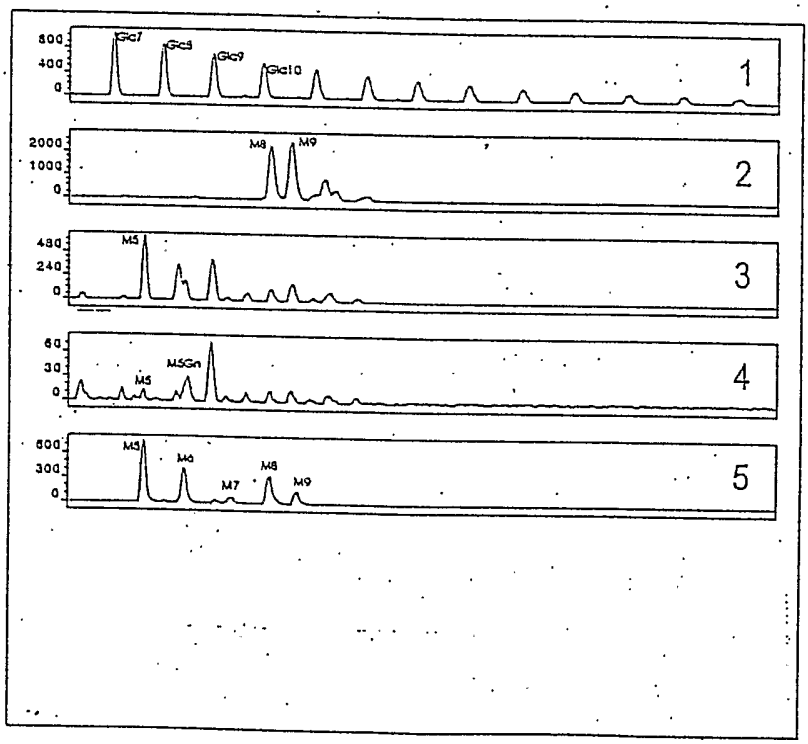

800
400
0
Glc7
Glc8
Glc9
Glc10
1
2000
1000
0
M8
M9
2
480
240
0
M5
3
60
30
0
M5
M5Gn
4
600
300
0
M5
M6
M7
M8
M9
5

FIG. 27

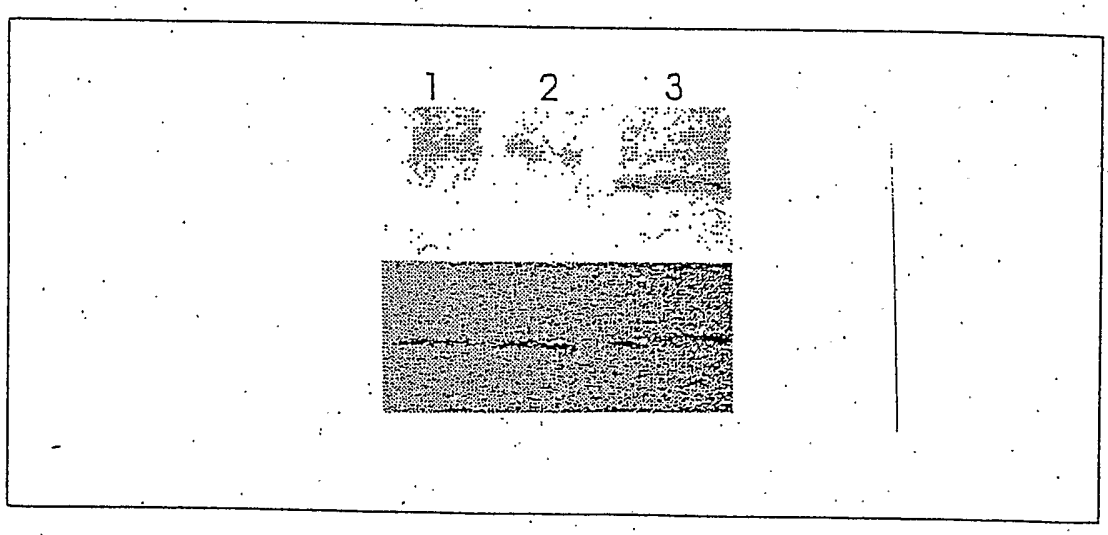


FIG. 28

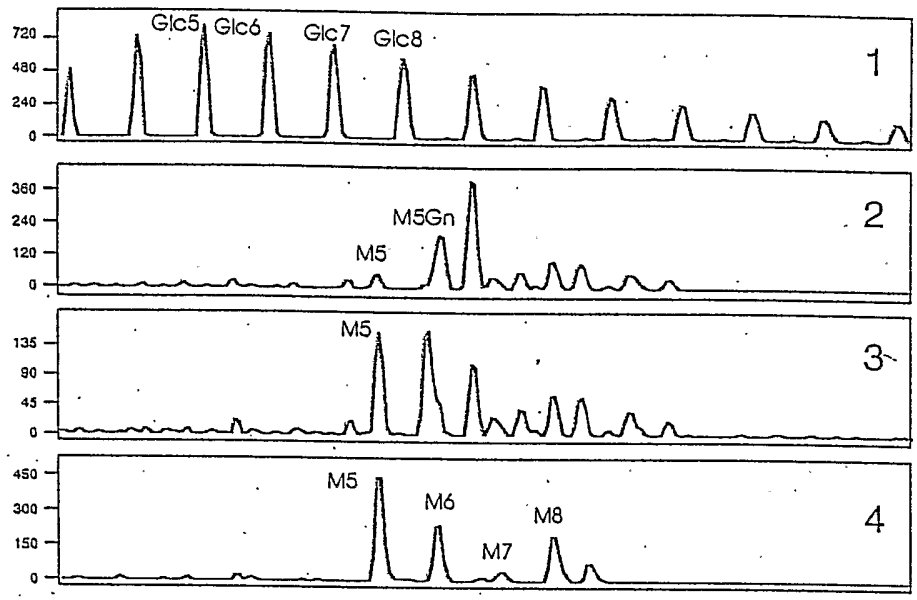

Glc5
Glc6
Glc7
Glc8
1
720
480
240
0
M5Gn
M5
2
360
240
120
0
M5
3
135
90
45
0
M5
M6
M8
M7
4
450
300
150
0

FIG. 29

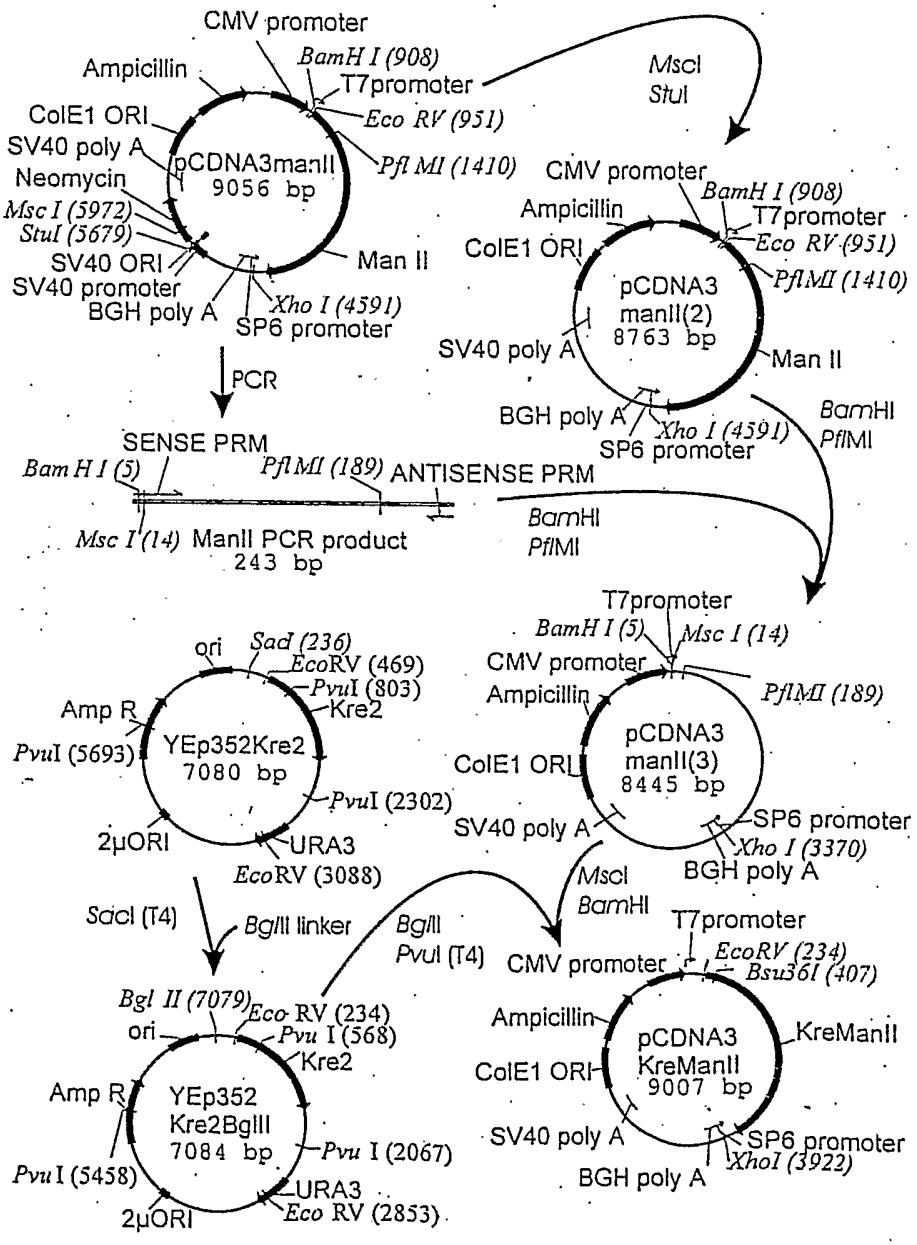

CMV promoter
BamH I (908)
T7promoter
Eco RV (951)
Pfl MI (1410)
Ampicillin
ColE1 ORI
SV40 poly A
Neomycin
Msc I (5972)
StuI (5679)
SV40 ORI
SV40 promoter
BGH poly A
SP6 promoter
Xho I (4591)
Man II
pCDNA3manII
9056 bp
MscI
StuI
PCR
SENSE PRM
Bam H I (5)
PflMI (189)
ANTISENSE PRM
Msc I (14)
ManII PCR product
243 bp
BamHI
PflMI
pCDNA3
manII(2)
8763 bp
T7promoter
BamH I (5)
Msc I (14)
PflMI (189)
pCDNA3
manII(3)
8445 bp
SP6 promoter
Xho I (3370)
ori
Sad (236)
EcoRV (469)
PvuI (803)
Kre2
Amp R
PvuI (5693)
YEp352Kre2
7080 bp
PvuI (2302)
2µORI
URA3
EcoRV (3088)
SacI (T4)
BglII linker
BglII
PvuI (T4)
Bgl II (7079)
Eco RV (234)
Pvu I (568)
YEp352
Kre2BglII
7084 bp
Pvu I (2067)
PvuI (5458)
URA3
Eco RV (2853)
EcoRV (234)
Bsu36I (407)
pCDNA3
KreManII
9007 bp
KreManII
SP6 promoter
XhoI (3922)

FIG. 30A

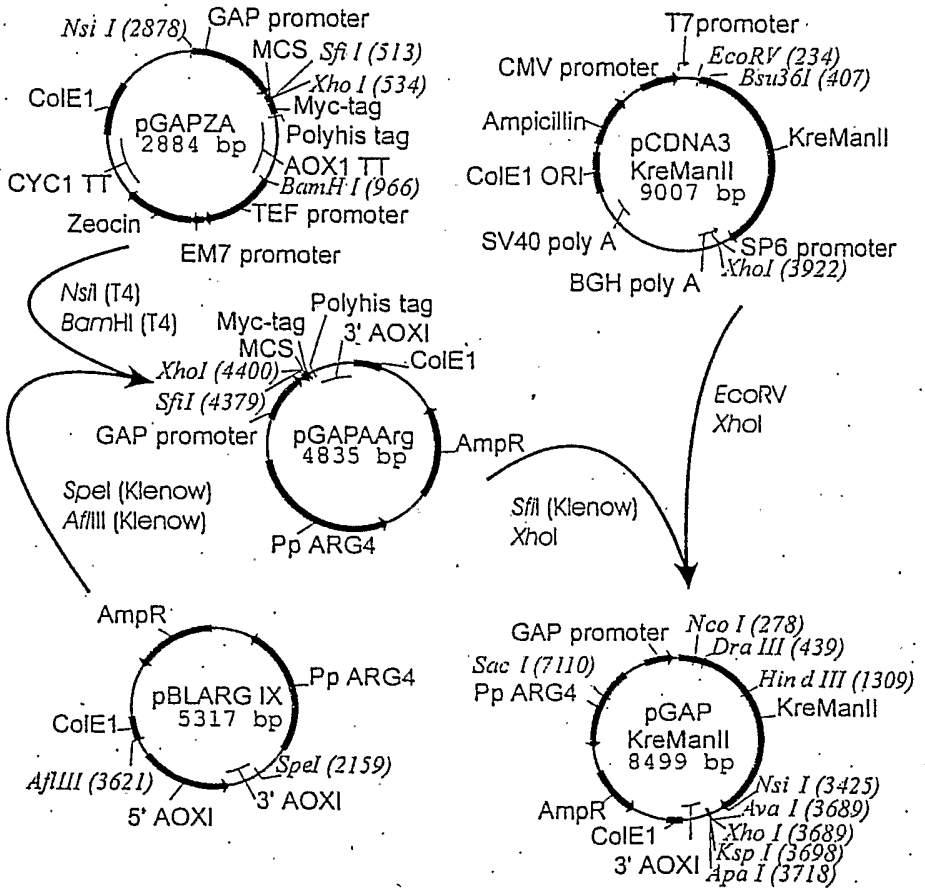

Nsi I (2878)
GAP promoter
MCS
Sfi I (513)
Xho I (534)
Myc-tag
Polyhis tag
AOX1 TT
BamH I (966)
TEF promoter
EM7 promoter
Zeocin
CYC1 TT
ColE1
pGAPZA
2884 bp
T7promoter
EcoRV (234)
Bsu36I (407)
CMV promoter
Ampicillin
ColE1 ORI
pCDNA3
KreManII
9007 bp
KreManII
SV40 poly A
BGH poly A
SP6 promoter
XhoI (3922)
NsiI (T4)
BamHI (T4)
Myc-tag
MCS
Polyhis tag
3' AOXI
ColE1
XhoI (4400)
SfiI (4379)
GAP promoter
pGAPAArg
4835 bp
AmpR
SpeI (Klenow)
AflII (Klenow)
Pp ARG4
EcoRV
XhoI
SfiI (Klenow)
XhoI
AmpR
pBLARG IX
5317 bp
Pp ARG4
ColE1
AflIII (3621)
SpeI (2159)
5' AOXI
3' AOXI
GAP promoter
Nco I (278)
Dra III (439)
Sac I (7110)
Pp ARG4
Hin d III (1309)
KreManII
pGAP
KreManII
8499 bp
Nsi I (3425)
Ava I (3689)
AmpR
ColE1
Xho I (3689)
Ksp I (3698)
3' AOXI
Apa I (3718)

FIG. 30B

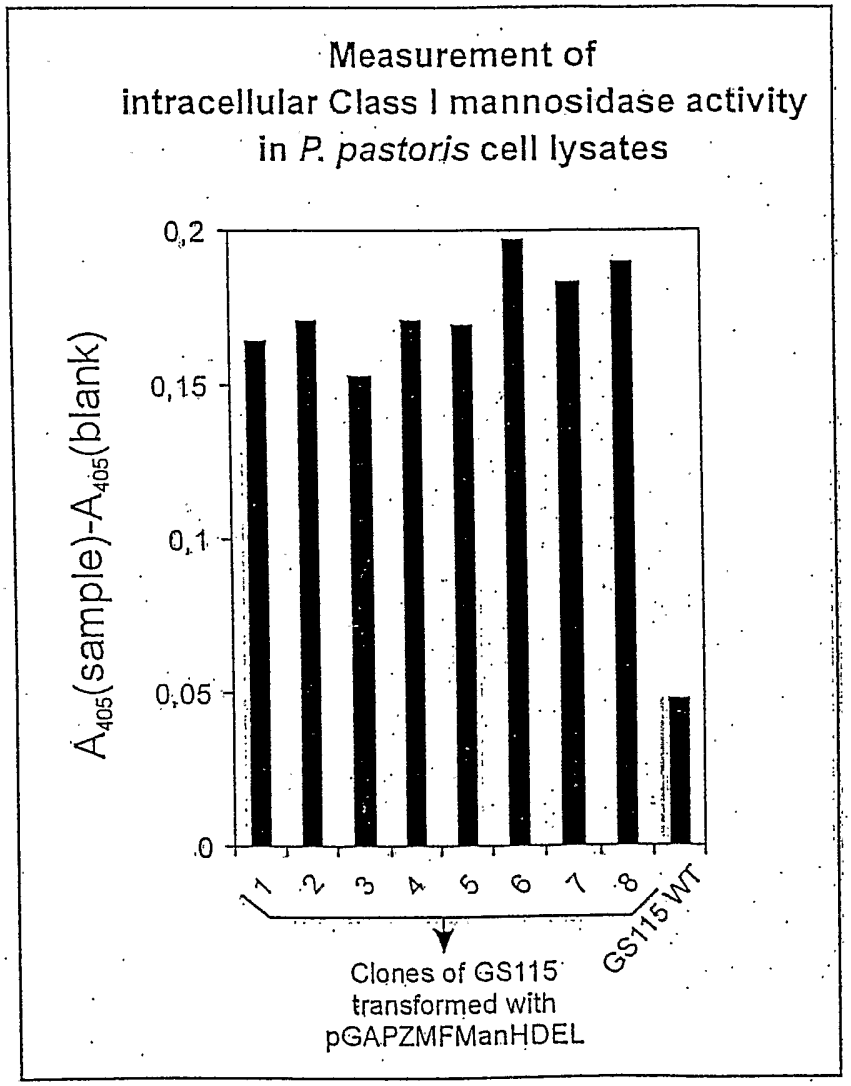
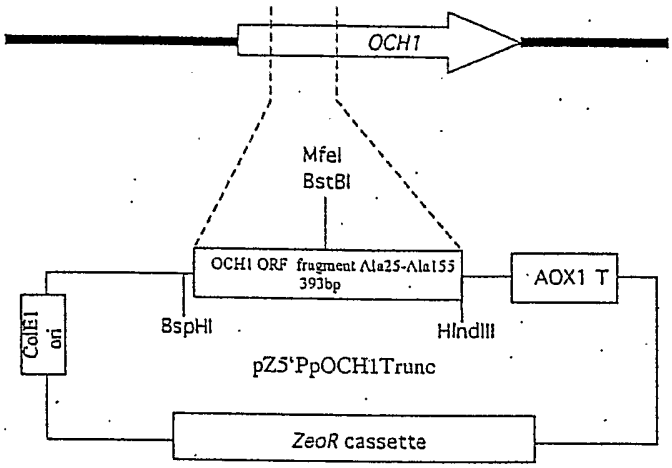
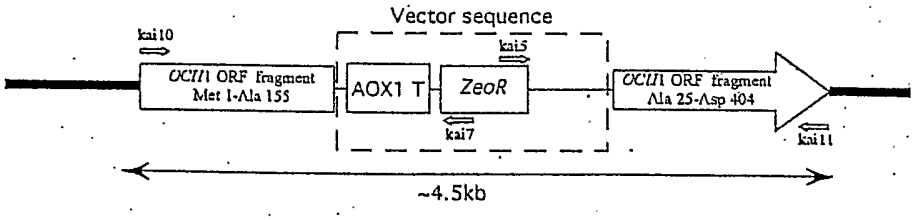

Measurement of
intracellular Class I mannosidase activity
in P. pastoris cell lysates
A405(sample)-A405(blank)
0,2
0,15
0,1
0,05
0
1
2
3
4
5
6
7
8
GS115 WT
Clones of GS115
transformed with
pGAPZMFManHDEL

FIG. 31

Construct for single homologous recombination in the Pichia OCH1 locus.



Structure of *OCH1* locus after recombination with pZ5'PpOCH1Trunc.



Kai 10 - Kai7 = ~ 1050 bp

Kai 11 - Kai5 = ~ 2100 bp

FIG. 32

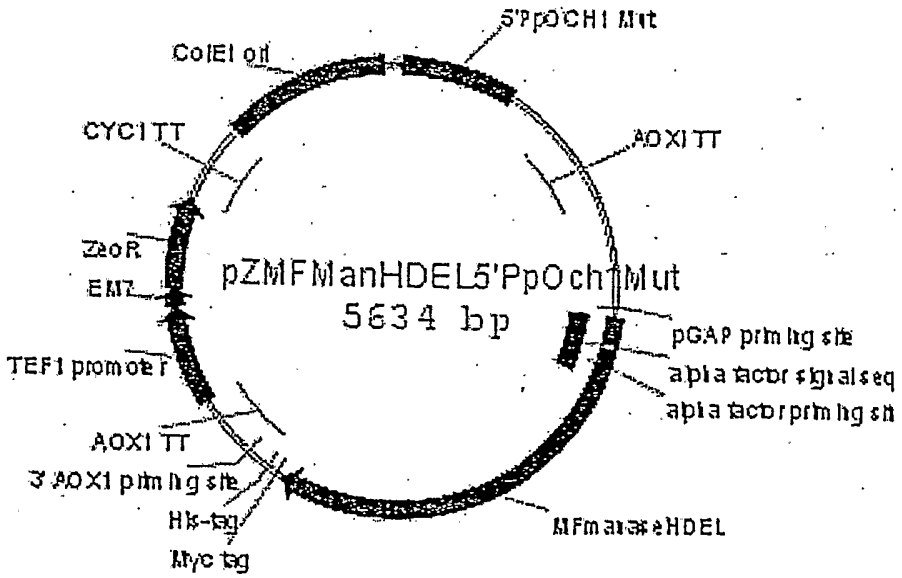

5'PpOCH1 Mut
ColEI ori
CYC1 TT
AOX1 TT
ZeoR
EM7
pZMFManHDEL5'PpOch1Mut
5634 bp
pGAP prim hg site
alpha factor signal seq
alpha factor prim hg sit
TEF1 promoter
AOX1 TT
3' AOX1 prim hg site
His-tag
Myc tag
MFmannaseHDEL

FIG. 33

Clones of GS115 transformed with pZ5'PpOCH1Trunc

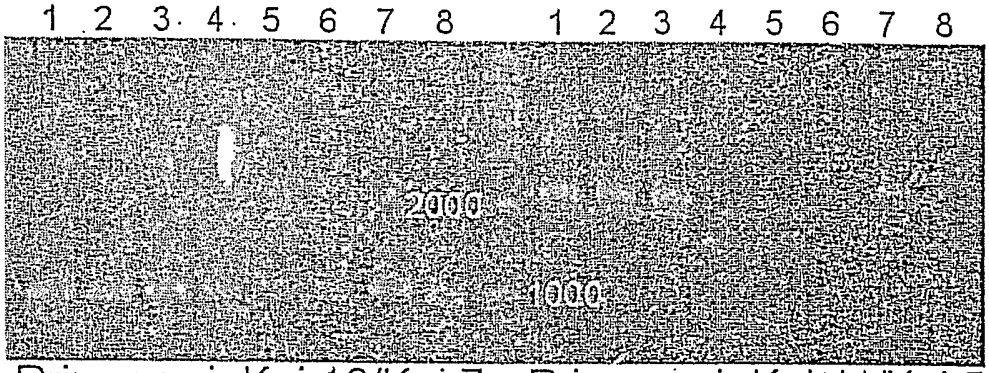


FIG. 34

SEQUENCE LISTING

SEQ ID NO: 1
HDEL (peptide)

SEQ ID NO: 2
5'-GACTGGTTCCAATTGACAAGC-3'

SEQ ID NO: 3
5'-AGTCTAGATTACAACTCGTCGTGAGCAAGGTGGCCGCCCCG TCG-3'

SEQ ID NO: 4
CCATTGAGGACGCATGCCGCGCC

SEQ ID NO: 5
GTATCTAGATTACAACTCGTCGTGCAGATCCTCTTCTGAGATGAGTTTTTGT TCAGCAAGGTGGCCGCCCCGTCGTGATGATGAA

SEQ ID NO: 6
AACTCGAGATGGACTCTTCAAAACACAAACGC

SEQ ID NO: 7
TTGCGGCCGCTTACAACTCGTCGTGTCGGACAGCAGGATTACCTGA

SEQ ID NO: 8
CCATTGAGGACGCATGCCGCGCC

SEQ ID NO: 9
GCAAATGGCATTCTGACATCCT

SEQ ID NO: 10
GTCCCTATTTCAATCAATTGAA

SEQ ID NO:11
GACTGGTTCCAATTGACAAGC

SEQ ID NO:12
CCG CTC GAG ATG GTC CTT TTG AAA TGG CTC

SEQ ID NO:13

CCG GGC CCA AAA ATA ACT TCC CAA TCT TCAG

SEQ ID NO: 14

The ORF sequence of the MFManHDEL fusion in pGAPZMFManHDEL:

*atg*agatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaa
acggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaaca
gcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaagag
aggctgaagctgaattcgccacaaaacgtggatctcccaaccctacgagggcggcagcagtcaaggccgcattccagac
gtcgtggaacgcttaccaccattttgcctttccccatgacgacctccacccggtcagcaacagctttgatgatgagagaaacg
gctggggctcgtcggcaatcgatggcttggacacggctatcctcatgggggatgccgacattgtgaacacgatccttcagta
tgtaccgcagatcaacttcaccacgactgcggttgccaaccaaggatcctccgtgttcgagaccaacattcggtacctcggt
ggcctgctttctgcctatgacctgttgcgaggtcctttcagctccttggcgacaaaccagaccctggtaaacagccttctgag
gcaggctcaaacactggccaacggcctcaaggttgcgttcaccactcccagcggtgtcccggaccctaccgtcttcttcaac
cctactgtccggagaagtggtgcatctagcaacaacgtcgctgaaattggaagcctggtgctcgagtggacacggttgagc
gacctgacgggaaacccgcagtatgcccagcttgcgcagaagggcgagtcgtatctcctgaatccaaagggaagcccgg
aggcatggcctggcctgattggaacgtttgtcagcacgagcaacggtacctttcaggatagcagcggcagctggtccggc
ctcatggacagcttctacgagtacctgatcaagatgtacctgtacgacccggttgcgtttgcacactacaaggatcgctgggt
ccttggtgccgactcgaccattgggcatctcggctctcacccgtcgacgcgcaaggacttgacctttttgtcttcgtacaacg
gacagtctacgtcgccaaactcaggacatttggccagttttggcggtggcaacttcatcttgggaggcattctcctgaacgag
caaaagtacattgactttggaatcaagcttgccagctcgtactttggcacgtacacccagacggcttctggaatcggccccga
aggcttcgcgtgggtggacagcgtgacgggcgccggcggctcgccgccctcgtcccagtccgggttctactcgtcggca
ggattctgggtgacggcaccgtattacatcctgcggccggagacgctggagagcttgtactacgcataccgcgtcacggg
cgactccaagtggcaggacctggcgtgggaagcgttgagtgccattgaggacgcatgccgcgccggcagcgcgtactcg
tccatcaacgacgtgacgcaggccaacggcgggggtgcctctgacgatatggagagcttctggtttgccgaggcgctcaa
gtatgcgtacctgatctttgcggaggagtcggatgtgcaggtgcaggccaccggcgggaacaaatttgtctttaacacgga
ggcgcaccccttagcatccgttcatcatcacgacggggcggccaccttgctcacgacgagttg*taa*

SEQ ID NO: 15

The ORF sequence of the MFmManHDEL fusion in pGAPZMFmManHDEL:

*atg*agatttccttcaatttttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaa
acggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaaca
gcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagatggact
cttcaaaacacaaacgctttgatctgggcttagaagatgtgttaattcctcacgtagatgccggcaaaggagctaaaaacccc
ggcgtcttcctgatccatggacccgacgaacacagacacagggaagaagaagagcgtctgagaaataagattagagctg
accatgagaaagccctggaagaagcaaaagaaaaattaagaaagtcaagagaggaaatccgtgcagaaattcagacaga
gaaaaacaaagtagcccaagcaatgaagacaaaagagaccagggtactgccgcctgtccctgtcccacaacgtgtaggg
gtcagtggtggggatccagaagacatggagatcaagaagaaaagagacaaaattaaagagatgatgaaacatgcctggg
ataattacagaacatacggatggggacataatgaactaaggcctattgcaaggaaaggccattccactaacatattcggaag
ctcacagatgggtgccaccatagtggatgctttggataccctttatatcatggggcttcatgatgaattcatggatgggcaaag
atggattgaagaaaaccttgatttcagtgtgaattcagaagtgtctgtctttgaagttaacattcgctttattggagggctcctcg
ctgcatattacctgtcaggagaggaaatattcaagactaaagcagtgcagttggctgagaaactccttcctgcctttaacacac
ctactgggattccctgggcaatggtgaacctgaaaagtggagtaggtcgaaactggggctgggcgtctgcaggcagcagc
atcctggctgagttcggcaccctgcacatggagtttgtgcacctcagctacttgaccggtgacttgacttactataataaggtc
atgcacattcggaaactactgcagaaaatggaacgcccaaatggtctttatccaaattatttaaacccaagaacagggcgctg
gggtcagtatcacacatcagttggtggtctgggagatagtttttatgaatacttactgaaagcatggctgacgtcagataaaac
agaccacgaggcaagaaggatgtatgacgatgctgttgaggctatagaaaaacatcttattaagaagtcccgaggaggtct
ggtttttattggagaatggaagaatggacacttggaaaggaagatggggcacttggcctgctttgctgggggaatgcttgcc
cttggagcagatggttccagaaaggataaagctggccactacttagaactaggggcagaaattgcacgaacatgtcatgag
tcatatgacagaactgcattgaaactaggtccggagtcattcaagtttgatggtgcagtggaagccgtggctgtgcggcagg
ctgaaaagtattacatccttcgtccagaagtaattgaaacctattggtatctatggcgatttacccacgacccaagatacaggc
agtggggctgggaagcagcactggctattgagaagtcgtgccgggtcagcggtgggttttctggtgtcaaggatgtatacg
ccccgacccctgtgcatgacgacgtgcagcagagcttttctcttgctgaaacattaaaatacttgtacctgctgttctctggcga
tgaccttctacctttagaccactgggtgtttaacacagaggcgcaccctctgccggtgttgcgcttagccaacagcactctttc
aggtaatcctgctgtccgacacgacgagttg*taa*

SEQ ID NO: 16 pAOX2ZAGLSII

catggccaagttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccgg
ctcgggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtc
caggaccaggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcg
gaggtcgtgtccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggag
gttcgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcc
cacgggtcccaggcctcggagatccgtccccctttccctttgtcgatatcatgtaattagttatgtcacgcttacattcacgccct
ccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagt
attaagaacgttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttga
gaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccagga
accgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtca
gaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgac
cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagt
tcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccccgttcagcccgaccgctgcgccttatccggtaact
atcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggt
atgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctg
ctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtt
tgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga
acgaaaactcacgttaagggattttggtcatgagatcagatctttttttcagaccatatgaccggtccatcttctacgggggggat
tatctatgctttgacctctatcttgattcttttatgattcaaatcacttttacgttatttattacttactggttatttacttagcgccttttctg
aaaaacatttactaaaaatcatacatcggcactctcaaacacgacagattgtgatcaagaagcagagacaatcaccactaag
gttgcacatttgagccagtaggctcctaatagaggttcgatacttattttgataatacgacatattgtcttacctctgaatgtgtca
atactctctcgttcttcgtctcgtcagctaaaaatataacacttcgagtaagatacgcccaattgaaggctacgagataccaga
ctatcactagtagaactttgacatctgctaaagcagatcaaatatccatttatccagaatcaattaccttcctttagcttgtcgaag
gcatgaaaaagctacatgaaaatccccatccttgaagttttgtcagcttaaaggactccatttcctaaaatttcaagcagtcctct
caactaaatttttttccattcctctgcacccagccctcttcatcaaccgtccagccttctcaaaagtccaatgtaagtagcctgca
aattcaggttacaaccccctcaattttccatccaagggcgatccttacaaagttaatatcgaacagcagagactaagcgagtcat
catcaccacccaacgatggtgaaaaactttaagcatagattgatggagggtgtatggcacttggcggctgcattagagtttga
aactatggggtaatacatcacatccggaactgatccgactccgagatcatatgcaaagcacgtgatgtaccccgtaaactgc

tcggattatcgttgcaattcatcgtcttaaacagtacaagaaactttattcatgggtcattggactctgatgaggggcacatttcc
ccaatgatttttgggaaagaaagccgtaagaggacagttaagcgaaagagacaagacaacgaacagcaaaagtgacag
ctgtcagctacctagtggacagttgggagtttccaattggttggttttgaatttttacccatgttgagttgtccttgcttctccttgca
aacaatgcaagttgataagacatcaccttccaagataggctatttttgtcgcataaatttttgtctcggagtgaaaaccccttttat
gtgaacagattacagaagcgtcctacccttcaccggttgagatggggagaaaattaagcgatgaggagacgattattggtat
aaaagaagcaaccaaaatcccttattgtccttttctgatcagcatcaaagaatattgtcttaaaacgggcttttaactacattgttc
ttacacattgcaaacctcttccttctatttcggatcaactgtattgactacattgatctttttttaacgaagtttacgacttactaaatcc
ccacaaacaaatcaactgagaaaagaattcacgtggcccagccggccgtctcggatcggtacctcgagatggtccttttga
aatggctcgtatgccaattggtcttctttaccgctttttcgcatgcgtttaccgactatctattaaagaagtgtgcgcaatctgggt
tttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctattacaaagtggacgccgagtctattgcac
acgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatgatatagccgttcagttccc
attctctctctctttttacaggatcactcagtaaggttcactataaatgagaaagagagaatgccaaccaacagcagcggtttg
ttgatctcttcacaacggttcaatgagacctggaagtacgcattcgacaagaaatttcaagaggaggcgaacaggaccagta
ttccacaattccacttccttaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcatttttgtcactttcaaactccac
tgcagacacatttcatcttcgaaacggtgatgtatccgtagaaatctttgctgaacctttcaattgaaagtttactggcaaaatg
cgctgaaacttattgtaaacgagcaaaatttcctgaacattgaacatcatagaactaagcaggaaaacttcgcacacgtgctg
ccagaagaaacaactttcaacatgtttaaggacaatttcttgtattcaaagcatgactctatgccttttggggcctgaatcggttg
cgctagatttctctttcatgggttctactaatgtctacggtataccggaacatgcgacgtcgctaaggctgatggacacttcagg
tggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaacatcggtaccagccaaccaatgtacggttcgatcc
cattcatgttttcatcttcgtccacatctatcttttgggtcaatgcagctgacacttgggtagacataaagtatgacaccagtaaa
aataaaacgatgactcattggatctccgaaaatggtgtcatagatgtagtcatgtccctggggccagatattccaactatcattg
acaaatttaccgatttgactggtagacccttttaccgcccatttcctctatagggtaccatcaatgtagatggaattataatgatg
agatggacgttctcacagtggactctcagatggatgctcatatgattccttacgattttatttggttggacttggagtatacgaac
gacaaaaaatattttacttggaagcagcactcctttcccaatccaaaaaggctgttatccaaattaaaaaagttgggtagaaat
cttgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgacagggtaattaatgaaaatgtagcagtcaaggatc
acaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgataccataagcaaatatggccaaaagattt
ggaagtccttttcgaacggtttatggatctgccggctgatttaactaatttattcatttggaatgatatgaacgagccttcgatttt
cgatggcccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtccataacatatatggtc

tatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccatccgataagcgtcctttccttctaacaagggcttttt
ttgccggctctcaacgtactgctgccacatggactggtgacaatgtggccaattgggattacttaaagatttccattcctatggt
tctgtcaaacaacattgctggtatgccatttataggagccgacatagctggctttgctgaggatcctacacctgaattgattgca
cgttggtaccaagcgggcttatggtacccatttttagagcacacgcccatatagacaccaagagaagagaaccatacttatt
caatgaacctttgaagtcgatagtacgtgatattatccaattgagatatttcctgctacctaccttatacaccatgtttcataaatca
agtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttgctgaattgtatcatatcgataaccaattttact
ggagtaattcaggtctattagtcaaacctgtcacggagcctggtcaatcagaaacggaaatggttttcccacccggtatattct
atgaattcgcatctttacactcttttataaacaatggtactgatttgatagaaaagaatatttctgcaccattggataaaattccatt
atttattgaaggcggtcacattatcactatgaaagataagtatagaagatcttcaatgttaatgaaaaacgatccatatgtaatag
ttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatggagaaacttttggctaccaaagaggtgag
tacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatattcccgagaatttgacaggcattcacc
acaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataatttacaacacaacataacgttgaaagacagta
ttaaagtcaaaaaaaatggcgaagaaagttcattgccgactagatcgtcatatgagaatgataataagatcaccattcttaacc
tatcgcttgacataactgaagattgggaagttatttttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgcc
gtcgaccatcatcatcatcattgagttttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccg
gtcttgctagattctaatcaagaggatgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaac
ctatatagtataggatttttttttgtcattttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgt
ggtaggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagac
cttcgtttgtgcggatcccccacacaccatagcttcaaaatgtttctactccttttttactcttccagattttctcggactccgcgca
tcgccgtaccacttcaaaacacccaagcacagcatactaaatttccctctttcttcctctagggtgtcgttaattacccgtacta
aaggtttggaaaagaaaaaagagaccgcctcgtttcttttttcttcgtcgaaaaaggcaataaaaatttttatcacgtttctttttctt
gaaatttttttttttttagttttttttctctttcagtgacctccattgatatttaagttaataaacggtcttcaatttctcaagtttcagtttcattt
ttcttgttctattacaacttttttttacttcttgttcattagaaagaaagcatagcaatctaatctaagggcggtgttgacaattaatcat
cggcatagtatatcggcatagtataatacgacaaggtgaggaactaaac

SEQ ID NO: 17 pAOX2ADE1glsII

tcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcac
tgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatca

ggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggc
gtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact
ataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccg
cctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaaga
cacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttg
aagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaag
agttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaa
aaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttgg
tcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaa
cttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactcc
ccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcac
cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctcca
tccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacagg
catcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccat
gttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatg
gcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgaga
atagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtg
ctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgt
gcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaa
gggaataagggcgacacggaaatgttgaatactcatactcttccttttcaatagctccaaggcaacaaattgactactcagac
cgacattcattcgttattgattttaaatcaacgataaacggaatggttacttgaatgatttcactttatgatcattgtttactaattacc
taaataggattttatatggaattggaagaataagggaaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaag
cccatttcttgccagtaattgcttcataagcttcaatatacttttctttactcttgatagcaatttctgcatccatggctacgccctctt
tgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttacccacttggtaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgttttcgtccagtccaaattcgaatttcgt
atcagcaataatgatccccttcaaaagggcgaagtttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
aaatgtctttacctacaatctcagcagcttgttcaatagagatgttttcatcgtgttcaccctgttcagctttcgttgaaggtgtgaa

aatcggagttggaaaggcgtcgctctcttgaaggttctcgttttcaaccttgactccatggacagtttttgagttcttgtactctttc
catgcacttccagtgatgtaacctctgacaatggcttccaaaggtatcagtctgtgctttttttactatcaaggatcgtccctctaat
tgagatttgtatttttcttcagacagttttgatggtagtaaagcaaagacttccttgtcattagaagcaaccaaatgattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatc
gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttcctttggcgatt
aaaggtaggattccgtccagatcagtgttcacaatggacatacttggaaggatacagcaaagtgtgttggaagcgatgacac
atggaaaggaatttttcgagtttcctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttc
aaagcgcccaagtgtttctgccaacctgacttttttctgaataatgaatcgttcaagtggagtatttaaaccatgattaagttacgt
gatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcggtgcgggcctcttcgctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaactcgaatgattttcccaaaccccctaccacaagatattcatcagctgcgagat
aggctgatcaggagcaagctcgtacgagaagaaacaaaatgacaaaaaaaatcctatactatataggttacaaataaaaaa
gtatcaaaaatgaagcctgcatctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgt
aagtgcccaacttgaactgaggaacagtcatgtctaaggctaaaactcaatgatgatgatgatgatggtcgacggcgctattc
agatcctcttctgagatgagttttgttcgggcccaaaaataacttcccaatcttcagttatgtcaagcgataggttaagaatggt
gatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccatttttttgactttaatactgtctttcaacgttatgt
tgtgttgtaaattattctttgcgataatgatttttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcgggaa
tatgacttcgaacatttttaaggtattgttttcgaaaatgaactgagtttctacgtactcacctctttggtagccaaaagtttctcca
tcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgtttttcattaacattga
agatcttctatacttatctttcatagtgataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattctttt
ctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgtttct
gattgaccaggctccgtgacaggtttgactaatagacctgaattactccagtaaaattggttatcgatatgatacaattcagcaa
attcagggtgttcaataaacattggattcattatcggaaatccagtgacacttgatttatgaaacatggtgtataaggtaggtag
caggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcattgaataagtatggttctcttctcttggtgtctata
tgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattcaggtgtaggatcctcagcaa
agccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaaccataggaatggaaatctttaagtaatccc
aattggccacattgtcaccagtccatgtggcagcagtacgttgagagccggcaaaaaaagcccttgttagaaggaaaggac

gcttatcggatggtgaataaatcgattttattgcgtcgtaagtagcttcatgcactgatagaccatatatgttatggacggatcttt
cctcaatgtaattgtcgtgaatcaaatcttttggagctgtggtctctgggccatcgaaaatcgaaggctcgttcatatcattccaa
atgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaaggacttccaaatcttttggccatatttgcttatggt
atcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttgactgctacattttcattaattacc
ctgtcactgatttcataatctttctttaaatgaggatcgattagtacgacaagatttctacccaacttttttaatttggataacagcct
ttttggattgggaaaggagtgctgcttccaagtaaaatatttttttgtcgttcgtatactccaagtccaaccaaataaaatcgtaag
gaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccatctacattgatggtaccctata
gaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgtcaatgatagttggaatatctggccccagggaca
tgactacatctatgacaccattttcggagatccaatgagtcatcgttttatttttactggtgtcatactttatgtctacccaagtgtca
gctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggctggtaccg
atgttgtactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtccatcagccttagcgacgtcgca
tgttccggtataccgtagacattagtagaacccatgaaagagaaatctagcgcaaccgattcaggccccaaaggcatagag
tcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgtttcttctggcagcacgtgtgcgaagttttcctgcttagttc
tatgatgttcaatgttcaggaaattttgctcgtttacaataagtttcagcgcattttgccagtaaactttcaattgaaaaggttcagc
aaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtttgaaagtgacaaaaatgaagatattttc
gaccagaatgagttcacagtttgtttttgcttaaggaagtggaattgtggaatactggtcctgttcgcctcctcttgaaatttcttgt
cgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttggttggcattctctctttctcattt
atagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaactgaacggctatatcatcgccctccaatcttg
gtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagtg
atgagatttggcaatattttctgcataaaccctgtttctatggcaaaacccagattgcgcacacttctttaatagatagtcggtaa
acgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttcaaaaggaccatctcgaggtaccgatccgag
acggccggctgggccacgtgaattcttttctcagttgatttgtttgtggggatttagtaagtcgtaaacttcgttaaaaaagatca
atgtagtcaatacagttgatccgaaatagaaggaagaggtttgcaatgtgtaagaacaatgtagttaaaagcccgttttaagac
aatattctttgatgctgatcagaaaaggacaataagggattttggttgcttcttttataccaataatcgtctcctcatcgcttaatttt
ctccccatctcaaccggtgaagggtaggacgcttctgtaatctgttcacataaaaggggttttcactccgagacaaaaatttat
gcgacaaaaatagcctatcttggaaggtgatgtcttatcaacttgcattgtttgcaaggagaagcaaggacaactcaacatgg
gtaaaaattcaaaaccaaccaattggaaactcccaactgtccactaggtagctgacagctgtcacttttgctgttcgttgtcttgt
ctctttcgcttaactgtcctcttacggctttctttcccaaaaaatcattggggaaatgtgcccctcatcagagtccaatgacccat

gaataaagtttcttgtactgtttaagacgatgaattgcaacgataatccgagcagtttacggggtacatcacgtgctttgcatat
gatctcggagtcggatcagttccggatgtgatgtattaccccatagtttcaaactctaatgcagccgccaagtgccatacacc
ctccatcaatctatgcttaaagtttttcaccatcgttgggtggtgatgatgactcgcttagtctctgctgttcgatattaactttgtaa
ggatcgcccttggatggaaaattgaggggttgtaacctgaatttgcaggctacttacattggacttttgagaaggctggacgg
ttgatgaagagggctgggtgcagaggaatggaaaaaaatttagttgagaggactgcttgaaatttttaggaaatggagtccttt
aagctgacaaaacttcaaggatggggattttcatgtagctttttcatgccttcgacaagctaaaggaaggtaattgattctggat
aaatggatatttgatctgctttagcagatgtcaaagttctactagtgatagtctggtatctcgtagccttcaattgggcgtatctta
ctcgaagtgttatatttttagctgacgagacgaagaacgagagagtattgacacattcagaggtaagacaatatgtcgtattat
caaaataagtatcgaacctctattaggagcctactggctcaaatgtgcaaccttagtggtgattgtctctgcttcttgatcacaat
ctgtcgtgtttgagagtgccgatgtatgatttttagtaaatgtttttcagaaaaggcgctaagtaaataaccagtaagtaataaat
aacgtaaaagtgatttgaatcataaaagaatcaagatagaggtcaaagcatagataatccccccgtagaagatggaccggtc
atatggtctgaaaaaaagatctgatctcatg

SEQ ID NO: 18 pGAPZAGLSII

tcgagatggtccttttgaaatggctcgtatgccaattggtcttctttaccgctttttcgcatgcgtttaccgactatctattaaagaa
gtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctattacaaagtgga
cgccgagtctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatga
tatagccgttcagttcccattctctctctttttttacaggatcactcagtaaggttcactataaatgagaaagagagaatgccaa
ccaacagcagcggtttgttgatctcttcacaacggttcaatgagacctggaagtacgcattcgacaagaaatttcaagagga
ggcgaacaggaccagtattccacaattccacttccttaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcattt
ttgtcactttcaaactccactgcagacacatttcatcttcgaaacggtgatgtatccgtagaaatctttgctgaacctttttcaattg
aaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttcctgaacattgaacatcatagaactaagcagga
aaacttcgcacacgtgctgccagaagaaacaactttcaacatgtttaaggacaatttcttgtattcaaagcatgactctatgcctt
tggggcctgaatcggttgcgctagatttctctttcatgggttctactaatgtctacggtataccggaacatgcgacgtcgctaag
gctgatggacacttcaggtggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaacatcggtaccagccaac
caatgtacggttcgatcccattcatgttttcatcttcgtccacatctatcttttgggtcaatgcagctgacacttgggtagacataa
agtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatggtgtcatagatgtagtcatgtccctggggcca
gatattccaactatcattgacaaatttaccgatttgactggtagacccttttaccgcccatttcctctatagggtaccatcaatgta

gatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattccttacgattttatttggttgg
acttggagtatacgaacgacaaaaaatattttacttggaagcagcactcctttcccaatccaaaaaggctgttatccaaattaaa
aaagttgggtagaaatcttgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgacagggtaattaatgaaaat
gtagcagtcaaggatcacaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgataccataagcaaa
tatggccaaaagatttggaagtccttttcgaacggtttatggatctgccggctgatttaactaatttattcatttggaatgatatga
acgagccttcgattttcgatggcccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtc
cataacatatatggtctatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccatccgataagcgtcctttcct
tctaacaagggctttttttgccggctctcaacgtactgctgccacatggactggtgacaatgtggccaattgggattacttaaag
atttccattcctatggttctgtcaaacaacattgctggtatgccatttataggagccgacatagctggctttgctgaggatcctac
acctgaattgattgcacgttggtaccaagcgggcttatggtacccatttttagagcacacgcccatatagacaccaagagaa
gagaaccatacttattcaatgaacctttgaagtcgatagtacgtgatattatccaattgagatatttcctgctacctaccttataca
ccatgtttcataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttgctgaattgtatcatat
cgataaccaattttactggagtaattcaggtctattagtcaaacctgtcacggagcctggtcaatcagaaacggaaatggttttc
ccacccggtatattctatgaattcgcatctttacactcttttataaacaatggtactgatttgatagaaaagaatatttctgcaccat
tggataaaattccattatttattgaaggcggtcacattatcactatgaaagataagtatagaagatcttcaatgttaatgaaaaac
gatccatatgtaatagttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatggagaaacttttggct
accaaagaggtgagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatattcccgagaattt
gacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataatttacaacacaacataac
gttgaaagacagtattaaagtcaaaaaaaatggcgaagaaagttcattgccgactagatcgtcatatgagaatgataataaga
tcaccattcttaacctatcgcttgacataactgaagattgggaagttatttttgggcccgaacaaaaactcatctcagaagagg
atctgaatagcgccgtcgaccatcatcatcatcattgagttttagccttagacatgactgttcctcagttcaagttgggcact
tacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgcctgagagatgcaggcttcatttttgata
ctttttttatttgtaacctatatagtataggatttttttttgtcattttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagct
gatgaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaa
gattaagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatgtttctactccttttttactcttccagattttctc
ggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaaatttccctctttcttcctctagggtgtcgtta
attacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttcttttttcttcgtcgaaaaaggcaataaaaatttttatc
acgtttcttttttcttgaaatttttttttttagttttttttctctttcagtgacctccattgatatttaagttaataaacggtcttcaatttctcaa

gtttcagtttcatttttcttgttctattacaacttttttttacttcttgttcattagaaagaaagcatagcaatctaatctaagggcggtgt
tgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatggccaagttgaccagtg
ccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttctcccgggacttc
gtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggaccaggtggtgccgg
acaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaacttc
cgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgccctgcgcgacccgg
ccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcccaggcctcgga
gatccgtccccctttccttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctccccccacatccgctctaacc
gaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagtattaagaacgttatttatatttca
aatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggttttgggacgctcgaa
ggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttg
ctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca
ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacct
gtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctcc
aagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccgg
taagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagag
ttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcgga
aaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcg
cagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggg
attttggtcatgcatgagatcagatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctgg
ctccgttgcaactccgaacgacctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaa
cgtctcttcccttctctctccttccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccctt
gcagcaatgctcttcccagcattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaag
tcccggccgtcgctggcaataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggc
gaacacctttcccaattttggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactatttc
gaaacgaggaattcacgtggcccagccggccgtctcggatcggtacc

SEQ ID NO: 19 pGAPADE1glsII

tcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcac
tgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatca
ggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggc
gtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact
ataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccg
cctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaaga
cacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttg
aagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaag
agttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaa
aaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttgg
tcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaa
cttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactcc
ccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcac
cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctcca
tccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacagg
catcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccat
gttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatg
gcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgaga
atagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtg
ctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgt
gcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaa
gggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatagctccaaggcaacaaattgactactcagac
cgacattcattcgttattgattttaaatcaacgataaacggaatggttacttgaatgatttcactttatgatcattgtttactaattacc
taaataggattttatatggaattggaagaataagggaaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaag
cccatttcttgccagtaattgcttcataagcttcaatatacttttctttactcttgatagcaatttctgcatccatggctacgccctctt
tgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttacccacttggtaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgttttcgtccagtccaaattcgaatttcgt

atcagcaataatgatccccttcaaaagggcgaagtttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
aaatgtctttacctacaatctcagcagcttgttcaatagagatgttttcatcgtgttcaccctgttcagctttcgttgaaggtgtgaa
aatcggagttggaaaggcgtcgctctcttgaaggttctcgttttcaaccttgactccatggacagtttttgagttcttgtactctttc
catgcacttccagtgatgtaacctctgacaatggcttccaaaggtatcagtctgtgctttttactatcaaggatcgtccctctaat
tgagatttgtatttttcttcagacagttttgatggtagtaaagcaaagacttccttgtcattagaagcaaccaaatgattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatc
gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttcctttggcgatt
aaaggtaggattccgtccagatcagtgttcacaatggacatacttggaaggatacagcaaagtgtgttggaagcgatgacac
atggaaaggaatttttcgagtttcctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttc
aaagcgcccaagtgtttctgccaacctgacttttttctgaataatgaatcgttcaagtggagtatttaaaccatgattaagttacgt
gatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcggtgcgggcctcttcgctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaactcgaatgattttcccaaacccctaccacaagatattcatcagctgcgagat
aggctgatcaggagcaagctcgtacgagaagaaacaaaatgacaaaaaaaatcctatactatataggttacaaataaaaaa
gtatcaaaaatgaagcctgcatctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgt
aagtgcccaacttgaactgaggaacagtcatgtctaaggctaaaactcaatgatgatgatgatgatggtcgacggcgctattc
agatcctcttctgagatgagttttgttcgggcccaaaaataacttcccaatcttcagttatgtcaagcgataggttaagaatggt
gatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccatttttttttgactttaatactgtctttcaacgttatgt
tgtgttgtaaattattctttgcgataatgatttttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcgggaa
tatgacttcgaacatttttttaaggtattgttttcgaaaatgaactgagtttctacgtactcacctctttggtagccaaaagtttctcca
tcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgtttttcattaacattga
agatcttctatacttatctttcatagtgataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattcttttt
ctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgtttct
gattgaccaggctccgtgacaggtttgactaatagacctgaattactccagtaaaattggttatcgatatgatacaattcagcaa
attcagggtgttcaataaacattggattcattatcggaaatccagtgacacttgatttatgaaacatggtgtataaggtaggtag
caggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcattgaataagtatggttctcttctcttggtgtctata
tgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattcaggtgtaggatcctcagcaa

agccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaaccataggaatggaaatctttaagtaatccc
aattggccacattgtcaccagtccatgtggcagcagtacgttgagagccggcaaaaaaagcccttgttagaaggaaaggac
gcttatcggatggtgaataaatcgattttattgcgtcgtaagtagcttcatgcactgatagaccatatatgttatggacggatcttt
cctcaatgtaattgtcgtgaatcaaatcttttggagctgtggtctctgggccatcgaaaatcgaaggctcgttcatatcattccaa
atgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaaggacttccaaatcttttggccatatttgcttatggt
atcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttgactgctacattttcattaattacc
ctgtcactgatttcataatctttctttaaatgaggatcgattagtacgacaagatttctacccaacttttttaatttggataacagcct
ttttggattgggaaaggagtgctgcttccaagtaaaatatttttttgtcgttcgtatactccaagtccaaccaaataaaatcgtaag
gaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccatctacattgatggtaccctata
gaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgtcaatgatagttggaatatctggccccagggaca
tgactacatctatgacaccattttcggagatccaatgagtcatcgttttatttttactggtgtcatactttatgtctacccaagtgtca
gctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggctggtaccg
atgttgtactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtccatcagccttagcgacgtcgca
tgttccggtataccgtagacattagtagaacccatgaaagagaaatctagcgcaaccgattcaggccccaaaggcatagag
tcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgtttcttctggcagcacgtgtgcgaagttttcctgcttagttc
tatgatgttcaatgttcaggaaattttgctcgtttacaataagtttcagcgcatttgccagtaaactttcaattgaaaaggttcagc
aaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtttgaaagtgacaaaaatgaagatattttc
gaccagaatgagttcacagtttgtttttgcttaaggaagtggaattgtggaatactggtcctgttcgcctcctcttgaaatttcttgt
cgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttggttggcattctctctttctcattt
atagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaactgaacggctatatcatcgccctccaatcttg
gtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagtg
atgagatttggcaatattttctgcataaaccctgtttctatggcaaaacccagattgcgcacacttctttaatagatagtcggtaa
acgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttcaaaaggaccatctcgaggtaccgatccgag
acggccggctgggccacgtgaattcctcgtttcgaaatagttgttcaattgattgaaatagggacaaataaattaaatttaaagt
ctttgggtcaggagaaaccaaaattgggaaaggtgttcgcctttatattcgattctggtggtttccaataatctcatgacatgcg
tccgcccgctattattgccagcgacggccgggacttttccatccctgggctgctaggtcgggtacacgacctccgttttaccc
gcaacgtaatgctgggaagagcattgctgcaaggggccgtagaagaagctctccagcagagtaaaatttcctagggacg
gtaacgggcggtggaaggagagagaagggaagagacgtttctggttccattactccacatttaagttttacccccggagaattt

tacgttgccagcaggtcgttcggagttgcaacggagccagatatttcagagatggctacctgattggacgaggacaccaag
acatttctacaaaaaagatctgatctca

SEQ ID NO: 20 pPICZAGLSII

cgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttgtagccttagac
atgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccattt
gcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcattttgtttcttctcgtacga
gcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgtttttcttg
gtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatg
tttctactccttttttactcttccagattttctcggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaa
attttccctctttcttcctctagggtgtcgttaattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttcttttc
ttcgtcgaaaaaggcaataaaaatttttatcacgtttcttttcttgaaatttttttttttagttttttctctttcagtgacctccattgatat
ttaagttaataaacggtcttcaatttctcaagtttcagtttcatttttcttgttctattacaacttttttacttcttgttcattagaaagaaa
gcatagcaatctaatctaaggggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgag
gaactaaaccatggccaagttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctgga
ccgaccggctcgggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatca
gcgcggtccaggaccaggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccg
agtggtcggaggtcgtgtccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtggg
ggcgggagttcgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgac
ggcggcccacgggtcccaggcctcggagatccgtccccctttccttgtcgatatcatgtaattagttatgtcacgcttacatt
cacgccctcccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtcccta ttatttttttatagt
tatgttagtattaagaacgttatttatatttcaaatttttctttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaacc
ttgcttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaag
gccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgc
tcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcct
gttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtagg
tatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatc
cggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag

agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatct
gcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtg
gttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgc
tcagtggaacgaaaactcacgttaagggattttggtcatgagatcagatctaacatccaaagacgaaaggttgaatgaaacct
ttttgccatccgacatccacaggtccattctcacacataagtgccaaacgcaacaggaggggatacactagcagcagaccg
ttgcaaacgcaggacctccactcctcttctcctcaacacccacttttgccatcgaaaaaccagcccagttattgggcttgattg
gagctcgctcattccaattccttctattaggctactaacaccatgactttattagcctgtctatcctggcccccctggcgaggttc
atgtttgtttatttccgaatgcaacaagctccgcattacacccgaacatcactccagatgagggctttctgagtgtggggtcaa
atagtttcatgttccccaaatggcccaaaactgacagtttaaacgctgtcttggaacctaatatgacaaaagcgtgatctcatcc
aagatgaactaagtttggttcgttgaaatgctaacggccagttggtcaaaaagaaacttccaaaagtcggcataccgtttgtct
tgtttggtattgattgacgaatgctcaaaaataatctcattaatgcttagcgcagtctctctatcgcttctgaaccccggtgcacct
gtgccgaaacgcaaatggggaaacacccgctttttggatgattatgcattgtctccacattgtatgcttccaagattctggtgg
gaatactgctgatagcctaacgttcatgatcaaaatttaactgttctaacccctacttgacagcaatatataaacagaaggaag
ctgccctgtcttaaacctttttttttatcatcattattagcttactttcataattgcgactggttccaattgacaagcttttgattttaacg
acttttaacgacaacttgagaagatcaaaaaacaactaattattcgaaacgaggaattcacgtggcccagccggccgtctcg
gatcggtacctcgagatggtccttttgaaatggctcgtatgccaattggtcttctttaccgcttttcgcatgcgtttaccgactat
ctattaaagaagtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctatt
acaaagtggacgccgagtctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattgg
agggcgatgatatagccgttcagttcccattctctctctttttacaggatcactcagtaaggttcactataaatgagaaagag
agaatgccaaccaacagcagcggtttgttgatctcttcacaacggttcaatgagacctggaagtacgcattcgacaagaaatt
tcaagaggaggcgaacaggaccagtattccacaattccacttccttaagcaaaaacaaactgtgaactcattctggtcgaaa
atatcttcatttttgtcactttcaaactccactgcagacacatttcatcttcgaaacggtgatgtatccgtagaaatctttgctgaac
cttttcaattgaaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttcctgaacattgaacatcatagaact
aagcaggaaaacttcgcacacgtgctgccagaagaaacaactttcaacatgtttaaggacaatttcttgtattcaaagcatga
ctctatgcctttggggcctgaatcggttgcgctagatttctctttcatgggttctactaatgtctacggtataccggaacatgcga
cgtcgctaaggctgatggacacttcaggtggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaacatcggt
accagccaaccaatgtacggttcgatcccattcatgttttcatcttcgtccacatctatcttttgggtcaatgcagctgacacttg
ggtagacataaagtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatggtgtcatagatgtagtcatgt

ccctggggccagatattccaactatcattgacaaatttaccgatttgactggtagacccttttttaccgcccatttcctctataggg
taccatcaatgtagatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattccttacg
attttatttggttggacttggagtatacgaacgacaaaaaatattttacttggaagcagcactcctttcccaatccaaaaaggctg
ttatccaaattaaaaaagttgggtagaaatcttgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgacagggt
aattaatgaaaatgtagcagtcaaggatcacaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgat
accataagcaaatatggccaaaagatttggaagtccttttttcgaacggtttatggatctgccggctgatttaactaatttattcatt
tggaatgatatgaacgagccttcgattttcgatggcccagagaccacagctccaaaagatttgattcacgacaattacattga
ggaaagatccgtccataacatatatggtctatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccatccgat
aagcgtcctttccttctaacaagggcttttttgccggctctcaacgtactgctgccacatggactggtgacaatgtggccaatt
gggattacttaaagatttccattcctatggttctgtcaaacaacattgctggtatgccatttataggagccgacatagctggcttt
gctgaggatcctacacctgaattgattgcacgttggtaccaagcgggcttatggtacccatttttttagagcacacgcccatata
gacaccaagagaagagaaccatacttattcaatgaacctttgaagtcgatagtacgtgatattatccaattgagatatttcctgc
tacctaccttatacaccatgtttcataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttgc
tgaattgtatcatatcgataaccaattttactggagtaattcaggtctattagtcaaacctgtcacggagcctggtcaatcagaaa
cggaaatggttttcccacccggtatattctatgaattcgcatctttacactcttttataaacaatggtactgatttgatagaaaaga
atatttctgcaccattggataaaattccattatttattgaaggcggtcacattatcactatgaaagataagtatagaagatcttcaa
tgttaatgaaaaacgatccatatgtaatagttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatg
gagaaacttttggctaccaaagaggtgagtacgtagaaactcagttcattttcgaaaacaatacctttaaaaaatgttcgaagtc
atattcccgagaatttgacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataattt
acaacacaacataacgttgaaagacagtattaaagtcaaaaaaaatggcgaagaaagttcattgccgactagatcgtcatat
gagaatgataataagatcaccattcttaacctatcgcttgacataactgaagattgggaagttatttttgggcc

SEQ ID NO: 21 pPICADE1glsII

aaattcctcgtttcgaataattagttgttttttgatcttctcaagttgtcgttaaaagtcgttaaaatcaaaagcttgtcaattggaac
cagtcgcaattatgaaagtaagctaataatgatgataaaaaaaaaaggtttaagacagggcagcttccttctgtttatatattgct
gtcaagtaggggttagaacagttaaattttgatcatgaacgttaggctatcagcagtattcccaccagaatcttggaagcatac
aatgtggagacaatgcataatcatccaaaaagcgggtgtttccccatttgcgtttcggcacaggtgcaccggggttcagaag
cgatagagagactgcgctaagcattaatgagattatttttgagcattcgtcaatcaataccaaacaagacaaacggtatgccg

actttggaagtttcttttgaccaactggccgttagcatttcaacgaaccaaacttagttcatcttggatgagatcacgcttttgtc
atattaggttccaagacagcgtttaaactgtcagttttgggccatttggggaacatgaaactatttgaccccacactcagaaag
ccctcatctggagtgatgttcgggtgtaatgcggagcttgttgcattcggaaataaacaaacatgaacctcgccaggggggc
caggatagacaggctaataaagtcatggtgttagtagcctaatagaaggaattggaatgagcgagctccaatcaagcccaat
aactgggctggttttcgatggcaaaagtgggtgttgaggagaagaggagtggaggtcctgcgtttgcaacggtctgctgct
agtgtatcccctcctgttgcgtttggcacttatgtgtgagaatggacctgtggatgtcggatggcaaaaaggtttcattcaacct
ttcgtctttggatgttgtcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctctt
ccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacg
gttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgctt
accggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtag
gtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttga
gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg
gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagcca
gttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcag
cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactc
acgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatcta
aagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatc
catagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccg
cgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctg
caactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgtt
gttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcga
gttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtg
ttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactca
accaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacata
gcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagtt
cgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggc

aaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttcaatagctccaaggcaac
aaattgactactcagaccgacattcattcgttattgattttaaatcaacgataaacggaatggttacttgaatgatttcactttatga
tcattgtttactaattacctaaataggattttatatggaattggaagaataagggaaatttcagatgtctgaaaaaggcgaggag
ggtactaatcattcaagcccatttcttgccagtaattgcttcataagcttcaatatacttttctttactcttgatagcaatttctgcatc
catggctacgccctctttgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttac
ccacttggtaagtcttttgattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgttttcgtccag
tccaaattcgaatttcgtatcagcaataatgatccccttcaaaagggcgaagtttttgcagcagaatacaactcgaccgccttg
acagcgaccttctcacaaatgtctttacctacaatctcagcagcttgttcaatagagatgttttcatcgtgttcaccctgttcagct
ttcgttgaaggtgtgaaaatcggagttggaaaggcgtcgctctcttgaaggttctcgttttcaaccttgactccatggacagtttt
tgagttcttgtactctttccatgcacttccagtgatgtaacctctgacaatggcttccaaaggtatcagtctgtgctttttactatca
aggatcgtccctctaattgagatttgtatttttcttcagacagttttgatggtagtaaagcaaagacttccttgtcattagaagcaa
ccaaatgattctttatgtagggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaatac
cgtttgtcataatcacatcgtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctc
taacctttcctttggcgattaaaggtaggattccgtccagatcagtgttcacaatggacatacttggaaggatacagcaaagtg
tgttggaagcgatgacacatggaaaggaatttttcgagtttcctagagtagtatattggggcggtgaaagttcagatgtttaatg
cttaatactcttatactcttcaaagcgcccaagtgtttctgccaacctgacttttttctgaataatgaatcgttcaagtggagtattta
aaccatgattaagttacgtgatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaa
ataccactaattacggagaagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgc
gcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagggcctcgaggcacaaacgaacgtctcactta
atcttctgtactctgaagaggagtgggaaataccaagaaaaacatcaaactcgaatgattttcccaaacccctaccacaagat
attcatcagctgcgagataggctgatcaggagcaagctcgtacgagaagaaacaaaatgacaaaaaaaatcctatactatat
aggttacaaataaaaaagtatcaaaaatgaagcctgcatctctcaggcaaatggcattctgacatcctcttgattagaatctag
caagaccggtcttctcgtaagtgcccaacttgaactgaggaacagtcatgtctaaggctacaaactcaatgatgatgatgatg
atggtcgacggcgctattcagatcctcttctgagatgagttttgttcgggcccaaaaataacttcccaatcttcagttatgtcaa
gcgataggttaagaatggtgatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccatttttttgacttta
atactgtctttcaacgttatgttgtgttgtaaattattctttgcgataatgatttttcaatattggtattcctcaaagtattgtggtgaat
gcctgtcaaattctcgggaatatgacttcgaacatttttaaggtattgttttcgaaaatgaactgagtttctacgtactcacctctt
ggtagccaaaagtttctccatcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatg

gatcgtttttcattaacattgaagatcttctatacttatctttcatagtgataatgtgaccgccttcaataaataatggaattttatcca
atggtgcagaaatattcttttctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggt
gggaaaaccatttccgtttctgattgaccaggctccgtgacaggtttgactaatagacctgaattactccagtaaaattggttat
cgatatgatacaattcagcaaattcagggtgttcaataaacattggattcattatcggaaatccagtgacacttgatttatgaaac
atggtgtataaggtaggtagcaggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcattgaataagtatg
gttctcttctcttggtgtctatatgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattca
ggtgtaggatcctcagcaaagccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaaccataggaat
ggaaatctttaagtaatcccaattggccacattgtcaccagtccatgtggcagcagtacgttgagagccggcaaaaaaagcc
cttgttagaaggaaaggacgcttatcggatggtgaataaatcgattttattgcgtcgtaagtagcttcatgcactgatagaccat
atatgttatggacggatctttcctcaatgtaattgtcgtgaatcaaatcttttggagctgtggtctctgggccatcgaaaatcgaa
ggctcgttcatatcattccaaatgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaaggacttccaaatc
ttttggccatatttgcttatggtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttga
ctgctacattttcattaattaccctgtcactgatttcataatctttctttaaatgaggatcgattagtacgacaagatttctacccaac
ttttttaatttggataacagccttttggattgggaaaggagtgctgcttccaagtaaaatatttttgtcgttcgtatactccaagtc
caaccaaataaaatcgtaaggaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccat
ctacattgatggtaccctatagaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgtcaatgatagttgg
aatatctggccccagggacatgactacatctatgacaccattttcggagatccaatgagtcatcgttttatttttactggtgtcata
ctttatgtctacccaagtgtcagctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccg
tacattggttggctggtaccgatgttgtactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtcca
tcagccttagcgacgtcgcatgttccggtataccgtagacattagtagaacccatgaaagagaaatctagcgcaaccgattc
aggccccaaaggcatagagtcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgtttcttctggcagcacgtgt
gcgaagttttcctgcttagttctatgatgttcaatgttcaggaaattttgctcgtttacaataagtttcagcgcattttgccagtaaa
ctttcaattgaaaaggttcagcaaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtttgaaag
tgacaaaaatgaagatattttcgaccagaatgagttcacagtttgtttttgcttaaggaagtggaattgtggaatactggtcctgt
tcgcctcctcttgaaatttcttgtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttg
gttggcattctctctttctcatttatagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaactgaacggcta
tatcatcgccctccaatcttggtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcg
tccactttgtaatagcagtgatgagatttggcaatattttctgcataaaccctgtttctatggcaaaacccagattgcgcacactt

ctttaatagatagtcggtaaacgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttcaaaaggaccatc
tcg

SEQ ID NO: 22 pYPT1ZAGLSII

cgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttgtagccttagac
atgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccattt
gcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcattttgtttcttctcgtacga
gcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgttttcttg
gtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatg
tttctactccttttttactcttccagattttctcggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaa
attttccctctttcttcctctagggtgtcgttaattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttctttttc
ttcgtcgaaaaaggcaataaaaatttttatcacgtttctttttcttgaaatttttttttttagttttttttctctttcagtgacctccattgatat
ttaagttaataaacggtcttcaatttctcaagtttcagtttcatttttcttgttctattacaacttttttttacttcttgttcattagaaagaaa
gcatagcaatctaatctaaggggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgag
gaactaaaccatggccaagttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctgga
ccgaccggctcgggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatca
gcgcggtccaggaccaggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccg
agtggtcggaggtcgtgtccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtggg
ggcgggagttcgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgac
ggcggcccacgggtcccaggcctcggagatccgtccccctttccctttgtcgatatcatgtaattagttatgtcacgcttacatt
cacgccctccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagt
tatgttagtattaagaacgttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaacc
ttgcttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaag
gccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgc
tcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcct
gttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtagg
tatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatc
cggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag

agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatct
gcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtg
gttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgc
tcagtggaacgaaaactcacgttaagggattttggtcatgagatcagatctatgatgagtcacaatctgcttccacagacgagt
acaaggacaggcaaaaggaattggaagaagttgctaacccaataatgagcaagttctatggagctgctggtggagctcctg
gtggagctcctggtggcttccctggaggtttccctggcggagctggcgcagctggcggtgccccaggtggtgctgcccca
ggcggagacagcggaccaaccgtggaagaagtcgattaagcaattcaacggataaattctggttaatatatataacgtgaat
aggaaattaaggaaattttggatctaataatgtgctgtatgccgacatcgggcatcgtagattgtatagtatcgctgacactata
ataagccagccaaaaccctaaaccagttgccctccactaattagtgtactacccaatcttgcctcttcgggtgtcttttataag
gacagattcacaagctcttgttgcccaatacacacatacacacagagataatagcagtcgaattcacgtggcccagccggc
cgtctcggatcggtacctcgagatggtccttttgaaatggctcgtatgccaattggtcttctttaccgcttttcgcatgcgtttac
cgactatctattaaagaagtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcac
tgctattacaaagtggacgccgagtctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaa
gattggagggcgatgatatagccgttcagttcccattctctctctttttttacaggatcactcagtaaggttcactataaatgag
aaagagagaatgccaaccaacagcagcggtttgttgatctcttcacaacggttcaatgagacctggaagtacgcattcgaca
agaaatttcaagaggaggcgaacaggaccagtattccacaattccacttccttaagcaaaaacaaactgtgaactcattctgg
tcgaaaatatcttcatttttgtcactttcaaactccactgcagacacatttcatcttcgaaacggtgatgtatccgtagaaatctttg
ctgaacctttttcaattgaaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttcctgaacattgaacatcat
agaactaagcaggaaaacttcgcacacgtgctgccagaagaaacaactttcaacatgtttaaggacaatttcttgtattcaaa
gcatgactctatgcctttggggcctgaatcggttgcgctagatttctctttcatgggttctactaatgtctacggtataccggaac
atgcgacgtcgctaaggctgatggacacttcaggtggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaac
atcggtaccagccaaccaatgtacggttcgatcccattcatgttttcatcttcgtccacatctatcttttgggtcaatgcagctga
cacttgggtagacataaagtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatggtgtcatagatgtag
tcatgtccctggggccagatattccaactatcattgacaaatttaccgatttgactggtagacccttttaccgcccatttcctcta
tagggtaccatcaatgtagatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattcc
ttacgattttatttggttggacttggagtatacgaacgacaaaaaatattttacttggaagcagcactcctttcccaatccaaaaa
ggctgttatccaaattaaaaaagttgggtagaaatcttgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgac
agggtaattaatgaaaatgtagcagtcaaggatcacaatggaaatgactatgtaggtcattgctggccaggtaattctatatgg

attgataccataagcaaatatggccaaaagatttggaagtccttttcgaacggtttatggatctgccggctgatttaactaattta
ttcatttggaatgatatgaacgagccttcgattttcgatggcccagagaccacagctccaaaagatttgattcacgacaattac
attgaggaaagatccgtccataacatatatggtctatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccat
ccgataagcgtcctttccttctaacaagggcttttttgccggctctcaacgtactgctgccacatggactggtgacaatgtggc
caattgggattacttaaagatttccattcctatggttctgtcaaacaacattgctggtatgccatttataggagccgacatagctg
gctttgctgaggatcctacacctgaattgattgcacgttggtaccaagcgggcttatggtacccatttttagagcacacgccc
atatagacaccaagagaagagaaccatacttattcaatgaacctttgaagtcgatagtacgtgatattatccaattgagatattt
cctgctacctaccttatacaccatgtttcataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctga
atttgctgaattgtatcatatcgataaccaattttactggagtaattcaggtctattagtcaaacctgtcacggagcctggtcaatc
agaaacggaaatggttttcccacccggtatattctatgaattcgcatctttacactcttttataaacaatggtactgatttgataga
aaagaatatttctgcaccattggataaaattccattatttattgaaggcggtcacattatcactatgaaagataagtatagaagat
cttcaatgttaatgaaaaacgatccatatgtaatagttatagcccctgataccgagggacgagccgttggagatctttatgttga
tgatggagaaacttttggctaccaaagaggtgagtacgtagaaactcagttcattttcgaaaacaatacattaaaaaatgttcg
aagtcatattcccgagaatttgacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaat
aatttacaacacaacataacgttgaaagacagtattaaagtcaaaaaaaatggcgaagaaagttcattgccgactagatcgtc
atatgagaatgataataagatcaccattcttaacctatcgcttgacataactgaagattgggaagttatttttgggcc

SEQ ID NO: 23 pYPT1ADE1glsII

gtcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctca
ctgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc
aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctgg
cgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagga
ctataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtcc
gcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagc
tgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaag
acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctt
gaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcaga

aaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttg
gtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaa
acttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactc
cccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctca
ccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctcc
atccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacag
gcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccca
tgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttat
ggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgag
aatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagt
gctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcg
tgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaa
agggaataagggcgacacggaaatgttgaatactcatactcttccttttcaatagctccaaggcaacaaattgactactcaga
ccgacattcattcgttattgattttaaatcaacgataaacggaatggttacttgaatgatttcactttatgatcattgtttactaattac
ctaaataggattttatatggaattggaagaataagggaaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaa
gcccatttcttgccagtaattgcttcataagcttcaatatacttttctttactcttgatagcaatttctgcatccatggctacgccctct
ttgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttacccacttggtaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgttttcgtccagtccaaattcgaatttcgt
atcagcaataatgatccccttcaaaagggcgaagttttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
aaatgtctttacctacaatctcagcagcttgttcaatagagatgttttcatcgtgttcaccctgttcagctttcgttgaaggtgtgaa
aatcggagttggaaaggcgtcgctctcttgaaggttctcgttttcaaccttgactccatggacagtttttgagttcttgtactctttc
catgcacttccagtgatgtaacctctgacaatggcttccaaaggtatcagtctgtgctttttttactatcaaggatcgtccctctaat
tgagatttgtatttttcttcagacagttttgatggtagtaaagcaaagacttccttgtcattagaagcaaccaaatgattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatc
gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttccttggcgatt
aaaggtaggattccgtccagatcagtgttcacaatggacatacttggaaggatacagcaaagtgtgttggaagcgatgacac
atggaaaggaatttttcgagtttcctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttc
aaagcgcccaagtgtttctgccaacctgacttttttctgaataatgaatcgttcaagtggagtatttaaaccatgattaagttacgt

gatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcggtgcgggcctcttcgctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaactcgaatgattttcccaaacccctaccacaagatattcatcagctgcgagat
aggctgatcaggagcaagctcgtacgagaagaaacaaaatgacaaaaaaaatcctatactatataggttacaaataaaaaa
gtatcaaaaatgaagcctgcatctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgt
aagtgcccaacttgaactgaggaacagtcatgtctaaggctacaaactcaatgatgatgatgatgatggtcgacggcgctatt
cagatcctcttctgagatgagttttgttcgggcccaaaaataacttcccaatcttcagttatgtcaagcgataggttaagaatgg
tgatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccatttttttgactttaatactgtctttcaacgttatg
ttgtgttgtaaattattctttgcgataatgatttttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcggga
atatgacttcgaacatttttttaaggtattgttttcgaaaatgaactgagtttctacgtactcacctctttggtagccaaaagtttctcc
atcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgtttttcattaacattg
aagatcttctatacttatctttcatagtgataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattcttt
tctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgtttc
tgattgaccaggctccgtgacaggtttgactaatagacctgaattactccagtaaaattggttatcgatatgatacaattcagca
aattcagggtgttcaataaacattggattcattatcggaaatccagtgacacttgatttatgaaacatggtgtataaggtaggta
gcaggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcattgaataagtatggttctcttctcttggtgtctat
atgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattcaggtgtaggatcctcagca
aagccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaaccataggaatggaaatctttaagtaatcc
caattggccacattgtcaccagtccatgtggcagcagtacgttgagagccggcaaaaaaagcccttgttagaaggaaaggga
cgcttatcggatggtgaataaatcgattttattgcgtcgtaagtagcttcatgcactgatagaccatatatgttatggacggatctt
tcctcaatgtaattgtcgtgaatcaaatcttttggagctgtggtctctgggccatcgaaaatcgaaggctcgttcatatcattcca
aatgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaaggacttccaaatcttttggccatatttgcttatg
gtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttgactgctacattttcattaatta
ccctgtcactgatttcataatctttctttaaatgaggatcgattagtacgacaagatttctacccaacttttttaatttggataacagc
ctttttggattgggaaaggagtgctgcttccaagtaaaatatttttttgtcgttcgtatactccaagtccaaccaaataaaatcgtaa
ggaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccatctacattgatggtaccctat
agaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgtcaatgatagttggaatatctggccccagggac

atgactacatctatgacaccattttcggagatccaatgagtcatcgttttatttttactggtgtcatactttatgtctacccaagtgtc
agctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggctggtacc
gatgttgtactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtccatcagccttagcgacgtcgc
atgttccggtataccgtagacattagtagaacccatgaaagagaaatctagcgcaaccgattcaggccccaaaggcataga
gtcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgtttcttctggcagcacgtgtgcgaagttttcctgcttagtt
ctatgatgttcaatgttcaggaaattttgctcgtttacaataagtttcagcgcattttgccagtaaactttcaattgaaaaggttcag
caaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtttgaaagtgacaaaaatgaagatatttt
cgaccagaatgagttcacagtttgtttttgcttaaggaagtggaattgtggaatactggtcctgttcgcctcctcttgaaatttctt
gtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttggttggcattctctctttctcat
ttatagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaactgaacggctatatcatcgccctccaatctt
ggtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagt
gatgagatttggcaatattttctgcataaaccctgtttctatggcaaaacccagattgcgcacacttctttaatagatagtcggta
aacgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttcaaaaggaccatctcgaggtaccgatccga
gacggccggctgggccacgtgaattcgactgctattatctctgtgtgtatgtgtgtattgggcaacaagagcttgtgaatctgt
ccttataaaagacacccgaagaggcaagattgggtagtacactaattagtggagggcaactggtttaggggttttggctggc
ttattatagtgtcagcgatactatacaatctacgatgcccgatgtcggcatacagcacattattagatccaaaatttccttaatttc
ctattcacgttatatatattaaccagaatttatccgttgaattgcttaatcgacttcttccacggttggtccgctgtctccgcctggg
gcagcaccacctggggcaccgccagctgcgccagctccgccagggaaacctccagggaagccaccaggagctccacc
aggagctccaccagcagctccatagaacttgctcattattgggttagcaacttcttccaattccttttgcctgtccttgtactcgtc
tgtggaagcagattgtgactcatcatagatctgatctcat

SEQ ID NO: 24 pGAPZAglsIIHDEL

tcgagatggtccttttgaaatggctcgtatgccaattggtcttctttaccgcttttttcgcatgcgtttaccgactatctattaaagaa
gtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctattacaaagtgga
cgccgagtctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatga
tatagccgttcagttcccattctctctctttttttacaggatcactcagtaaggttcactataaatgagaaagagagaatgccaa
ccaacagcagcggtttgttgatctcttcacaacggttcaatgagacctggaagtacgcattcgacaagaaatttcaagagga
ggcgaacaggaccagtattccacaattccacttccttaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcattt

ttgtcactttcaaactccactgcagacacatttcatcttcgaaacggtgatgtatccgtagaaatctttgctgaacctttcaattg
aaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttcctgaacattgaacatcatagaactaagcagga
aaacttcgcacacgtgctgccagaagaaacaactttcaacatgtttaaggacaatttcttgtattcaaagcatgactctatgcctt
tggggcctgaatcggttgcgctagatttctctttcatgggttctactaatgtctacggtataccggaacatgcgacgtcgctaag
gctgatggacacttcaggtggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaacatcggtaccagccaac
caatgtacggttcgatcccattcatgttttcatcttcgtccacatctatcttttgggtcaatgcagctgacacttgggtagacataa
agtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatggtgtcatagatgtagtcatgtccctggggcca
gatattccaactatcattgacaaatttaccgatttgactggtagacccttttttaccgcccatttcctctatagggtaccatcaatgta
gatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattccttacgattttatttggttgg
acttggagtatacgaacgacaaaaaatattttacttggaagcagcactcctttcccaatccaaaaaggctgttatccaaattaaa
aaagttgggtagaaatcttgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgacagggtaattaatgaaaat
gtagcagtcaaggatcacaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgataccataagcaaa
tatggccaaaagatttggaagtccttttttcgaacggtttatggatctgccggctgatttaactaatttattcatttggaatgatatga
acgagccttcgattttcgatggcccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtc
cataacatatatggtctatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccatccgataagcgtccttttcct
tctaacaagggctttttttgccggctctcaacgtactgctgccacatggactggtgacaatgtggccaattgggattacttaaag
atttccattcctatggttctgtcaaacaacattgctggtatgccatttataggagccgacatagctggctttgctgaggatcctac
acctgaattgattgcacgttggtaccaagcgggcttatggtacccatttttttagagcacacgcccatatagacaccaagagaa
gagaaccatacttattcaatgaacctttgaagtcgatagtacgtgatattatccaattgagatatttcctgctacctaccttataca
ccatgtttcataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttgctgaattgtatcatat
cgataaccaattttactggagtaattcaggtctattagtcaaacctgtcacggagcctggtcaatcagaaacggaaatggttttc
ccacccggtatattctatgaattcgcatctttacactcttttataaacaatggtactgatttgatagaaaagaatatttctgcaccat
tggataaaattccattatttattgaaggcggtcacattatcactatgaaagataagtatagaagatcttcaatgttaatgaaaaac
gatccatatgtaatagttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatggagaaacttttggct
accaaagaggtgagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatattcccgagaattt
gacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataatttacaacacaacataac
gttgaaagacagtattaaagtcaaaaaaaatggcgaagaaagttcattgccgactagatcgtcatatgagaatgataataaga
tcaccattcttaacctatcgcttgacataactgaagattgggaagttatttttgggcccgaacaaaaactcatctcagaagagg

atctgaatagcgccgtcgaccacgacgaactgtgagttttagccttagacatgactgttcctcagttcaagttgggcacttacg
agaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgcctgagagatgcaggcttcatttttgatactttt
ttatttgtaacctatatagtataggatttttttgtcattttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgat
gaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagatt
aagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatgtttctactccttttttactcttccagattttctcgga
ctccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaaattttccctctttcttcctctagggtgtcgttaatta
cccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttctttttcttcgtcgaaaaaggcaataaaaatttttatcacg
tttctttttcttgaaatttttttttttagttttttctctttcagtgacctccattgatatttaagttaataaacggtcttcaatttctcaagtttc
agtttcatttttcttgttctattacaacttttttacttcttgttcattagaaagaaagcatagcaatctaatctaagggcggtgttgac
aattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatggccaagttgaccagtgccgt
tccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttctcccgggacttcgtgg
aggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggaccaggtggtgccggacaa
caccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaacttccgg
gacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgccctgcgcgacccggccg
gcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcccaggcctcggagat
ccgtcccccttttccttttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctcccccccacatccgctctaaccgaa
aaggaaggagttagacaacctgaagtctaggtcccctatttatttttttatagttatgttagtattaagaacgttatttatatttcaaatt
tttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggttttgggacgctcgaaggct
ttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctgg
cgtttttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagga
ctataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtcc
gcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagc
tgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaag
acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctt
gaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcaga
aaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttg
gtcatgcatgagatcagatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgt

tgcaactccgaacgacctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctc
ttcccttctctctccttccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccctgcagc
aatgctcttcccagcattacgttgcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccg
gccgtcgctggcaataatagcgggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaac
accttcccaattttggtttctcctgacccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactatttcgaaa
cgaggaattcacgtggcccagccggccgtctcggatcggtacc

SEQ ID NO: 25 pGAPADE1glsIIHDEL

cgtactcacctctttggtagccaaaagtttctccatcatcaacataaagatctccaacggctcgtccctcggtatcaggggctat
aactattacatatggatcgtttttcattaacattgaagatcttctatacttatctttcatagtgataatgtgaccgccttcaataaataa
tggaatttttatccaatggtgcagaaatattcttttctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatag
aatataccgggtgggaaaaccatttccgtttctgattgaccaggctccgtgacaggtttgactaatagacctgaattactccag
taaaattggttatcgatatgatacaattcagcaaattcagggtgttcaataaacattggattcattatcggaaatccagtgacact
tgatttatgaaacatggtgtataaggtaggtagcaggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcat
tgaataagtatggttctcttctcttggtgtctatatgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgt
gcaatcaattcaggtgtaggatcctcagcaaagccagctatgtcggctcctataaatggcataccagcaatgttgtttgacag
aaccataggaatggaaatctttaagtaatcccaattggccacattgtcaccagtccatgtggcagcagtacgttgagagccg
gcaaaaaaagcccttgttagaaggaaaggacgcttatcggatggtgaataaatcgattttattgcgtcgtaagtagcttcatgc
actgatagaccatatatgttatggacggatctttcctcaatgtaattgtcgtgaatcaaatcttttggagctgtggtctctgggcca
tcgaaaatcgaaggctcgttcatatcattccaaatgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaag
gacttccaaatcttttggccatatttgcttatggtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccat
tgtgatccttgactgctacattttcattaattaccctgtcactgatttcataatctttctttaaatgaggatcgattagtacgacaaga
tttctacccaacttttttaatttggataacagccttttggattgggaaaggagtgctgcttccaagtaaaatatttttttgtcgttcgta
tactccaagtccaaccaaataaaatcgtaaggaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatc
attataattccatctacattgatggtaccctatagaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgtca
atgatagttggaatatctggccccagggacatgactacatctatgacaccattttcggagatccaatgagtcatcgttttatttttta
ctggtgtcatactttatgtctacccaagtgtcagctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgg
gatcgaaccgtacattggttggctggtaccgatgttgtactcaaagacatcaacgttgaaaagcctgtagggttcctttccacc

tgaagtgtccatcagccttagcgacgtcgcatgttccggtataccgtagacattagtagaacccatgaaagagaaatctagc
gcaaccgattcaggccccaaaggcatagagtcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgtttcttctg
gcagcacgtgtgcgaagttttcctgcttagttctatgatgttcaatgttcaggaaattttgctcgtttacaataagtttcagcgcatt
ttgccagtaaactttcaattgaaaaggttcagcaaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtg
gagtttgaaagtgacaaaaatgaagatattttcgaccagaatgagttcacagtttgtttttgcttaaggaagtggaattgtggaat
actggtcctgttcgcctcctcttgaaatttcttgtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaac
cgctgctgttggttggcattctctctttctcatttatagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaa
ctgaacggctatatcatcgccctccaatcttggtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaa
tagactcggcgtccactttgtaatagcagtgatgagatttggcaatattttctgcataaaccctgtttctatggcaaaacccagat
tgcgcacacttctttaatagatagtcggtaaacgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttca
aaaggaccatctcgaggtaccgatccgagacggccggctgggccacgtgaattcctcgtttcgaaatagttgttcaattgatt
gaaatagggacaaataaattaaatttaaagtctttgggtcaggagaaaccaaaattgggaaaggtgttcgccttttatattcgat
tctggtggtttccaataatctcatgacatgcgtccgcccgctattattgccagcgacggccgggacttttccatccctgggctg
ctaggtcgggtacacgacctccgttttacccgcaacgtaatgctgggaagagcattgctgcaaggggccgtagaagaag
ctctccagcagagtaaaatttcctagggacggtaacgggcggtggaaggagagagaagggaagagacgtttctggttcca
ttactccacatttaagttttaccccggagaattttacgttgccagcaggtcgttcggagttgcaacggagccagatatttcagag
atggctacctgattggacgaggacaccaagacatttctacaaaaaagatctgatctcatcgaccggctgcattaatgaatcgg
ccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcgg
ctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacat
gtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttttccataggctccgccccctg
acgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccct
ggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggc
gctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgt
tcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagca
gccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac
actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa
acaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctt
tgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatctt

cacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgctta
atcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatac
gggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataa
accagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccggg
aagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtt
tggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctc
cttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactg
tcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagtt
gctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcg
gggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatct
tttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaa
atgttgaatactcatactcttcctttttcaatagctccaaggcaacaaattgactactcagaccgacattcattcgttattgattttaa
atcaacgataaacggaatggttacttgaatgatttcactttatgatcattgtttactaattacctaaataggattttatatggaattgg
aagaataagggaaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaagcccatttcttgccagtaattgcttca
taagcttcaatatacttttctttactcttgatagcaatttctgcatccatggctacgccctctttgccattcaatccgttggccgtcaa
ccaatctctgagaaactgcttatcgtaactctcttgcgatttacccacttggtaagtcttttgattccaaaatctagaagaatctgg
agttaaaacttcatctactagtaccaattcattgttttcgtccagtccaaattcgaatttcgtatcagcaataatgatccccttcaaa
agggcgaagttttttgcagcagaatacaactcgaccgccttgacagcgaccttctcacaaatgtctttacctacaatctcagca
gcttgttcaatagagatgttttcatcgtgttcaccctgttcagctttcgttgaaggtgtgaaaatcggagttggaaaggcgtcgct
ctcttgaaggttctcgttttcaaccttgactccatggacagtttttgagttcttgtactctttccatgcacttccagtgatgtaacctc
tgacaatggcttccaaaggtatcagtctgtgctttttactatcaaggatcgtccctctaattgagatttgtatttttcttcagacagt
tttgatggtagtaaagcaaagacttccttgtcattagaagcaaccaaatgattctttatgtagggtgccaaaaaatcaaaccag
aaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatcgtaagcggagatacggtcagtt
gcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttccttggcgattaaaggtaggattccgtccagat
cagtgttcacaatggacatacttggaaggatacagcaaagtgtgttggaagcgatgacacatggaaaggaatttttcgagttt
cctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttcaaagcgcccaagtgtttctgc
caacctgacttttttctgaataatgaatcgttcaagtggagtatttaaaccatgattaagttacgtgatttggcactggataaggtc
gaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggagaagcttagtaacagttatcatct

cttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcg
ctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaagaggagtgggaaataccaagaa
aaacatcaaactcgaatgattttcccaaacccctaccacaagatattcatcagctgcgagataggctgatcaggagcaagct
cgtacgagaagaaacaaaatgacaaaaaaaatcctatactatataggttacaaataaaaaagtatcaaaaatgaagcctgca
tctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgtaagtgcccaacttgaactgag
gaacagtcatgtctaaggctaaaactcacagttcgtcgtggtcgacggcgctattcagatcctcttctgagatgagtttttgttc
gggcccaaaaataacttcccaatcttcagttatgtcaagcgataggttaagaatggtgatcttattatcattctcatatgacgatc
tagtcggcaatgaactttcttcgccatttttttgactttaatactgtctttcaacgttatgttgtgttgtaaattattctttgcgataatg
atttttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcgggaatatgacttcgaacatttttttaaggtattg
ttttcgaaaatgaactgagtttcta

SEQ ID NO: 26
GACGAGATCTTTTTTTCAGACCATATGACCGG

SEQ ID NO: 27
GCGGAATTCTTTTCTCAGTTGATTTGTTTGT

SEQ ID NO: 28
GCG GGT CGA CCA CGA CGA ACT GTG AGT TTT AGC CTT AGA CAT GAC

SEQ ID NO: 29
CAG GAG CAAA GCT CGT ACG AG

SEQ ID NO: 30 ScKre2 cDNA
atggccctctttctcagtaagagactgttgagatttaccgtcattgcaggtgcggttattgttctcctcctaacattgaattccaac
agtagaactcagcaatatattccgagttccatctccgctgcatttgattttacctcaggatctatatcccctgaacaacaagtcat
ctctgaggaaaatgatgctaaaaaattagagcaaagtgctctgaattcagaggcaagcgaagactccgaagccatggatga
agaatccaaggctctgaaagctgccgctgaaaaggcagatgccccgatcgacactaaaacaaccatggattatatcactcc
atcttttgctaacaaagctggtaagccaaaagcttgttacgtcactttggtgagaaacaaggagttgaaaggtttgctaagctc
cattaaatatgtggaaaacaaaattaacaagaaattcccatatccttgggttttcctaaacgatgaaccttttactgaagaattca

aggaagcagtcaccaaagctgtttcttccgaagttaagtttggtattttgcccaaggaacattggtcatatcctgaatggattaat
caaaccaaggctgctgaaattcgtgcagatgctgccaccaaatacatatacggtggctccgaatcttatagacacatgtgtcg
ttaccaatctgggtttttctggagacatgaattattagaagagtacgattggtactggcgtgtggaaccagacatcaagttatac
tgtgatattaattacgacgtttttaagtggatgcaagaaaacgaaaaagtttacggctttaccgtttctattcatgaatatgaagtg
acgatcccaacactatggcaaacgtccatggatttcatcaaaaagaaccccgaatacttagatgaaaacaacctgatgagttt
tctttcgaacgataatggtaaaacatacaatctgtgccatttctggtcaaactttgaaattgcaaacttgaatttgtggaggtcac
cagcctacagagagtattttgacactttggatcatcaaggtggatttttctacgaaagatggggcgatgctcccgttcattctatt
gctgctgctttgtttttgccaaaggataaaatccattatttttcagacattggttaccatcatccaccttatgataactgcccattgg
acaaggaggtctataacagtaacaactgtgaatgtgaccaaggtaatgatttcactttccaaggttactcttgtggtaaggaat
attatgatgctcaagggttggtaaagccaaaaaactggaaaaaattccgtgagtagaaatcttggaacatactgtttctttgtttt
gactttatactttctatttatatttattttttataactggttaagtacacataggactgcgtatcaaacatataagtgaggcaatccac
attttttttaaagattcgaatatttttattctcattagcgtattccgagaatagttcgaaaaaatataaggtatatcaagagtttttaca
agtgagaggaaagaggaataagctataagcaacaaaagcgtaaaaaaattagctgaagacatagaactatggatgttctca
aagaggtgttgtcactagaccaagataaatttgaccagctgaaggaaacgagccgagataaaacaaatgaaacggatgatc
ctttgaaaactatttgaaggattgtaaatttaaagcgccttcaaacaaagatcagtcaccatttgctaaacttaaatcattacag
gaaactcattctaacaatgaagcggctattaatataattattcctcaattgattgattacttaaccgaattcactaataggttatcaa
attacacacaagatttagacttcattaaaaaaaagtccaatgaattacagtcattgctcgaatacaactccactaaactggcaca
tatctctcctatggttaatgatttgatgattcctcctgaactcattgatgacatcattaaagggaagatcaatgaaagctggcagg
ataatataacattcatagcagataaagaagaaatttataacaagtataggtccaataatctcgatcaagacaacaaggacgca
gaaaattcagcaatgctagcaccaaaggattttgataagttatgtcaactcctggacatcctaaaaaatgttattctaga

SEQ ID NO: 31 Human mannosidase II

aaaatgaagttaagccgccagttcaccgtgttcggcagtgcgatcttctgtgtggtgattttctcgctctacctgatgctggacc
ggggtcacttagactaccccaggaacccgcgccgcgagggctccttccctcagggccagctctcaatgttgcaagaaaaa
atagaccatttggagcgtttgctagctgagaataatgagatcatctcaaatattagagactcagtcatcaatttgagtgagtctgt
ggaggatggtccgaaaagttcacaaagcaatttcagccaaggtgctggctcacatcttctgccctcacaattatccctctcagt
tgacactgcagactgtctgtttgcttcacaaagtggaagtcacaattcagatgtgcagatgttggatgtttacagtctaatttcttt
tgacaatccagatggtggagtttggaagcaaggatttgacattacttatgaatctaatgaatgggacactgaacccttcaagt

ctttgtggtgcctcattcccataacgacccaggttggttgaagactttcaatgactactttagagacaagactcagtatatttttaa
taacatggtcctaaagctgaaagaagactcacggaggaagtttatttggtctgagatctcttacctttcaaagtggtgggatatt
atagatattcagaagaaggatgctgttaaaagtttaatagaaaatggtcagcttgaaattgtgacaggtggctgggttatgcct
gatgaagctactccacattattttgccttaattgatcaactaattgaaggacatcagtggctggaaaataatataggagtgaaac
ctcggtccggctgggctattgatccctttggacactcaccaacaatggcttatcttctaaaccgtgctggactttctcacatgctt
atccagagagttcattatgcagttaaaaaacactttgcactgcataaaacattggagtttttttggagacagaattgggatctgg
gatctgtcacagatattttatgccacatgatgcccttctacagctatgacatccctcacacttgtggacctgatcctaaaatatgc
tgccagtttgattttaaacgtcttcctggaggcagatttggttgtccctggggagtccccccagaaacaatacatcctggaaat
gtccaaagcagggctcggatgctactagatcagtaccgaaagaagtcaaagctttttcgaaccaaagttctcctggctccact
aggagatgatttccgctactgtgaatacacggaatgggatttacagtttaagaattatcagcagctttttgattatatgaattctca
gtccaagtttaaagttaagatacagtttggaacttttatcagatttttttgatgcgctggataaagcagatgaaactcagagagac
aagggccaatcgatgttccctgttttaagtggagatttttttcacttatgccgatcgagatgatcattactggagtggctattttaca
tccagacccttttacaaacgaatggacagaatcatggaatctcatttaagggctgctgaaattctttactatttcgccctgagac
aagctcacaaatacaagataaataaatttctctcatcatcactttacacggcactgacagaagccagaaggaatttgggactgt
ttcaacatcatgatgctatcacaggaactgcaaaagactgggtggttgtggattatggtaccagacttttttcattcgttaatggttt
tggagaagataattggaaattctgcatttcttcttattgggaaggacaaactcacatacgactcttactctcctgataccttcctg
gagatggatttgaaacaaaaatcacaagattctctgccacaaaaaaatataataaggctgagtgcggagccaaggtaccttg
tggtctataatcctttagaacaagaccgaatctcgttggtctcagtctatgtgagttccccgacagtgcaagtgttctctgcttca
ggaaaacctgtggaagttcaagtcagcgcagtttgggatacagcaaatactatttcagaaacagcctatgagatctcttttcga
gcacatataccgccattgggactgaaagtgtataagattttggaatcagcaagttcaaattcacatttagctgattatgtcttgtat
aagaataaagtagaagatagcggaattttcaccataaagaatatgataaatactgaagaaggtataacactagagaactcctt
tgttttacttcggtttgatcaaactggacttatgaagcaaatgatgactaaagaagatggtaaacaccatgaagtaaatgtgcaa
ttttcatggtatggaaccacaattaaaagagacaaaagtggtgcctacctcttcttacctgatggtaatgccaagccttatgttta
cacaacaccgccctttgtcagagtgacacatggaaggatttattcggaagtgacttgctttttgaccatgttactcatagagtc
cgactataccacatacagggaatagaaggacagtctgtggaagtttccaatattgtggacatccgaaaagtatataaccgtga
gattgcaatgaaaatttcttctgatataaaaagccaaaatagattttatactgacctaaatgggtaccagattcaacctagaatga
cactgagcaaattgcctcttcaagcaaatgtctatcccatgaccacaatggcctatatccaggatgccaaacatcgtttgacac
tgctctctgctcagtcattaggggtttcgagtttgaatagtggtcagattgaagttatcatggatcgaagactcatgcaagatga

taatcgtggccttgagcaaggtatccaggataacaagattacagctaatctatttcgaatactactagaaaaaagaagtgctgt
taatacggaagaagaaaagaagtcggtcagttatccttctctccttagccacataacttcttctctcatgaatcatccagtcattc
caatggcaaataagttctcctcacctacccttgagctgcaaggtgaattctctccattacagtcatctttgccttgtgacattcatc
tggttaatttgagaacaatacagtcaaaggtgggcaatgggcactccaatgaggcagccttgatcctccacagaaaagggtt
tgattgtcggttctctagcaaaggcacagggctgttttgttctactactcagggaaagatattggtacagaaacttttaaacaag
tttattgtcgaaagtctcacaccttcatcactatccttgatgcattcacctcccggcactcagaatataagtgagatcaacttgag
tccaatggaaatcagcacattccgaatccagttgaggtgaacctgactttcacatttggattgagaatcattggcttttataccttt
cttggtttgacgtgcaataaagaagcacattattttagcttctggctactgtgagaacatgaattctgtgattctgtgggtttttctt
ttttctttaccagtacagtaaga

SEQ ID NO: 32 sense primer

CGCGGATCCATGGCCAAAAAGTTCACAAAGCAATTTC

SEQ ID NO: 33 antisense primer

GTGTCCCATTCATTAGATTC

SEQ ID NO: 34 human GlcNAc-transferase I

ATGCTGAAGAAGCAGTCTGCAGGGCTTGTGCTGTGGGGCGCTATCCTCTTT
GTGGCCTGGAATGCCCTGCTGCTCCTCTTCTTCTGGACGCGCCCAGCACCT
GGCAGGCCACCCTCAGTCAGCGCTCTCGATGGCGACCCCGCCAGCCTCAC
CCGGGAAGTGATTCGCCTGGCCCAAGACGCCGAGGTGGAGCTGGAGCGGC
AGCGTGGGCTGCTGCAGCAGATCGGGGATGCCCTGTCGAGCCAGCGGGGG
AGGGTGCCCACCGCGGCCCCTCCCGCCCAGCCGCGTGTGCCTGTGACCCCC
GCGCCGGCGGTGATTCCCATCCTGGTCATCGCCTGTGACCGCAGCACTGTT
CGGCGCTGCCTGGACAAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTC
CCCATCATCGTTAGCCAGGACTGCGGGCACGAGGAGACGGCCCAGGCCAT
CGCCTCCTACGGCAGCGCGGTCACGCACATCCGGCAGCCCGACCTGAGCA
GCATTGCGGTGCCGCCGGACCACCGCAAGTTCCAGGGCTACTACAAGATC
GCGCGCCACTACCGCTGGGCGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTC
CCCGCGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCCCCGGACTTCTT
CGAGTACTTTCGGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTG

GTGCGTCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCA
GCAGGCCTGAGCTGCTCTACCGCACCGACTTTTTCCCTGGCCTGGGCTGGC
TGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCC
TTCTGGGACGACTGGATGCGGCGGCCGGAGCAGCGGCAGGGGCGGGCCTG
CATACGCCCTGAGATCTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGA
GCCACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGCTGAACCAGC
AGTTTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCT
ATGACCGAGATTTCCTCGCCCGCGTCTACGGTGCTCCCCAGCTGCAGGTGG
AGAAAGTGAGGACCAATGACCGGAAGGAGCTGGGGGAGGTGCGGGTGCA
GTATACGGGCAGGGACAGCTTCAAGGCTTTCGCCAAGGCTCTGGGTGTCA
TGGATGACCTTAAGTCGGGGGTTCCGAGAGCTGGCTACCGGGGTATTGTC
ACCTTCCAGTTCCGGGGCCGCCGTGTCCACCTGGCGCCCCCACCGACGTGG
GAGGGCTATGATCCTAGCTGGAATATGCTGAAGAAGCAGTCTGCAGGGCT
TGTGCTGTGGGGCGCTATCCTCTTTGTGGCCTGGAATGCCCTGCTGCTCCTC
TTCTTCTGGACGCGCCCAGCACCTGGCAGGCCACCCTCAGTCAGCGCTCTC
GATGGCGACCCCGCCAGCCTCACCCGGGAAGTGATTCGCCTGGCCCAAGA
CGCCGAGGTGGAGCTGGAGCGGCAGCGTGGGCTGCTGCAGCAGATCGGGG
ATGCCCTGTCGAGCCAGCGGGGGAGGGTGCCCACCGCGGCCCCTCCCGCC
CAGCCGCGTGTGCCTGTGACCCCCGCGCCGGCGGTGATTCCCATCCTGGTC
ATCGCCTGTGACCGCAGCACTGTTCGGCGCTGCCTGGACAAGCTGCTGCAT
TATCGGCCCTCGGCTGAGCTCTTCCCCATCATCGTTAGCCAGGACTGCGGG
CACGAGGAGACGGCCCAGGCCATCGCCTCCTACGGCAGCGCGGTCACGCA
CATCCGGCAGCCCGACCTGAGCAGCATTGCGGTGCCGCCGGACCACCGCA
AGTTCCAGGGCTACTACAAGATCGCGCGCCACTACCGCTGGGCGCTGGGC
CAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGAC
CTGGAGGTGGCCCCGGACTTCTTCGAGTACTTTCGGGCCACCTATCCGCTG
CTGAAGGCCGACCCCTCCCTGTGGTGCGTCTCGGCCTGGAATGACAACGG
CAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCACCG

ACTTTTTCCCTGGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGC
TGGAGCCCAAGTGGCCAAAGGCCTTCTGGGACGACTGGATGCGGCGGCCG
GAGCAGCGGCAGGGGCGGGCCTGCATACGCCCTGAGATCTCAAGAACGAT
GACCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTTCTTTGACCAGCACCT
CAAGTTTATCAAGCTGAACCAGCAGTTTGTGCACTTCACCCAGCTGGACCT
GTCTTACCTGCAGCGGGAGGCCTATGACCGAGATTTCCTCGCCCGCGTCTA
CGGTGCTCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGACCGGAAGG
AGCTGGGGGAGGTGCGGGTGCAGTATACGGGCAGGGACAGCTTCAAGGCT
TTCGCCAAGGCTCTGGGTGTCATGGATGACCTTAAGTCGGGGGTTCCGAGA
GCTGGCTACCGGGGTATTGTCACCTTCCAGTTCCGGGGCCGCCGTGTCCAC
CTGGCGCCCCCACCGACGTGGGAGGGCTATGATCCTAGCTGGAAT

SEQ ID NO: 35 ScKre2 Golgi localization signal (peptide)

MALFLSKRLLRFTVIAGAVIVLLLTLNSNSRTQQYIPSSISAAFDFTSGSISPEQQVISEENDAKKL
EQSALNSEASEDSEAMDEESKALKAAAEKADAP

SEQ ID NO: 36 pPIC6AKrecoGnTI

GAAATTTTTTTTTTAGTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAA
GTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCT
ATTACAACTTTTTTTACTTCTTGTTCATTAGAAAGAAAGCATAGCAATCTA
ATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATA
GTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGCCTTTGTCTCA
AGAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCC
CCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCC
GCATCTTCACTGGTGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAG
AACTCGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTT
GTATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTGAGCCCCTGCGGA
CGGTGCCGACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTG

AAGGACAGTGATGGACAGCCGACGGCAGTTGGGATTCGTGAATTGCTGCC
CTCTGGTTATGTGTGGGAGGGCTAAGCACTTCGTGGCCGAGGAGCAGGAC
TGACACGTCCGACGGCGGCCCACGGGTCCCAGGCCTCGGAGATCCGTCCC
CCTTTTCCTTTGTCGATATCATGTAATTAGTTATGTCACGCTTACATTCACG
CCCTCCCCCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTG
AAGTCTAGGTCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAGAACGTT
ATTTATATTTCAAATTTTTCTTTTTTTTCTGTACAGACGCGTGTACGCATGT
AACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCT
TTAATTTGCAAGCTGGAGACCAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCC
CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC
TTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA
GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATCAGATCTAA
CATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCAC
AGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTA
GCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACC
CACTTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCG
CTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTG

TCTATCCTGGCCCCCCTGGCGAGGTTCATGTTTGTTTATTTCCGAATGCAAC
AAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGT
GGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAAACTGACAGTTTAAA
CGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAA
CTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAAACTT
CCAAAAGTCGGCATACCGTTTGTCTTGTTTGGTATTGATTGACGAATGCTC
AAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCC
CGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGATG
ATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACT
GCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTAACCCCTACTT
GACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTT
TATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTTCCAATTGACAAG
CTTTTGATTTTAACGACTTTTAACGACAACTTGAGAAGATCAAAAAACAAC
TAATTATTCGAAACGAGGAATTATCCATATTCGCAAGCAGTTCCACTCGAA
AGCATGGCCCTCTTTCTCAGTAAGAGACTGTTGAGATTTACCGTCATTGCA
GGTGCGGTTATTGTTCTCCTCCTAACATTGAATTCCAACAGTAGAACTCAG
CAATATATTCCGAGTTCCATCTCCGCTGCATTTGATTTTACCTCAGGATCTA
TATCCCCTGAACAACAAGTCATCTCTGAGGAAAATGATGCTAAAAAATTA
GAGCAAAGTGCTCTGAATTCAGAGGCAAGCGAAGACTCCGAAGCCATGGA
TGAAGAATCCAAGGCTCTGAAAGCTGCCGCTGAAAAGGCAGATGCCCCGC
CGGCGGTGATTCCCATCCTGGTCATCGCCTGTGACCGCAGCACTGTTCGGC
GCTGCCTGGACAAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTCCCCA
TCATCGTTAGCCAGGACTGCGGGCACGAGGAGACGGCCCAGGCCATCGCC
TCCTACGGCAGCGCGGTCACGCACATCCGGCAGCCCGACCTGAGCAGCAT
TGCGGTGCCGCCGGACCACCGCAAGTTCCAGGGCTACTACAAGATCGCGC
GCCACTACCGCTGGGCGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTCCCCG
CGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCCCCGGACTTCTTCGAG
TACTTTCGGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTGC

GTCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAG
GCCTGAGCTGCTCTACCGCACCGACTTTTTCCCTGGCCTGGGCTGGCTGCT
GTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCCTTCT
GGGACGACTGGATGCGGCGGCCGGAGCAGCGGCAGGGGCGGGCCTGCAT
ACGCCCTGAGATCTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGAGCC
ACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGCTGAACCAGCAGT
TTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCTATG
ACCGAGATTTCCTCGCCCGCGTCTACGGTGCTCCCCAGCTGCAGGTGGAGA
AAGTGAGGACCAATGACCGGAAGGAGCTGGGGGAGGTGCGGGTGCAGTA
TACGGGCAGGGACAGCTTCAAGGCTTTCGCCAAGGCTCTGGGTGTCATGG
ATGACCTTAAGTCGGGGGTTCCGAGAGCTGGCTACCGGGGTATTGTCACCT
TCCAGTTCCGGGGCCGCCGTGTCCACCTGGCGCCCCCACCGACGTGGGAG
GGCTATGATCCTAGCTGGAATTAGCACCTGTCGACTGGAGACCTGCAGGC
ATGCAAGCTTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGAC
ATGACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCT
AGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGG
CTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT
GTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAG
CTGATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGT
TTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAC
CTTCGTTTGTGCGGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCC
TTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCA
AAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGTG
TCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTC
GTTTCTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTT
TTCTT

SEQ ID NO: 37 pGAPKreManII

ATCCATATTCGCAAGCAGTTCCACTCGAAAGCATGGCCCTCTTTCTCAGTA
AGAGACTGTTGAGATTTACCGTCATTGCAGGTGCGGTTATTGTTCTCCTCC
TAACATTGAATTCCAACAGTAGAACTCAGCAATATATTCCGAGTTCCATCT
CCGCTGCATTTGATTTTACCTCAGGATCTATATCCCCTGAACAACAAGTCA
TCTCTGAGGAAAATGATGCTAAAAAATTAGAGCAAAGTGCTCTGAATTCA
GAGGCAAGCGAAGACTCCGAAGCCATGGATGAAGAATCCAAGGCTCTGA
AAGCTGCCGCTGAAAAGGCAGATGCCCCGCCAAAAAGTTCACAAAGCAAT
TTCAGCCAAGGTGCTGGCTCACATCTTCTGCCCTCACAATTATCCCTCTCA
GTTGACACTGCAGACTGTCTGTTTGCTTCACAAAGTGGAAGTCACAATTCA
GATGTGCAGATGTTGGATGTTTACAGTCTAATTTCTTTTGACAATCCAGAT
GGTGGAGTTTGGAAGCAAGGATTTGACATTACTTATGAATCTAATGAATG
GGACACTGAACCCCTTCAAGTCTTTGTGGTGCCTCATTCCCATAACGACCC
AGGTTGGTTGAAGACTTTCAATGACTACTTTAGAGACAAGACTCAGTATAT
TTTTAATAACATGGTCCTAAAGCTGAAAGAAGACTCACGGAGGAAGTTTA
TTTGGTCTGAGATCTCTTACCTTTCAAAGTGGTGGGATATTATAGATATTC
AGAAGAAGGATGCTGTTAAAAGTTTAATAGAAAATGGTCAGCTTGAAATT
GTGACAGGTGGCTGGGTTATGCCTGATGAAGCTACTCCACATTATTTTGCC
TTAATTGATCAACTAATTGAAGGACATCAGTGGCTGGAAAATAATATAGG
AGTGAAACCTCGGTCCGGCTGGGCTATTGATCCCTTTGGACACTCACCAAC
AATGGCTTATCTTCTAAACCGTGCTGGACTTTCTCACATGCTTATCCAGAG
AGTTCATTATGCAGTTAAAAAACACTTTGCACTGCATAAAACATTGGAGTT
TTTTTGGAGACAGAATTGGGATCTGGGATCTGTCACAGATATTTTATGCCA
CATGATGCCCTTCTACAGCTATGACATCCCTCACACTTGTGGACCTGATCC
TAAAATATGCTGCCAGTTTGATTTTAAACGTCTTCCTGGAGGCAGATTTGG
TTGTCCCTGGGGAGTCCCCCCAGAAACAATACATCCTGGAAATGTCCAAA
GCAGGGCTCGGATGCTACTAGATCAGTACCGAAAGAAGTCAAAGCTTTTT
CGTACCAAAGTTCTCCTGGCTCCACTAGGAGATGATTTCCGCTACTGTGAA
TACACGGAATGGGATTTACAGTTTAAGAATTATCAGCAGCTTTTTGATTAT

ATGAATTCTCAGTCCAAGTTTAAAGTTAAGATACAGTTTGGAACTTTATCA
GATTTTTTTGATGCGCTGGATAAAGCAGATGAAACTCAGAGAGACAAGGG
CCAGTCGATGTTCCCTGTTTTAAGTGGAGATTTTTTCACTTATGCCGATCGA
GATGATCATTACTGGAGTGGCTATTTTACATCCAGACCCTTTTACAAACGA
ATGGACAGAATCATGGAATCTCATTTAAGGGCTGCTGAAATTCTTTACTAT
TTCGCCCTGAGACAAGCTCACAAATACAAGATAAATAAATTTCTCTCATCA
TCACTTTACACGGCACTGACAGAAGCCAGAAGGAATTTGGGACTGTTTCA
ACATCATGATGCTATCACAGGAACTGCAAAAGACTGGGTGGTTGTGGATT
ATGGTACCAGACTTTTTCATTCGTTAATGGTTTTGGAGAAGATAATTGGAA
ATTCTGCATTTCTTCTTATTTTGAAGGACAAACTCACATACGACTCTTACTC
TCCTGATACCTTCCTGGAGATGGATTTGAAACAAAAATCACAAGATTCTCT
GCCACAAAAAAAATATAATAAGGCTGAGTGCGGAGCCAAGGTACCTTGTGG
TCTATAATCCTTTAGAACAAGACCGAATCTCGTTGGTCTCAGTCTATGTGA
GTTCCCCGACAGTGCAAGTGTTCTCTGCTTCAGGAAAACCTGTGGAAGTTC
AAGTCAGCGCAGTTTGGGATACAGCAAATACTATTTCAGAAACAGCCTAT
GAGATCTCTTTTCGAGCACATATACCGCCATTGGGACTGAAAGTGTATAAG
ATTTTGGAATCAGCAAGTTCAAATTCACATTTAGCTGATTATGTCTTGTAT
AAGAATAAAGTAGAAGATAGCGGAATTTTCACCATAAAGAATATGATAAA
TACTGAAGAAGGTATAACACTAGAGAACTCCTTTGTTTTACTTCGGTTTGA
TCAAACTGGACTTATGAAGCAAATGATGACTAAAGAAGATGGTAAACACC
ATGAAGTAAATGTGCAATTTTCATGGTATGGAACCACAATTAAAAGAGAC
AAAAGTGGTGCCTACCTCTTCTTACCTGATGGTAATGCCAAGCCTTATGTT
TACACAACACCGCCCTTTGTCAGAGTGACACATGGAAGGATTTATTCGGA
AGTGACTTGCTTTTTTGACCATGTTACTCATAGAGTCCGACTATACCACAT
ACAGGGAATAGAAGGACAGTCTGTGGAAGTTTCCAATATTGTGGACATCC
GAAAAGTATATAACCGTGAGATTGCAATGAAAATTTCTTCTGATATAAAA
AGCCAAAATAGATTTTATACTGACCTAAATGGGTACCAGATTCAACCTAG
AATGACACTGAGCAAATTGCCTCTTCAAGCAAATGTCTATCCCATGACCAC

AATGGCCTATATCCAGGATGCCAAACATCGTTTGACACTGCTCTCTGCTCA
GTCTTTAGGGGTTTCGAGTTTGAATAGTGGTCAGATTGAAGTTATCATGGA
TCGAAGACTCATGCAAGATGATAATCGTGGCCTTGAGCAAGGTATCCAGG
ATAACAAGATTACAGCTAATCTATTTCGAATACTACTAGAAAAAAGAAGT
GCTGTTAATACGGAAGAAGAAAAGAAGTCGGTCAGTTATCCTTCTCTCCTT
AGCCACATAACTTCTTCTCTCATGAATCATCCAGTCATTCCAATGGCAAAT
AAGTTCTCACCTACCCTTGAGCTGCAAGGTGAATTCTCTCCATTACAGTCA
TCTTTGCCTTGTGACATTCATCTGGTTAATTTGAGAACAATACAGTCAAAG
GTGGGCAATGGGCACTCCAATGAGGCAGCCTTGATCCTCCACAGAAAAGG
GTTTGATTGTCGGTTCTCTAGCAAAGGCACAGGGCTGTTTTGTTCTACTACT
CAGGGAAAGATATTGGTACAGAAACTTTTAAACAAGTTTATTGTCGAAAG
TCTCACACCTTCATCACTATCCTTGATGCATTCACCTCCCGGCACTCAGAAT
ATAAGTGAGATCAACTTGAGTCCAATGGAAATCAGCACATTCCGAATCCA
GTTGAGGTGAACCTGACTTTCACATTTGGATTGAGAATCATTGGCTTTTAT
ACCTTTCTTGGTTTGACGTGCAATAAAGAAGCACATTATTTTAGCTTCTGG
CTACTGTGAGAACATGAATTCTGTGATTCTGTGGGTTTTTTCTTTTTTCTTTT
ACCAGTACAGTAAGAAAAAAAAAAAAAAAAAAAAAACTCGAGCCGCGGC
GGCCGCCAGCTTGGGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAA
TAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTTAGCCTTAGACAT
GACTGTTCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAG
ATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTT
CATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTC
ATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTG
ATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTT
TCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGACCTT
CGTTTGTGCGGATCCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT
CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCA
CCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATAT
ATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA
TCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGT
GTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA
TGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGG
TCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT
TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCG
GCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACT

CGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGAC
ACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATGCTTCTAGTGGTAG
GAATTAATTCTGTACCGGTTTACAGAAGGACGACTCTTGATGCGCCAACCA
CAGTGACAATAGACATAGAGGAAATGACAAAAGGATTATGCGAGGATGCT
GCTGGAGACGATTCAAAGTTTAGTTTAGAAAGGTCCTCCATTTATGCTGAT
AGAATACTAGATACCCGTGAACTTTGTCTCAGGAGATCCGCATCAGACGA
AGGATGTTCCGACCTGCAAATAATCGAAGAAGAGACCCCTAGGCCGTACG
ACAGTCAGTTAGTAGATATTTATACCATTCTGCGAGAAGGTCCCCCAGTAT
AGATAAGAGAGTTTCAATTGTATGTACTATAACAGTTTATTGTATCATTAA
CAAACTCAGTATAAAACCTTCAAGACTCTAGCTTTTCATTCAGTGCATCCA
ATTGCTTTAAAACAGCAGTCTTAGAAGTTCCTCCGGTGGCAGTTCTTTTTTC
AACACTGGCTTCAAAGTCAAACGTTGAAGCCACATCAGCCTCAAAACGGG
AGTCAATGGATTTCAATTGTTCGAGGGACAACTGATCAATACCAGAAAGG
TTCAACTCCTCGGCTTGTCTGACACATTCACCAGAAATGTGGTGAGTTTCT
CTGAATGGAACTCCCCTTCTAACTAAATAGTCGGCAAGATCTGTAGCCAGC
ATATCCATAGTTAGAGCATTCTTCATTCGTTCGGCATCAATGTTCAAGGTA
GAAACTACACCGGATGCTATCAAAATCGAGTGCTCTACAGTGATTAGAGT
ATCAAATAAAGGCTCCTTATCCTCTTGCATATCTTTGTTATAGGTTGACGG
AATGGACTTAATAGACATGAGGAAACCAGCCAAGGCCCCAAAACATCTAC
CAGATTTACCCCTCAATAACTCCAAAGAGTCTGGGTTTTTTTTTTGAGGCA
TCAGAGAAGATCCAGTAGAATAAGCATCTGCCAACTTGATAAATCCAAAC
TCTCCAGTGGAGTAAATGATCAAATCTTCTGAGAATCGAGAAATATGATTC
ATAAACAACGAAGACCAGAACATGGTTTCGACTACAAAATCTCTGTCTGA
AACAGCGGCCAAAGAATTACCAATAACAGAATCAAACCCTAATCTCTCAG
CAATGTATTCACGATCAATTCCATAAGGATGACCAGCCAAAGCTCCAGCTC
CCAATGGGGATTTGTTCAACCTTTTAACGATTTGATTCAGTCTCTCATAATC
TTCAGTGAAATAGGTAGCATACATGCTCAACCAGTGAGACCATCTGATTG

GTTGAGCTCTTTGCAAGTGAGTATAACCGGGCATCAAGACGTCTATTTCCT
GTTCAGCTCTCTTGATGATTACTTGAATGAACTGCTTCAGATAGTCAGCCA
ACTGAGTTAGATTGTCTCTGACATACAACCGCATATCAGTGGCAACTTGAT
CATTTCTAGACCTTCCGGTATGAACCTTACCAGAGATTCCACGACCAATCA
ACTCACCCAAGCGACGTTCATTAGCAGTGTGAATATCCTCATCCCCTGGCT
TCTCAACAAACTTCCCTTCTGCCCACTCTGCTTCAATCAATTTGAGACCAC
GATGAATCTCACTTAGTTCGTCTTTAGTTAGCAAATTAATTTTCTCCAGGCC
CTCAGTGTAAACTTTTGTTCCTTCTAAATCCACCTTGTACATTTTCTTGTCG
TAAGGTAAGGAAGCGTTATACAAATCCATCAAGGGGTCAGTAGCCCCAGT
AAACCTGCCACCCCACAGTTTAAGTCTTTCTTCTTGATTCGACATAGATAG
CTGGTAATAAGTTTAGAACAAAAGGAAAGAGAAGGTAGAATATAGGTGA
AAGAATTGCCAAAAGAGGCAGCGCGGTAAAAAGAAAAATGAATCATCGA
AATTAGTCTAGATTCTCGTATTTCTGTTGGTTATTGTATTATCTAATCAGGG
TAAAACACCTAAGGCTTACATCTCCCAGTGAAACGTGGGTATAATGAACA
GTCTTTCACAGTGAATCTGTCGCACAACCATGCTAAGATACGTTCCGTTCC
TAGACCGTAACCACCGTGAGGGCAGATTAACGCTTACAATTTCCATTCGCC
ATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGGGCACTAGTGAGATCAGATCTTTTTTGTAGAAATGTCTTGG
TGTCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAA
CTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAAT
GTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGC
CCCCTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGT
CGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTG
GCAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGA
ATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGTTTCTCCTGACCCA
AAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACAACTAT
TTCGAAACGAGGAATTCACGTGGCCC

SEQ ID NO: 38 ScKre2 Protein (amino acid sequence)

MALFLSKRLLRFTVIAGAVIVLLLTLNSNSRTQQYIPSSIAAFDFTSGSISPEQQV
ISEENDAKKLEQSALNSEASEDSEAMDEESKALKAAAEKADAPIDTKTTMDYI
TPSFANKAGKPKACYVTLVRNKELKGLLSSIKYVENKINKKFPYPWVFLNDEP
FTEEFKEAVTKAVSSEVKFGILPKEHWSYPEWINQTKAAEIRADAATKYIYGG
SESYRHMCRYQSGFFWRHELLEEYDWYWRVEPDIKLYCDINYDVFKWMQEN
EKVYGFTVSIHEYEVTIPTLWQTSMDFIKKNPEYLDENNLMSFLSNDNGKTYN
LCHFWSNFEIANLNLWRSPAYREYFDTLDHQGGFFYERWGDAPVHSIAAALF
LPKDKIHYFSDIGYHHPPYDNCPLDKEVYNSNNCECDQGNDFTFQGYSCGKE
YYDAQGLVKPKNWKKFRE

SEQ ID NO: 39 primer Kai5

AAggAgTTAgACAACCTgAAgTCTA

SEQ ID NO: 40 primer Kai7

AAggAgTAgAAACATTTTgAAgCTA

SEQ ID NO: 41 primer Kai10

CCATggCgAAggCAgATggCAgTT

SEQ ID NO: 42 primer Kai11

AgATCTTTAgTCCTTCCAACTTCCTT

SEQ ID NO: 43 primer Kai13

ATCTAAgCTATATTCgCCgTTTCTgTCATT

SEQ ID NO: 44 primer Kai14

CgATTATggATgTTAgATCTgATCTCATgA

SEQ ID NO: 45 pZ5'PpOCH1Trunc

tcatgagatcagatctaacatccaTAATCGATCTAAgctatattcgccgtttctgtcatttgcgttttgtacggaccct
cacaacaattatcatctccaaaaatagactatgatccattgacgctccgatcacttgatttgaagactttggaagctccttcaca
gttgagtccaggcaccgtagaagataatcttcgaagacaattggagtttcattttccttaccgcagttacgaacctttccccaa
catatttggcaaacgtggaaagtttctccctctgatagttcctttccgaaaaacttcaaagacttaggtgaaagttggctgcaaa
ggtccccaaattatgatcattttgtgataccgatgatgcagcatgggaacttattcaccatgaatacgaacgtgtaccagaag
tcttggaagcttttgattttaacgacttttaacgacaacttgagaagatcaaaaaacaactaattattcgcgaaacgaggaattc
acgtggcccagccggccgtctcggatcggtacctcgagccgcggcggccgccagctttctagagaacaaaaactcatctc
agaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttgtagccttagacatgactgttcctcagttcaa
gttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgcctgagagatgcaggct
tcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcattttgtttcttctcgtacgagcttgctcctgatcagcct
atctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactcctcttca
gagtacagaagattaagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatgtttctactccttttttactctt
ccagattttctcggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaaattttccctctttcttcctcta
gggtgtcgttaattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttcttttttcttcgtcgaaaaaggcaat
aaaaatttttatcacgtttctttttcttgaaatttttttttttagtttttttctctttcagtgacctccattgatatttaagttaataaacggtct
tcaatttctcaagtttcagtttcatttttcttgttctattacaacttttttacttcttgttcattagaaagaaagcatagcaatctaatcta
aggggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatggcca
agttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttc
tcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggacca
ggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtg
tccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgccctg
cgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcc
caggcctcggagatccgtcccccttttccttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctcccccacaca
tccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagtattaagaacg
ttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggttttgg

gacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgctt
accggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtag
gtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttga
gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg
gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagcca
gttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcag
cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactc
acgttaagggattttgg

SEQ ID NO: 46 pZMFManHDEL5'PpOCH1Mut

ggtcgactctagaatggcgaaggcacatggcagtttgctctagtataatcctcacaatccacccagaaggtattacttctacat
ggctatattcgccgtttctgtcatttgcgttttgtacggaccctcacaacaattatcatctccaaaaatagactatgatccattgac
gctccgatcacttgatttgaagactttggaagctccttcacagttgagtccaggcaccgtagaagataatcttcgaagacaatt
ggagtttcattttccttaccgcagttacgaacctttccccaacatatttggcaaacgtggaaagtttctccctctgatagttccttt
ccgaaaaacttcaaagacttaggtgaaagttggctgcaaaggtccccaaattatgatcatttgtgatacccgatgatgcagc
atgggaacttattcaccatgaatacgaacgtgtaccagaagtcttggaagcttttgattttaacgacttttaacgacaacttgag
aagatcaaaaaacaactaattattcgcgaaacgaggaattcacgtggcccagccggccgtctcggatcggtacctcgagcc
gcggcggccgccagctttctagagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatc
atcattgagtttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatc
aagaggatgtcagaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttt
tttgtcattttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaa
atcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcctt
ttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacgacctgct
ggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctccttccaccg
cccgttaccgtccctaggaaattttactctgctggagagcttcttctacggcccccttgcagcaatgctcttcccagcattacgtt

gcgggtaaaacggaggtcgtgtacccgacctagcagcccagggatggaaaagtcccggccgtcgctggcaataatagcg
ggcggacgcatgtcatgagattattggaaaccaccagaatcgaatataaaaggcgaacacctttcccaattttggtttctcctg
acccaaagactttaaatttaatttatttgtccctatttcaatcaattgaacaactatttcgcgaaacgatgagatttccttcaatttttta
ctgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaaacggcacaaattccggct
gaagctgtcatcggttactcagatttagaaggggatttcgatgttgctgttttgccattttccaacagcacaaataacgggttatt
gtttataaatactactattgccagcattgctgctaaagaagaggggtatctctcgagaaaagagaggctgaagctgaattcg
ccacaaaacgtggatctcccaaccctacgagggcggcagcagtcaaggccgcattccagacgtcgtggaacgcttacca
ccattttgcctttccccatgacgacctccacccggtcagcaacagctttgatgatgagagaaacggctggggctcgtcggca
atcgatggcttggacacggctatcctcatgggggatgccgacattgtgaacacgatccttcagtatgtaccgcagatcaactt
caccacgactgcggttgccaaccaaggatcctccgtgttcgagaccaacattcggtacctcggtggcctgctttctgcctatg
acctgttgcgaggtcctttcagctccttggcgacaaaccagacgcctggtaaacagccttctgaggcaggctcaaacactggc
caacggcctcaaggttgcgttcaccactcccagcggtgtcccggaccctaccgtcttcttcaaccctactgtccggagaagt
ggtgcatctagcaacaacgtcgctgaaattggaagcctggtgctcgagtggacacggttgagcgacctgacgggaaaccc
gcagtatgcccagcttgcgcagaagggcgagtcgtatctcctgaatccaaagggaagcccggaggcatggcctggcctg
attggaacgtttgtcagcacgagcaacggtacctttcaggatagcagcggcagctggtccggcctcatggacagcttctacg
agtacctgatcaagatgtacctgtacgacccggttgcgtttgcacactacaaggatcgctgggtccttggtgccgactcgacc
attgggcatctcggctctcacccgtcgacgcgcaaggacttgaccttttgtcttcgtacaacggacagtctacgtcgccaaa
ctcaggacatttggccagttttggcggtggcaacttcatcttgggaggcattctcctgaacgagcaaaagtacattgactttgg
aatcaagcttgccagctcgtactttggcacgtacacccagacggcttctggaatcggccccgaaggcttcgcgtgggtgga
cagcgtgacgggcgccggcggctcgccgccctcgtcccagtccgggttctactcgtcggcaggattctgggtgacggca
ccgtattacatcctgcggccggagacgctggagagcttgtactacgcataccgcgtcacgggcgactccaagtggcagga
cctggcgtgggaagcgttgagtgccattgaggacgcatgccgcgccggcagcgcgtactcgtccatcaacgacgtgacg
caggccaacggcgggggtgcctctgacgatatggagagcttctggtttgccgaggcgctcaagtatgcgtacctgatctttg
cggaggagtcggatgtgcaggtgcaggccaccggcgggaacaaatttgtctttaacacggaggcgcacccctttagcatc
cgttcatcatcacgacggggcggccaccttgctcacgacgagttgtaatctagggcggccgccagctttctagaacaaaaa
ctcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagttttagccttagacatgactgttcctc
agttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgcctgagagat
gcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcatttgtttcttctcgtacgagcttgctcctga

tcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactc
ctcttcagagtacagaagattaagtgagaccttcgtttgtgcggatcccccacacaccatagcttcaaaatgtttctactccttttt
tactcttccagattttctcggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaaattttccctctttct
tcctctagggtgtcgttaattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttctttttcttcgtcgaaaaa
ggcaataaaaatttttatcacgtttctttttcttgaaatttttttttttagttttttctctttcagtgacctccattgatatttaagttaataa
acggtcttcaatttctcaagtttcagtttcatttttcttgttctattacaacttttttacttcttgttcattagaaagaaagcatagcaat
ctaatctaagggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaacca
tggccaagttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctc
gggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtcca
ggaccaggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcgga
ggtcgtgtccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagtt
cgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggccca
cgggtcccaggcctcggagatccgtccccctttccttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctcc
ccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagtatt
aagaacgttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaa
ggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaacc
gtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcaga
ggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccct
gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcg
gtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatc
gtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctg
aagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgc
aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacg
aaaactcacgttaagggattttggtcatgc