



(86) Date de dépôt PCT/PCT Filing Date: 2003/06/25
(87) Date publication PCT/PCT Publication Date: 2004/01/08
(85) Entrée phase nationale/National Entry: 2004/12/20
(86) N° demande PCT/PCT Application No.: EP 2003/006711
(87) N° publication PCT/PCT Publication No.: 2004/003194
(30) Priorité/Priority: 2002/06/26 (10/185,475) US

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 9/24, C12N 15/81, C12N 9/10,
C12P 21/00

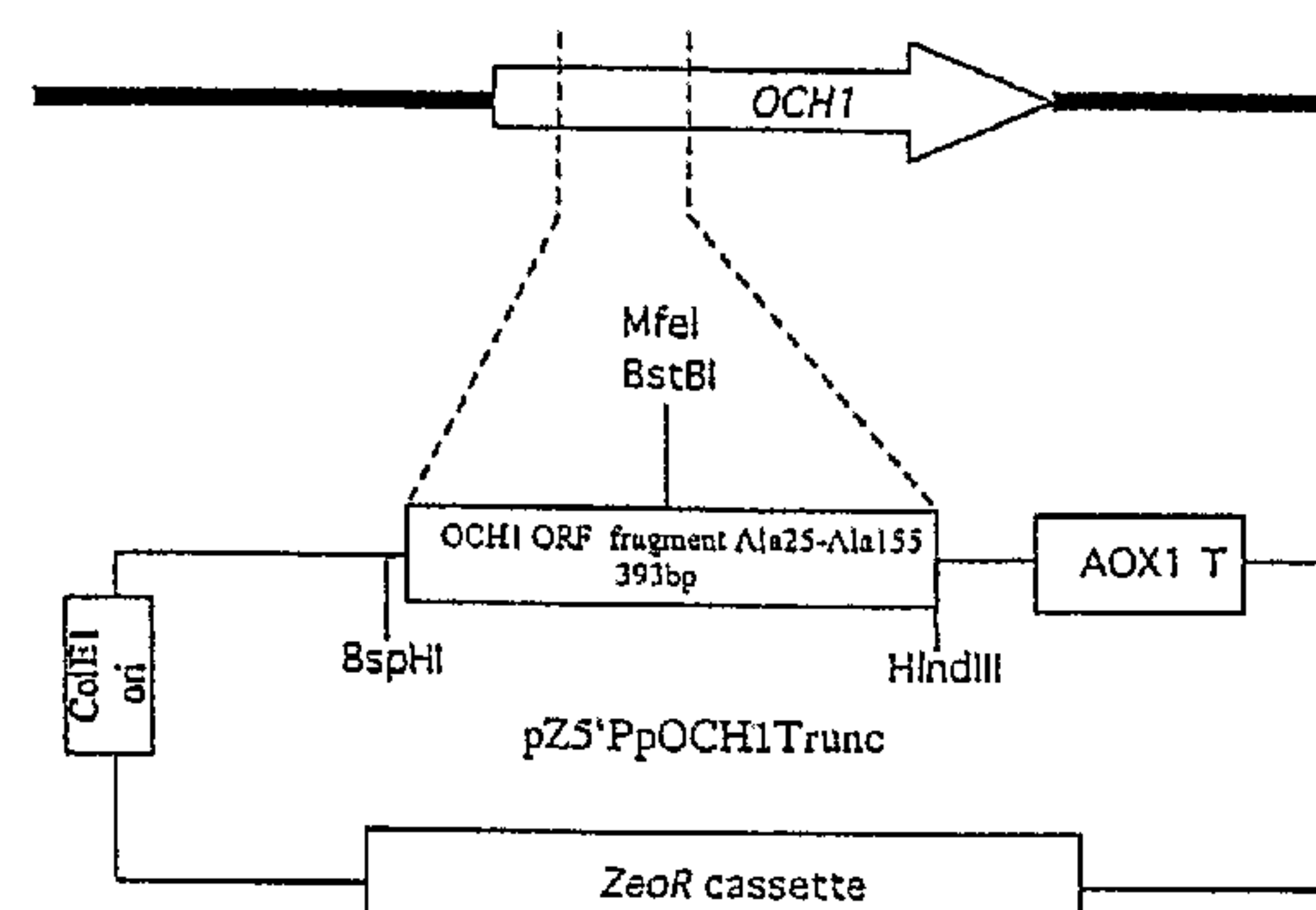
(71) Demandeur/Applicant:
FLANDERS INTERUNIVERSITY INSTITUTE FOR
BIOTECHNOLOGY, BE

(72) Inventeurs/Inventors:
CONTRERAS, ROLAND, BE;
CALLEWAERT, NICO L. M., BE;
GEYSENS, STEVEN C. J., BE;
KAIGORODOV, VLADIMIR, BE;
WOUTER, VERVECKEN, BE

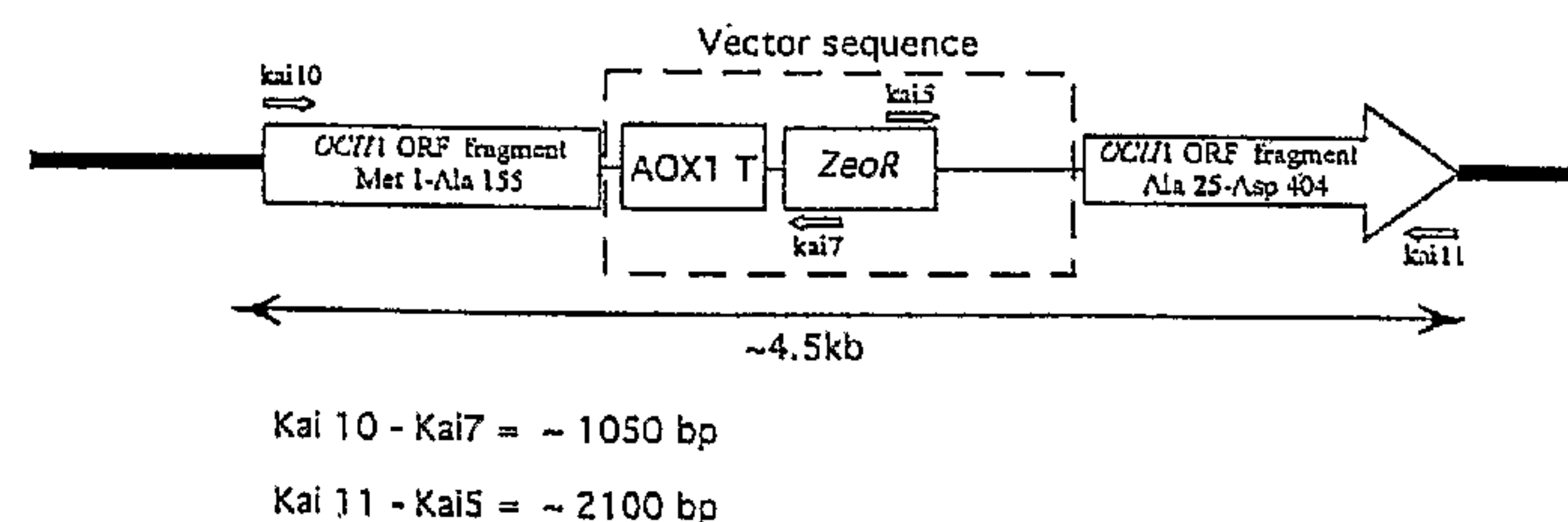
(74) Agent: MCFADDEN, FINCHAM

(54) Titre : MODIFICATION DE LA GLYCOSYLATION DES PROTEINES CHEZ PICHIA PASTORIS
(54) Title: PROTEIN GLYCOSYLATION MODIFICATION IN PICHIA PASTORIS

Construct for single homologous recombination in the Pichia OCH1 locus.



Structure of OCH1 locus after recombination with pZ5'PpOCH1Trunc.



(57) **Abrégé/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-transferase 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.



(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 A3

(51) International Patent Classification⁷: **C12N 9/24**,
15/81, C12P 21/00, C12N 9/10 // C12R 1:84

(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]; Rijvissches-
straat 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).

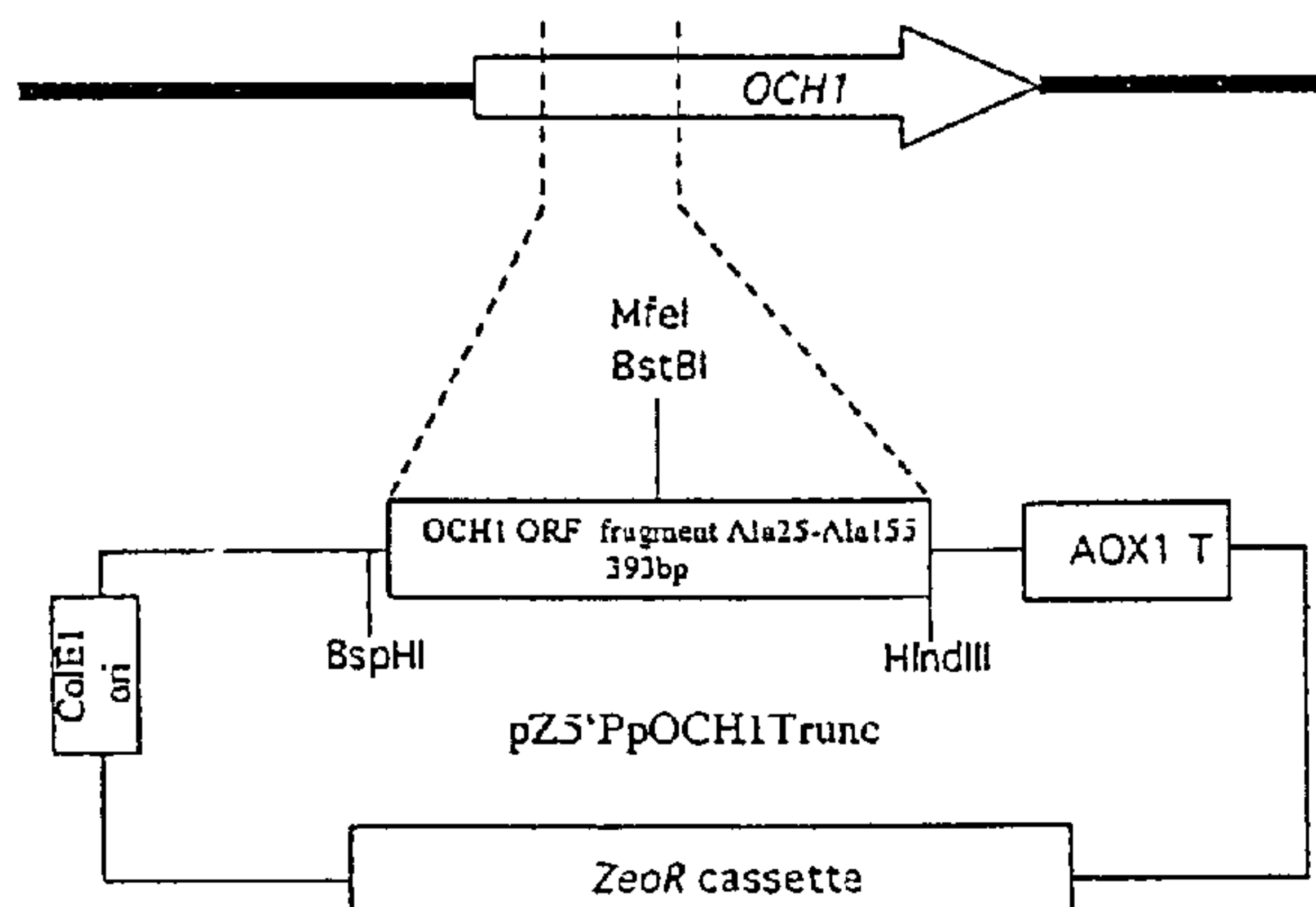
(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NI, NO, NZ, OM, PG, PH, 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),

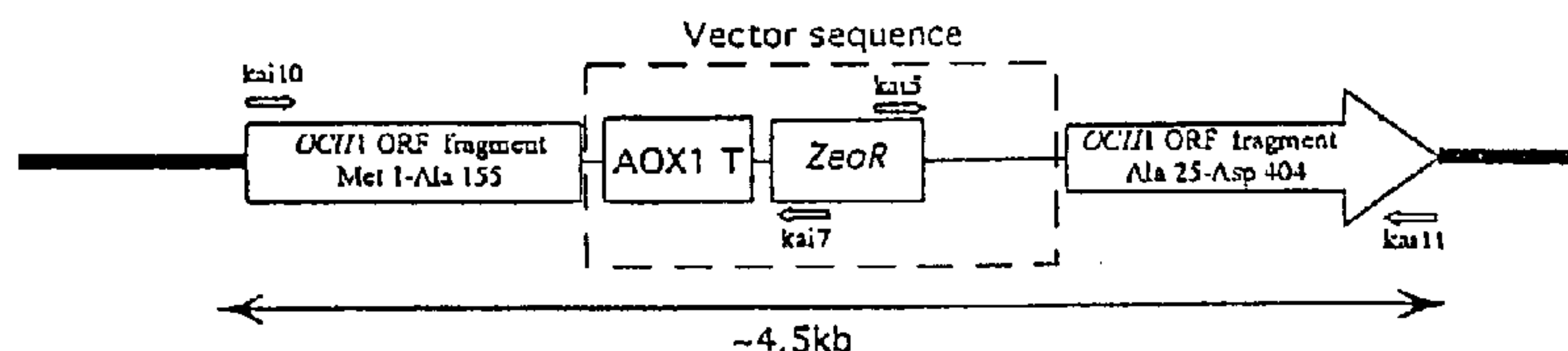
[Continued on next page]

(54) Title: PROTEIN GLYCOSYLATION MODIFICATION IN *PICHIA PASTORIS*

Construct for single homologous recombination in the *Pichia* OCH1 locus.



Structure of *UCH1* locus after recombination with pZ5⁺PpUCH1Trunc.



Kai 10 - Kai7 = ~ 1050 bp

Kai 11 - Kai5 = ~ 2100 bp

(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-transferase 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 A3

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).

*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,*

Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:

22 April 2004

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.

Protein Glycosylation Modification in *Pichia Pastoris*

Field of the Invention

The present invention relates to methods and vectors useful for genetically
5 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
15 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),
20 where the precursor N-linked oligosaccharide of 14 sugars ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) 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
25 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 $\text{Man}_8\text{GlcNAc}_2$. 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 $\text{Man}_5\text{GlcNAc}_2$ by removing three mannose residues by Golgi mannosidase I; $\text{Man}_5\text{GlcNAc}_2$ 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 $\text{Man}_{14}\text{GlcNAc}_2$ 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
5 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.*
10 *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.

15

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
20 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.

25 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
5 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
10 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.

15 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*
20 *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.
25 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
5 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
10 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
15 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.

20 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
25 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
5 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.

10 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
15 which, when introduced into a methylotrophic yeast strain, inactivates or disrupts the OCH1 gene. A preferred OCH1 inactivation vector is pBLURA5'PpOCH1 and 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
20 "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
25 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 exogenous 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
5 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*
10 *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
15 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
20 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
25 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
5 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
10 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 -
15 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
20 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 promoter EM7; Zeocin: Zeocin resistance gene; CYC1 TT: 3' end of the *S. cerevisiae* CYC1 gene; GAP: promoter
25 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 promoter (in pGAPZMFManHDEL).

5 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
10 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 promoter (in pGAPZmManHDEL). Furthermore, variants of the expression cassette were made in which the coding sequence for a cMyc epitope tag was inserted
15 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 promoter (in pGAPZmMycManHDEL).

20 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 promoter (in pGAPZMFManMycHDEL).

25 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
 5 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
 10 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
 15 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
 20 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
 25 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
5 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 Man₈GlcNAc₂. Panel C: same analytes as panel 2, but after overnight treatment with
10 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 Man₅GlcNAc₂ peak in the profile of RNase B (Panel F). Panel E: same analytes as panel D, but after overnight treatment with 3 mU/ml
15 purified recombinant *T. reesei* α -1,2-mannosidase. Panel F: N-glycans derived from bovine RNase B. These glycans consist of Man₅GlcNAc₂ to Man₈GlcNAc₂. Different isomers are resolved, accounting for the number of peaks for Man₇GlcNAc₂.

Figure 7 depicts the processing of influenza haemagglutinin N-glycans by HDEL-tagged *Trichoderma reesei* α -1,2-mannosidase and the HDEL-tagged catalytic
20 domain of murine α -1,2-mannosidase IB. The Man₅GlcNAc₂ 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 Man₉GlcNAc₂. Panel 3: N-glycans derived from recombinant
25 haemagglutinin co-expressed in *Pichia pastoris* with *T. reesei* mannosidase-HDEL (under control of the GAP promotor). The peak at scan 1950 corresponds to Man₆GlcNAc₂. 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.

5 **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
10 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
15 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}GlcNAc₂. As can be
20 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 Man₈GlcNAc₂ substrate N-glycan (named peak 1 in Panel 2) of OCH1p. Peak 2 represents the Man₉GlcNAc₂ product of OCH1p. Furthermore, upon supplementary transformation of pGAPZMFManHDEL, the major glycan on the cell wall
25 glycoproteins of three independent clones is the Man₅GlcNAc₂ end product (peak 3 in panel 4) of *T. reesei* α -1,2-mannosidase digestion of the Man₈GlcNAc₂ 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 $\text{Man}_5\text{GlcNAc}_2$ 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.

5 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*
 10 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:
 15 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:
 20 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:
 25 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 pGAPZAglslIIHDEL (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-methylumbelliferyl-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 pGAPZAglslIIHDEL).

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 $\text{Man}_8\text{GlcNAc}_2$ and $\text{Man}_9\text{GlcNAc}_2$. While TS produced in GS115 which was transformed with MannosidaseHDEL presented mainly $\text{Man}_5\text{GlcNAc}_2$, there were still some $\text{Man}_6\text{GlcNAc}_2$ and $\text{Man}_7\text{GlcNAc}_2$ 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 $\text{Man}_5\text{GlcNAc}_2$ was converted to $\text{GlcNAcMan}_5\text{GlcNAc}_2$ (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.

5 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
10 constitutive prokaryotic promoter 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
15 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
20 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
25 3' AOX1 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
5 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
10 (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
15 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
20 leaving the endoplasmic reticulum (ER) of methylotrophic yeasts, including *Pichia* and especially *Pichia pastoris*, have the Man₈GlcNAc₂ 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
25 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. Methylophilic 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 $\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_5\text{GlcNAc}_2$ and $\text{GlcNAcMan}_3\text{GlcNAc}_2$. 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 methylophilic 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 methylophilic yeast strain one or more proteins whose enzymatic activities lead to a reduction or modification of glycosylation in glycoproteins produced by the methylophilic 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
5 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 $\text{Man}_8\text{GlcNAc}_2$, and converts this core oligosaccharide on glycoproteins to $\text{Man}_5\text{GlcNAc}_2$. *In vitro*, $\text{Man}_5\text{GlcNAc}_2$ is a very poor substrate for
10 any *Pichia* Golgi mannosyltransferase, i.e., mannose residues can not be added to this sugar structure. On the other hand, $\text{Man}_5\text{GlcNAc}_2$ 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*,
15 glycoproteins with reduced mannose content glycans, such as $\text{Man}_5\text{GlcNAc}_2$, can be produced. The branching of mammalian-like oligosaccharide structures will then occur after trimming of the oligosaccharide to the $\text{Man}_5\text{GlcNAc}_2$.

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
20 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
25 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
5 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
10 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
15 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
20 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 $\text{Man}_8\text{GlcNAc}_2$ (the substrate of α -1,2-mannosidase) is already formed on a glycoprotein, but has not reached a Golgi
25 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
5 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
10 (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
15 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
20 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
25 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.

5 It has been established that the initial N-linked oligosaccharide (Glc₃Man₉GlcNAc₂), 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
10 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
15 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
20 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
25 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
5 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.

10 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
15 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.

20 A GlcNAc-Transferase I is responsible for the addition of β -1,2-GlcNAc to a $\text{Man}_5\text{GlcNAc}_2$, and converts this core oligosaccharide on glycoproteins to $\text{GlcNAcMan}_5\text{GlcNAc}_2$. The mannose residues of $\text{GlcNAcMan}_5\text{GlcNAc}_2$ 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
25 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 $\text{GlcNAcMan}_5\text{GlcNAc}_2$, 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 $\text{Man}_5\text{GlcNAc}_2$ (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.

5 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
10 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).

15 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
20 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
25 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 $\text{Man}_5\text{GlcNAc}_2$, and converts the core oligosaccharide $\text{Man}_5\text{GlcNAc}_2$ on
5 glycoproteins to $\text{GlcNAcMan}_5\text{GlcNAc}_2$. $\text{GlcNAcMan}_5\text{GlcNAc}_2$ 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
10 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
15 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).

20 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
25 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.

5 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.

10 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
20 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
25 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.

5 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
10 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*
15 *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.

20 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.

25 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
5 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
10 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
15 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
20 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,
25 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, pGAPZAglslIIHDEL and pGAPADE1glsIIHDEL, which are
5 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
10 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
15 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* OCH1 gene has been cloned (Nakayama et al. *EMBO J.*
20 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 (Man₅GlcNAc₂ and Man₈GlcNAc₂) (Nakanishi-Shindo et al. *J. Biol. Chem.* 268: 26338-26345, 1993).

A *Pichia* sequence has been described in Japanese Patent Application No.
25 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.

5 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
10 methylotrophic yeast, including *Pichia*, especially *Pichia pastoris*.

Preferred knock-in-and-inactivation vectors of the present invention include pZMFMManHDEL5'PpOCH1Mut, which are further described in the Examples hereinbelow.

15 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
20 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
25 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 anomala*, *Hansenula polymorpha* and *Candida boidinii* are examples of

methylophilic yeasts useful in the practice of the present invention. Preferred methylophilic 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; 5 PPY120H and YGC4; as well as strains derived therefrom.

Methylophilic yeast strains which can be modified using the present methods also include those methylophilic 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 10 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 methylophilic yeast strain using known methods such as the spheroplast technique, described by Cregg et al. 1985, or the whole-cell lithium chloride yeast transformation 15 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 20 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 limited to, culturing auxotrophic cells after transformation in the 25 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., $\text{Man}_8\text{GlcNAc}_2$, with the enzyme expressed and purified from a methylotrophic yeast strain, and assessing the ability of such enzyme to trim $\text{Man}_8\text{GlcNAc}_2$ to, e.g., $\text{Man}_5\text{GlcNAc}_2$. 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. Claeysens, Anal. Biochem., in press). In *in vivo* assays, the α -1,2-mannosidase or a functional part thereof is co-expressed in a

methylophilic 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 methylophilic 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 methylophilic yeast strain, and is, preferably, integrated into the genome of the methylophilic yeast strain.

The expression of Glucosidase II introduced in a methylophilic 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 methylophilic yeast strain can be assessed using a variety of assays. For example, methylophilic 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
5 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
10 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
15 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 GlcNAcMan₅GlcNAc₂, both *in vitro* and *in vivo* assays can be employed. In both *in vitro* and *in vivo* assays, the
20 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
25 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 genomic 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 (α mf), *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
5 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 inactivated 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.

10 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
15 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
20 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
25 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 1Introduction of α -1,2-Mannosidase to the ER-Golgi Border1.1 Plasmids

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

5

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'-AGTCTAGATTACAACCTCGTCGTGAGCAAGGTGGCCGCCCGTCG-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 pGAPZ α A (Invitrogen, Baarn, The Netherlands) which had been digested with *Eco* RI

10

15

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.

5 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

10 GTATCTAGATTACAACCTCGTCGTGCAGATCCTCTTCTGAGATGAGTTTTTGTTCAGCAAGGTGGCCGCCCCGTCGTGATGATGAA (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
15 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.

20 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:
25 5'AACTCGAGATGGACTCTTCAAAACACAAACGC3' (SEQ ID NO: 6) and 5'TTGCGGCCGCTTACAACCTCGTCGTGTCGGACAGCAGGATTACCTGA3' (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 pGAPZ α A 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 μ g/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

5 (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.

10 For the control of genomic integration of the mouse mannosidase transgenes, PCR was done using the sense primer 5'GAP 5'GTCCCTATTTC AATCAATTGAA3' (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

15 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.

20 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 OD₆₀₀=2-2.5, in a volume of 10 ml. After pelleting of the cells by

25 centrifugation at 3000 x g for 5 min, the cells were washed twice in sterile water and subsequently resuspended in 500 μ l 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
5 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.

10 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
15 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 CaCl₂. As blank control, 150 µl of the lysis buffer was used, supplemented with 50 microliter of the working
20 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-subtracted.

Expression of an α-1,2-Mannosidase in GS115 strains expressing influenza virus haemagglutinin was verified by qualitative Northern blot. Expression of an α-
25 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/NarI* fragment of the mannosidase, labelled with ^{32}P using HighPrime (Boehringer Mannheim).

5 *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
10 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
15 Mannheim) were used at a concentration of $1\mu\text{g/ml}$, 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
20 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
25 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.

5 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
10 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
15 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
20 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
25 (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.

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

Briefly, yeast cultures were grown to OD₆₀₀ in YPD (for pGAPZMFManMycHDEL) or in YMP following a YPGlycerol growth phase for
10 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 MgCl₂ and 4% formaldehyde and incubated for 2h at room temperature. After pelleting, the cells were resuspended to an OD₆₀₀=10 in 100 mM potassium phosphate
15 buffer pH 7.5 containing 1mM MgCl₂ and EDTA-free CompleteTM protease inhibitor cocktail (Boehringer Mannheim). To 100 µl of cell suspension, 0.6 µl of β-mercaptoethanol 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
20 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 fluorphores (Molecular probes) were used at 5µg/µl. The
25 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 fluoresensce 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
5 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

5

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
10 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
15 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
20 <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
25 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 μ g of the purified recombinant glycoproteins were used as the substrate. 1U of the α -1,2-mannosidase is defined as the amount of enzyme that releases 1 μ mol 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 Man₈GlcNAc₂ (Compare with panel F, representing an analysis of the N-glycans of bovine RNaseB. The one but last peak in this profile is Man₈GlcNAc₂, the first peak is Man₅GlcNAc₂). *In vitro*, this glycan was digestible to Man₅GlcNAc₂ 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 Man₅GlcNAc₂ (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 haemagglutinin 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 μ l of a 1:1 mixture of 20 mM APTS in 1.2M citric acid and 1M N_2CNBH_3 in DMSO and incubating for 16h at 37°C at the bottom of a 250 μ l PCR-tube. The reaction was stopped by the addition of 10 μ l 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 μ l 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 μ l gel loading buffer containing 50% formamide and 0.5 μ l Genescan 500TM, 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 μ l 10% ammoniumpersulfate solution in water and 20 μ l 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 $\text{Man}_5\text{GlcNAc}_2$ was obtained. Even after digestion of the N-glycans with 3 mU of purified *Trichoderma reesei* α -1,2-mannosidase, only $\text{Man}_6\text{GlcNAc}_2$ 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 $\text{Man}_8\text{GlcNAc}_2$ precursor to $\text{Man}_5\text{GlcNAc}_2$ 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 $\text{Man}_5\text{GlcNAc}_2$.

Example 3

Inactivation of the *Pichia* OCH1 Gene:

A *Pichia pastoris* sequence was found in the GenBank under Accession
5 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
10 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
15 *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
20 *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
25 500 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 OD₆₀₀=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 Callewaert et al. (2001) *Glycobiology* 11, 275-281. As shown in Figure 10, there was an increased amount of Man₈GlcNAc₂ glycan in the OCH1-inactivated 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 Man₅GlcNAc₂ was observed. Furthermore, upon digestion of the same glycan mixtures with 3 mU/ml purified recombinant *Trichoderma reesei* α -1,2-mannosidase, more Man₅GlcNAc₂ 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 Man₅ 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
5 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.

10 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-
15 glycosylation defective yeast cells (Proc Natl Acad Sci U S A 1991 Apr 15;88(8):3209-12).

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 promoter of pPICZB by the AOX2 promoter region of the AOX2 gene (Martinet et al.,
5 Biotechnology Letters 21). The AOX2 promoter 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'GCGGAATTCTTTTCTCAGTTGATTGTTTGT 3' (SEQ ID
10 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-
15 790, 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
20 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
25 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

5 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

10 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 pGAPZAglslIIHDEL. The sequence of plasmid pGAPZAglslIIHDEL is set forth in SEQ ID NO: 24. The construction strategy and the resulting final expression vectors (pGAPADE1glslIIHDEL and pGAPZAglslIIHDEL) are depicted in Figures 20-21.

15

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.

20 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 MgCl₂, 55°C for annealing). The primers were the same as the

25 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(myhis6) (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 Vivaspins 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

25 *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 μ l 80 mM phosphate-citrate buffer pH 6.8, 7 μ l 250 mM mannose, 3.5 μ l 250 mM 2-deoxy-D-glucose, 0.8 μ l 4-MeUmbelliferyl-alpha-D-glucopyranoside (1 μ M). Three assays were performed: one with 1 unit commercial enzyme, one without the enzyme and one with the enzyme but without the substrate.

5 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
10 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 μ l elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM
15 imidazol pH 8.0). From each eluate, 20 μ l was assayed for its glucosidase II activity. 0.33 units of the commercial enzyme diluted in 20 μ l 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
20 *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:

25 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
5 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
10 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

5

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
10 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
15 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.
20 Expression of GLSII from some of these transformants was observed.

Example 6

Introduction of GlcNAc-Transferase I in *Pichia pastoris*

5 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.

10 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

15 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 pUCKrecoGnTI 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

20 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.

25 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 integration 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 GlcNAcMan₅GlcNAc₂ 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-acetylhexosaminy l residues (presented in Figure 26).

In order to evaluate the transfer of GlcNAc to glycans of a heterologous protein or the percentage of Man₅GlcNAc₂ that was converted to GlcNAcMan₅GlcNAc₂, 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
5 glycans from TS produced in GS115 containing MannosidaseHDEL and Kre2-GlcNAc-transferase, almost all $\text{Man}_5\text{GlcNAc}_2$ was converted to $\text{GlcNAcMan}_5\text{GlcNAc}_2$.

Alternatively TS from the three different strains was separated by SDS-PAGE and blotted on a nitrocellulose membrane. First the proteins were visualized by
10 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
15 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 $\text{GlcNAcMan}_5\text{GlcNAc}_2$ 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
20 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 $\text{Man}_5\text{GlcNAc}_2$ was present in undigested N-glycans from TS produced in strain GS115 ManHDEL Kre2GnTI. Both the $\text{Man}_5\text{GlcNAc}_2$ and the $\text{Man}_6\text{GlcNAc}_2$ peaks were increased if the same glycans as
25 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

5 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
10 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 published in Cereghino et al. *Gene* 263(1-2):159-69, 2001.

15 The mannosidase cDNA fragment was purified by PCR. The primers used were the sense primer: 5'CGCGGATCCATGGCCAAAAAGTTCACAAAGCAATTTC3' (SEQ ID NO: 32) and antisense primer: 5'GTGTCCCATTTCATTAGATTTC3' (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
20 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-
25 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 pGAPAAArg. pGAPAAArg was further digested by Sfi I, blunted by Klenow, followed by Xho I digestion. The resulting vector was then ligated with the
5 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.

10 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*
15 *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 Man₅GlcNAc₂ was observed. Cell wall glycoproteins (mannoproteins) were prepared in a miniaturized way as described by Peat et al. (*J. Chem. Soc.* 29, 1961).

20 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 GlcNAcMan₅GlcNAc₂ comparable to the situation in the GS115 strain.

25 Clone 4 (YGC-4 M1 GnTI4) was transformed with plasmid pGAPKreManII by electroporation after linearisation 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 prepared 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

5 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
10 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
15 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
20 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
25 ligase, overnight at 16 °C.

The mutated 5'PpOCH1-3'AOXI 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.

SEQUENCE LISTING

SEQ ID NO: 1
5 HDEL (peptide)

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

SEQ ID NO: 3
10 5'-AGTCTAGATTACAACCTCGTCGTGAGCAAGGTGGCCGCCCCG TCG-3'

SEQ ID NO: 4
CCATTGAGGACGCATGCCGCGCC

SEQ ID NO: 5
15 GTATCTAGATTACAACCTCGTCGTGCAGATCCTCTTCTGAGATGAGTTTTTGT
TCAGCAAGGTGGCCGCCCCGTCGTGATGATGAA

SEQ ID NO: 6
20 AACTCGAGATGGACTCTTCAAAACACAAACGC

SEQ ID NO: 7
TTGCGGCCGCTTACAACCTCGTCGTGTCGGACAGCAGGATTACCTGA

SEQ ID NO: 8
25 CCATTGAGGACGCATGCCGCGCC

SEQ ID NO: 9
GCAAATGGCATTCTGACATCCT

SEQ ID NO: 10
30 GTCCCTATTTC AATCAATTGAA

SEQ ID NO: 11
35 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:

5
atgagatttccttcaatttttactgetgtttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaa
acggcacaaattccggctgaagctgtcatcggttactcagattagaaggggattcgaattgtgtgtttgccattttccaaca
gcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaagag
aggctgaagctgaattcgccacaaaacgtggatctcccaaccctacgagggcggcagcagtcaggccgcattccagac
10 gtcgtggaacgcttaccaccattttgcctttcccatgacgacctccaccgggtcagcaacagctttgatgatgagagaaacg
gctggggctcgtcggcaatcgaatggcttgacacggctatcctcatgggggatgccgacattgtgaacacgatccttcagta
tgtaccgcagatcaacttcaccacgactgcgggtgccaaccaaggatcctccgtgttcgagaccaacattcggtacctcggt
ggcctgctttctgcctatgacctgttgcgaggtcctttcagctccttggcgacaaaccagacctggtaaacagccttctgag
gcaggctcaaactggccaacggcctcaaggttgcgttcaccactcccagcggtgtcccggacctaccgtcttcttcaac
15 cctactgtccggagaagtgtgcatctagcaacaacgtcgtgaaattggaagcctggtgctcgagtggacacgggtgagc
gacctgacgggaaacccgcagtatgccagcttgccgagaagggcgagtcgtatctcctgaatccaaaggggaagcccgg
aggcatggcctggcctgattggaacgtttgtcagcacgagcaacggctacctttcaggatagcagcggcagctggtccggc
ctcatggacagcttctacgagtacctgatcaagatgtacctgtacgacctgggttgcgtttgcacactacaaggatcgctgggt
ccttgggtgccgactcgaccttgggcatctcggctctcaccgctcgacgcgcaaggacttgaccttttgtcttcgtacaacg
20 gacagtctacgtcgccaaactcaggacatttggccagtttggcggtggcaacttcatttgggaggcattctcctgaacgag
caaaagtacattgactttggaatcaagcttgccagctcgtactttggcacgtacaccagacggcttctggaatcggtcccca
aggcttcgcgtgggtggacagcgtgacgggcgcggcggtcgcgcctcgtcccagtcgggttctactcgtcggca
ggattctgggtgacggcaccgtattacatcctgcggccggagacgctggagagcttgtactacgcataccgcgtcacggg
cgactccaagtggcaggacctggcgtgggaagcgttgagtgcattgaggacgcatgccgcgccggcagcgcgtactcg
25 tccatcaacgacgtgacgcaggccaacggcgggggtgcctctgacgatatggagagcttctggttgccgaggcgctcaa
gtatgcgtacctgatctttgcggaggagtcggatgtgcaggtgcaggccaccggcggaacaaattgtctttaacacgga
ggcgcaccccttagcatccgttcacatcacgacggggcgccaccttgctcacgacgagttgtaa

SEQ ID NO: 15

30 The ORF sequence of the MFmManHDEL fusion in pGAPZMFmManHDEL:

atgagatttccttcaattttactgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaa
acggcacaaattccggctgaagctgtcatcggttactcagatttagaaggggattcgaatgttgctgtttgccattttccaaca
gcacaaataacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagatggact
5 cttcaaaacacaaacgctttgatctgggcttagaagatgtgttaattcctcacgtagatgccggcaaaggagctaaaaacccc
ggcgtcttctgatccatggacccgacgaacacagacacaggggaagaagaagagcgtctgagaaataagattagagctg
accatgagaaagccctggaagaagcaaaagaaaaattaagaaagtcaagagaggaaatccgtgcagaaattcagacaga
gaaaaacaaagtagcccaagcaatgaagacaaaagagaccaggggtactgccgcctgtccctgtcccacaacgtgtaggg
gtcagtgggtggggatccagaagacatggagatcaagaagaaaagagacaaaattaaagagatgatgaaacatgcctggg
10 ataattacagaacatacggatggggacataatgaactaaggcctattgcaaggaaaggccattccactaacatattcggaag
ctcacagatgggtgccaccatagtggatgcttggataccctttatatcatggggcttcatgatgaattcatggatgggcaaag
atggattgaagaaaaccttgattcagtgtgaattcagaagtgtctgtcttgaagttaacattcgctttattggagggtcctcg
ctgcatattacctgtcaggagaggaaatattcaagactaaagcagtgcagttggctgagaaactccttctgcctttaacacac
ctactgggattccctgggcaatggtgaacctgaaaagtggagtaggtcgaaactggggctgggcgtctgcaggcagcagc
15 atcctggctgagttcggcacccctgcacatggagtttgtcacctcagctacttgaccggtgacttgacttactataataaggtc
atgcacattcggaaactactgcagaaaatggaacgccc aaatggtctttatccaaattattaaacccaagaacagggcgctg
gggtcagtatcacacatcagttgggtggtctgggagatagttttatgaatacttactgaaagcatggctgacgtcagataaaac
agaccacgaggcaagaaggatgtatgacgatgctgttgaggctatagaaaaacatcttattaagaagtcccaggagggtct
ggtttttattggagaatggaagaatggacacttggaaaggaagatggggcacttggcctgctttgctgggggaatgcttgcc
20 cttggagcagatggttccagaaaggataaagctggccactacttagaactaggggcagaaattgcacgaacatgtcatgag
tcatatgacagaactgcattgaaactagggtccggagtcattcaagtttgatgggtgcagtggagccgtggctgtgcggcagg
ctgaaaagtattacatccttcgtccagaagtaattgaaacctattggtatctatggcgatttaccacgacccaagatacaggc
agtggggctgggaagcagcactggctattgagaagtcgtgccgggtcagcgggtgggtttctggtgtcaaggatgtatacg
ccccgaccttgtgcatgacgacgtgcagcagagcttttcttctgctgaaacattaaaatactgtacctgtgttctctggcga
25 tgaccttctaccttagaccactgggtgtttaa acacagagggcgaccctctgccgggtgttcgcttagccaacagcactcttc
aggtaatcctgctgtccgacacgacgagttgtaa

SEQ ID NO: 16 pAOX2ZAGLSII

catggccaagttgaccagtgccgttcgggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccgg
ctcgggttctcccgggacttcgtggaggacgacttcgccggtgtgggtccgggacgacgtgaccctgttcacagcgcggtc
caggaccaggtggtgccggacaacacctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcg
gaggtcgtgtccacgaacttcgggacgcctccgggcccggccatgaccgagatcggcgagcagccgtggggggcggga
5 gttcgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcc
cacgggtcccaggcctcgagatccgtcccccttttctttgtcgatacatgtaattagttatgtcacgcttacattcacgcct
ccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttttttatagttatgttagt
attaagaacgttatttataatttcaaatttttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaacctgttgga
gaaggtttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaaggccagcaaaaaggccagga
10 accgtaaaaaggccgcgttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtca
gaggtggcgaaacccgacaggactataaagataccaggcggttccccctggaagctccctcgtgcgctctcctgttccgac
cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagt
tcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccggttcagcccagccgctgcgccttatccggttaact
atcgtcttgagtccaacccggttaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggt
15 atgtaggcggtgctacagagttcttgaagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctg
ctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaacaaccaccgctggttagcggtggtttttgtt
tgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatccttgatctttctacgggggtctgacgctcagtgga
acgaaaactcacgtaagggttttggtcatgagatcagatcttttttcagaccatatgaccggtccatcttctacggggggat
tatctatgctttgacctctatcttgattctttatgattcaaatcacttttacgttatttattacttactggttatttacttagcgccctttctg
20 aaaaacatttactaaaaatcatacatcggcactctcaaacacgacagattgtgatcaagaagcagagacaatcaccactaag
gttgcacatttgagccagtaggctcctaataagagggttcgatacttattttgataatacgacataattgtcttacctctgaatgtgtca
atactctctcgttcttcgtctcgtcagctaaaaatataacacttcgagtaagatacgcccaattgaaggctacgagataccaga
ctatcactagtagaacttgacatctgctaaagcagatcaaataccatttatccagaatcaattac cttccttagcttgctgaag
gcatgaaaaagctacatgaaaatccccatccttgaagttttgtcagcttaaggactccatttcctaaaatttcaagcagtcctct
25 caactaaattttttccattcctctgcacccagccctcttcatcaaccgtccagccttctcaaaagtccaatgtaagtagcctgca
aattcaggttacaacccctcaattttccatccaaggcgatccttacaaagttaatatcgaacagcagagactaagcgagtcac
catcaccaccaacgatgggtgaaaaactttaagcatagattgatggagggtgtatggcacttggcggctgcattagagtttga
aactatggggtaatacatcacatccggaactgatccgactccgagatcatatgcaaagcacgtgatgtaccccgtaaactgc

tcggattatcgttgcaattcatcgtcttaaacagtacaagaaactttattcatgggtcattggactctgatgaggggcacatttcc
ccaatgatttttgggaaagaaagccgtaagaggacagttaagcgaaagagacaagacaacgaacagcaaaagtacag
ctgtcagctacctaagtggacagttgggagtttccaattggttgggttgaattttacccatgttgagttgccttgcttctccttgca
aacaatgcaagttgataagacatcaccttccaagataggctattttgtcgcataaattttgtctcggagtgaacccttttat
5 gtgaacagattacagaagcgtcctacccttaccgggttgagatggggagaaaattaagcgatgaggagacgattattggtat
aaaagaagcaaccaaatacccttattgtcctttctgatcagcatcaaagaatattgtcttaaacgggcttttaactacattgttc
ttacacattgcaaacctcttcttctatttcggatcaactgtattgactacattgatcttttaacgaagtttacgacttactaaatcc
ccacaaacaaatcaactgagaaaagaattcacgtggcccagccggccgtctcggatcgggtacctcgagatggccttttga
aatggctcgtatgccaatgggtcttctttaccgcttttcgcatgcgtttaccgactatctattaagaagtggtgcgcaatctgggt
10 tttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctattacaaagtggacgccgagctattgcac
acgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatgatatagccgttcagttccc
attctctctctctttttacaggatcactcagtaagggtcactataaatgagaaagagagaatgccaccaacagcagcgggttg
ttgatctcttcacaacgggtcaatgagacctggaagtacgcattcgacaagaaattcaagaggaggcgaacaggaccagta
ttccacaattccacttcttlaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcattttgtcactttcaaactccac
15 tgcagacacatttcatcttcgaaacgggtgatgtatccgtagaaatcttctgaaccttttcaattgaaagtttactggcaaatg
cgctgaaacttattgtaaacgagcaaaatttctgaacattgaacatcatagaactaagcaggaaaacttcgcacacgtgctg
ccagaagaaacaactttcaacatgtttaaggacaatttcttgattcaaagcatgactctatgcctttggggcctgaatcgggtg
cgctagatttctcttcatgggttctactaatgtctacgggtataccggaacatgcgacgtcgctaaggctgatggacacttcagg
tggaaggaaccctacaggcttttcaacgttgatgtctttgagtacaacatcgggtaccagccaaccaatgtacgggttcgatcc
20 cattcatgtttcatcttcgtccacatctatctttgggtcaatgcagctgacacttgggttagacataaagtatgacaccagtaaa
aataaaacgatgactcattggatctccgaaaatgggtgtcatagatgtagtcatgtccctggggccagatattccaactatcattg
acaaatttaccgatttgactggtagaccttttaccgccatttctctatagggtaccatcaatgtagatggaattataatgatg
agatggacgttctcacagtggactctcagatggatgctcatatgattccttacgattttatttggttgacttgaggtatacgaac
gacaaaaaatattttacttggaagcagcactccttcccaatccaaaaggctgttatccaaattaaaaaagttgggtagaat
25 ctgtcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgcagggttaattaatgaaatgtagcagtcaaggatc
acaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgataccataagcaaatatggccaaaagatt
ggaagtccttttcgaacgggttatggatctgccggctgatttaactaattttcatttggaatgatatgaacgagccttcgattt
cgatggcccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtccataacatatatggtc

tatcagtgcataagctacttacgacgcaataaaatcgattattcaccatccgataagcgtcctttccttctaacaagggctttt
 ttgccggctctcaacgtactgctgccacatggactgggtgacaatgtggccaattgggattacttaaagattccattcctatggt
 tctgtcaaacaacattgctggatgccatttataggagccgacatagctggcttgctgaggatcctacacctgaattgattgca
 cgttgggtaccaagcgggcttatgggtacccatttttagagcacacgcccataatagacaccaagagaagagaaccatacttatt
 5 caatgaaccttgaagtcgatagtagtgatattatccaattgagatatttctgctacctacctatacaccatgtttcataaatca
 agtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttctgaattgtatcatatcgataaccaatttact
 ggagtaattcaggtctattagtcacacctgtcacggagcctgggtcaatcagaaacggaaatggtttcccaccggtatattct
 atgaattcgcatactttacactctttataaacaatgggtactgatttgatagaaaagaatatttctgcaccattggataaaattccatt
 atttattgaaggcgggtcacattatcactatgaaagataagtatagaagatcttcaatgttaatgaaaaacgatccatatgtaatag
 10 ttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatggagaaacttttggctaccaagaggtgag
 tacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatattcccagagaatttgacaggcattcacc
 acaatactttgaggaataccaatattgaaaaatcattatcgcaaagaataattacaacacaacataacgttgaagacagta
 ttaaagtcaaaaaaatggcgaagaaagttcattgccgactagatcgatcatatgagaatgataataagatcaccattcttaacc
 tatcgcttgacataactgaagattgggaagttattttgggcccgaacaaaaactcatctcagaagaggatctgaatagcgcc
 15 gtcgaccatcatcatcatcattgagtttagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccg
 gtcttgctagattctaataagaggatgtcagaatgccatttgctgagagatgcaggcttcattttgatactttttatttgtaac
 ctatatagtataggattttttgtcattttgtttcttcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatctgt
 ggtaggggttgggaaaatcattcgagtttgatgttttcttggtatttccactcctcttcagagtacagaagattaagtgagac
 ctctgttgtgcggatccccacacaccatagcttcaaaatgtttctactcctttttactcttccagattttctcgactccgcgca
 20 tcgccgtaccacttcaaacacccaagcacagcatactaaattttccctctttctcctctagggtgtcgtaattaccggtacta
 aaggttggaaaagaaaaagagaccgcctcgtttcttttcttcgtcgaaaaaggcaataaaaaattttatcacgtttcttttctt
 gaaatttttttttagttttttcttttcagtgacctccattgataattaagttaataaacgggtcttcaattctcaagttcagttcatt
 ttctgttctattacaacttttttacttctgttcattagaaagaaagcatagcaatctaataaggcggtgttgacaattaatcat
 cggcatagtatatcggcatagtataatacgacaaggtgaggaactaaac

25

SEQ ID NO: 17 pAOX2ADE1glsII

tcgaccggctgcattaatgaatcgccaacgcgcggggagaggcggttgcgtattgggcgctcttccgcttctcgtcac
 tgactcgtcgcgtcggctcgggtcggcgagcggtatcagctcactcaaaggcggttaatacgggttatccacagaatca

ggggataacgcaggaaagaacatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggccgcgttgctggc
gttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact
ataaagataaccaggcggttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccg
cctttctccctcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagct
5 gggctgtgtgcacgaacccccgttcagccgaccgctgcgccttatccggttaactatcgtcttgagtccaacccggtaaga
cacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttg
aagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaag
agttggtagctcttgatccggcaacaaaccaccgctggtagcgggtggtttttgttgcaagcagcagattacgcgcagaa
aaaaaggatctcaagaagatccttgatctttctacgggggtctgacgctcagtggaaacgaaaactcacgttaagggttttgg
10 tcatgagattatcaaaaaggatcttcacctagatccttttaattaaaaatgaagtttaaatcaatctaaagtatatatgagtaaa
cttgggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgcctgactcc
ccgtcgtgtagataactacgatacgggaggggttaccatctggccccagtgctgcaatgataccgcgagaccacgctcac
cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggctctgcaactttatccgcctcca
tccagtctattaattgttgccgggaagctagagtaagtagttgccagttaatagtttgcgcaacgttggtgcatgtacagg
15 catcgtgggtgtcacgctcgtcgttgggtatggcttcattcagctccgggtcccaacgatcaaggcgagttacatgatccccat
gttggtgcaaaaaagcggttagctccttcggctcctccgatcgtgtcagaagtaagttggccgcagtggtatcactcatgttatg
gcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattctgaga
atagtgtatgcggcgaccgagttgctcttgccggcggtcaatacgggataataccgcgccacatagcagaactttaaaagtg
ctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcagatgaaccactcgt
20 gcaccaactgatcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaa
gggaataaggcgacacggaaatgttgaatactcactcttcttttcaatagctccaaggcaacaaattgactactcagac
cgacattcattcgttattgattttaaatcaacgataaacggaatgggtacttgaatgatttcatttatgatcattgttactaattacc
taaataaggattttatatggaattggaagaataagggaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaag
ccatttcttgccagtaattgttcataagcttcaatatacttttcttactcttgatagcaatttctgcatccatggctacgccctctt
25 tgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttaccacttggttaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtagtaccattcattgtttcgtccagtcctaaattcgaattcgt
atcagcaataatgatcccttcaaaaaggcggaagtttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
aaatgtctttacctacaatctcagcagcttgttcaatagagatgtttcatcgtgttcacctgttcagcttctgtgaagggtgtgaa

aatcggagttggaaggcgtcgtctcttgaaggttctcgtttcaaccttgactccatggacagttttgagttctgtactctttc
catgcacttccagtgatgtaacctctgacaatggcttccaaaggatcagtcctgtgctttttactatcaaggatcgtccctcta
tgagatttgatttttcttcagacagtttgatggtagtaaagcaaagacttccttgatagaagcaaccaaagattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatc
5 gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtcttaacctttcctttggcgatt
aaaggtaggattccgtccagatcagtggtcacatggacatacttgaaggatacagcaaagtgtgttgaagcgtatgacac
atggaaaggaattttcaggttcttagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttc
aaagcgcccaagtgttctgccaacctgactttttctgaataatgaatcgttcaagtggagtatttaaaccatgattaagtacgt
gatttggcactggataaggctgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
10 aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcgggtgcgggcctcttcgtattacgccagggcctcgaggcacaacgaacgtctcacttaattctctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaactcgaatgattttcccaaaccctaccacaagatattcatcagctgcgagat
aggctgatcaggagcaagctcgtacgagaagaacaaaatgacaaaaaaatcctatactatataggttacaataaaaaa
gtatcaaaaatgaagcctgcattctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgt
15 aagtgcccaactgaactgaggaacagtcattgtctaaggctaaaactcaatgatgatgatgatggctgacggcgctattc
agatcctctctgagatgagttttgttcgggccccaaaaataacttccaatcttcagttatgtcaagcgatagggttaagaatggt
gatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccattttttgactttaatactgtctttcaacgttatgt
tgtgttgtaaattattctttgcgataatgatttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaatctcgggaa
tatgacttcgaacatttttaaggattgttttcgaaaatgaactgagttctacgtactcacctctttggtagccaaaagtttctcca
20 tcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgttttcattaacattga
agatcttctatacttattcttcatagtgataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattcttt
ctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgtttct
gattgaccaggctccgtgacaggttgactaataagacctgaattactccagtaaaattggttatcgatatgatacaattcagcaa
attcagggtgttcaataaacattggattcattatcggaatccagtgacacttgatttatgaaacatggtgtataaggtaggttag
25 caggaaatatctcaattggataatatcacgtactatcgacttcaaagggttcattgaataagtatggttctcttcttgggtgtctata
tgggcgtgtgctctaaaaaatgggtaccataagcccgcttgggtaccaacgtgcaatcaattcagggttaggacctcagcaa
agccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaacataggaatggaaatctttaagtaatccc
aattggccacattgtcaccagtcctatgtggcagcagtagcttgagagccggcaaaaaaagcccttgtagaaggaaaggac

gcttatacggatggtgaataaatcgattttattgcgtcgtgaagtagcttcatgcactgatagaccatatatgttatggacggatctt
cctcaatgtaattgtcgtgaatcaaacttttggagctgtggictctgggccatcgaaaatcgaaggctcgttcatatcattcaa
atgaataaattagttaaatcagccggcagatccataaacggttcgaaaaaggacttccaaacttttggccatatttgcttatggt
atcaatccatatagaattacctggccagcaatgacctacatagtcattccattgtgatccttgactgctacattttcattaattacc
5 ctgtcactgatttcataatctttcttaaatgaggatcgattagtagcacaagaatttctaccaactttttaatttggataacagcct
tttggattgggaaaggagtgctgctccaagtaaaatatttttgtcgttcgtatactccaagtccaaccaaataaaatcgtaa
gaatcatatgagcatccatctgagagtcactgtgagaacgtccatctcatcattataattccatctacattgatgtaccctata
gaggaaatgggcggtaaaaagggtctaccagtc aaatcggtaaatttgtcaatgatagttggaatatctggccccagggaca
tgactacatctatgacaccatttcggagatccaatgagtcacgttttattttactggtgtcatactttatgtctaccaagtgtca
10 gctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggctggtaccg
atgttgtactcaaagacatcaacgttgaaaagcctgtagggttcccttccacctgaagtgtccatcagccttagcgacgtcgca
tgttccggtataccgtagacattagtagaacccatgaaagagaaatctagcgcaaccgattcaggcccaaaggcatagag
tcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgttcttctggcagcacgtgtgcgaagtttctgcttagttc
tatgatgttcaatgttcaggaaatttgcctcgtttacaataagtttcagcgcattttgccagtaaaccttcaattgaaaaggttcagc
15 aaagatttctacggatacatcaccgttcgaagatgaaatgtgtctgcagtggagtttgaaagtgacaaaaatgaagatatcttc
gaccagaatgagttcacagtttgttttgcttaagggaagtggaaattgtggaatactggtcctgttcgcctcctcttgaaatttctgt
cgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaccgctgctgttggttggcattctctcttctcattt
atagtgaaccttactgagtgatcctgtaaaaaagagagagagaatgggaactgaacggctatatcatcgcctccaatcttg
gtatagttttaattatggttagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagtg
20 atgagatttggcaatatcttctgcataaacctgttctatggcaaaaccagattgcgcacacttcttaatagatagtcggtaa
acgcatgcgaaaaagcggtaagaagaccaattggcatacagaccatttcaaaaggaccatctcgaggtaccgatccgag
acggccggctgggccacgtgaattctttctcagttgattgtttgtggggatttagtaagtcgtaaacttcgttaaaaaagatca
atgtagtcaatacagttgatccgaaatagaaggaagaggtttgcaatgtgtaagaacaatgtagttaaaagcccgttttaagac
aatattctttgatgctgatcagaaaaggacaataagggttttgggtgcttctttataccaataatcgtctcctcatcgtttaattt
25 ctccccatctcaaccggtgaagggtaggacgcttctgtaatctgttcacataaaaggggttttactccgagacaaaaattat
gcgacaaaaatagcctatcttgggaaggtgatgtcttatcaacttgcatgtttgcaaggagaagcaaggacaactcaacatgg
gtaaaaattcaaaaccaaccaattggaaactcccaactgtccactaggtagctgacagctgtcacttttgcgtgttcgttctgt
ctctttcgcttaactgtcctcttacggctttctttcccaaaaaatcattggggaaatgtgcccctcatcagagtccaatgacccat

gaataaagtttctgtactgtttaagacgatgaattgcaacgataatccgagcagtttacgggtacatcacgtgcttgcata
gatctcggagtcggatcagttccggatgtgatgtattaccccatagttcaaactctaatagcagccgccaagtgccatacacc
ctccatcaatctatgcttaagttttcaccatcgttgggtggtgatgatgactcgttagtctctgctgttcgatattaactttgtaa
ggatcgcccttggatggaaaattgaggggttgaacctgaattgcaggctacttacattggacttttgagaaggctggacgg
5 ttgatgaagagggtgggtgcagaggaatggaaaaaatttagttgagaggactgcttgaatttaggaaatggagtccttt
aagctgacaaaacttcaaggatggggattttcatgtagcttttcatgccttcgacaagctaaaggaaggtaattgattctggat
aaatggataattgatctgcttagcagatgtcaaagttctactagtatagctctggtatctcgtagccttcaattgggcgtatctta
ctcgaagtgttatatttttagctgacgagacgaagaacgagagagtattgacacattcagaggtaagacaatatgtcgtattat
caaaataagtatcgaaacctctattaggagcctactggctcaaagtgaaccttagtgggtgattgtctctgcttcttgatcacaat
10 ctgtcgtgttgagagtccgatgtatgatttttagtaaattgttttcagaaaaggcgctaagtaaataaccagtaagtaataaat
aacgtaaaagtgattgaatcataaaagaatcaagatagagggtcaaagcatagataatcccccgtagaagatggaccggctc
atatggctctgaaaaaagatctgatctcatg

SEQ ID NO: 18 pGAPZAGLSII

15 tcgagatggctctttgaaatggctcgtatgccaattggctcttaccgcttttcgcatgcgtttaccgactatctattaaagaa
gtgtgcgcaatctgggtttgccatagaaacagggtttatgcagaaaataattgccaaatctcatcactgctattacaaagtga
cgccgagtctattgcacacgatccttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatga
tatagccgttcagttcccattctctctctctttttacaggatcactcagtaagggtcactataaatgagaaagagagaatgccaa
ccaacagcagcggttgtgatctcttcacaacgggtcaatgagacctggaagtacgcattcgacaagaaattcaagagga
20 ggcgaaacaggaccagtatccacaattccacttccctaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcattt
ttgtcactttcaaactccactgcagacacatttcatcttcgaaacgggtgatgtatccgtagaaatctttgctgaacctttcaattg
aaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttctgaacattgaacatcatagaactaagcagga
aaacttcgcacacgtgctgccagaagaacaactttcaacatgtttaaggacaatttcttgattcaaagcatgactctatgcctt
tggggcctgaatcggttgcgctagatttctcttcatgggttctactaatgtctacgggtataccggaacatgcgacgtcgctaag
25 gctgatggacacttcagggtggaaaggaacctacaggctttcaacgttgatgtctttgagtacaacatcggtaccagccaac
caatgtacgggtcgatcccattcatgttttcatcttcgtccacatctatctttgggtcaatgcagctgacacttgggtagacataa
agtatgacaccagtaaaaaataaaacgatgactcattggatctccgaaaatgggtgtcatagatgtagtcattgcctggggcca
gatattccaactatcattgacaaatttaccgattgactggtagacccttttaccgcccatttctctatagggtaccatcaatgta

gatggaattataatgatgagatggacgttcacagtggactctcagatggatgctcatatgattccttacgattttatttggtgg
acttggagtatacgaacgacaaaaaatattttacttgaagcagcactcctttccaatccaaaaaggctgttatccaaattaa
aaagttgggtagaaatcttgtcgtactaatcgatcctcatttaaagaagattatgaaatcagtgcagggtaattaatgaaat
gtagcagtcaaggatcacatggaaatgactatgtaggtcattgctggccaggttaattctatatggattgataccataagcaaa
5 tatggccaaaagatttgaagtccttttcgaacggttatggatctgccggctgatttaactaatttattcatttgaatgatatga
acgagccttcgattttcgaaggccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtc
cataacatatatggtctatcagtgcataagctacttacgacgcaataaaatcgatttattcaccatccgataagcgtccttcc
tctaacaagggttttttggcggctctcaacgtactgctgccacatggactggtgacaatgtggccaattgggattacttaaag
attccattcctatggttctgtcaacaacattgctggatgccatttataggagccgacatagctggcttctgaggatcctac
10 acctgaattgattgcacgttgggtaccaagcgggcttatggtaccatttttagagcacacgcccataatagacaccaagagaa
gagaaccatacttattcaatgaaccttgaagtcgatagtagctgatattatccaattgagatatttctgtacctaccttataca
ccatgttccataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttctgaattgtatcatat
cgataaccaattttactggagtaattcaggtctattagtcacacctgtcacggagcctggccaatcagaaacggaaatggtttc
ccacccggtatattctatgaattcgcactctttacactctttataacaatgggtactgatttgatagaaaagaatatttctgcaccat
15 tggataaaattccattatttattgaaggcggtcacattatcactatgaaagataagtagaagatcttcaatgttaatgaaaaac
gatccatattgtaatagttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatggagaaacttttggct
accaaagaggtgagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatttcccagaaattt
gacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataattacaacacacataac
gttgaaagacagtattaaagtcaaaaaaatggcgaagaaagttcattgccgactagatcgcatatgagaatgataataaga
20 tcaccattctaacctatcgcttgacataactgaagattgggaagttattttgggcccgaacaaaaactcatctcagaagagg
atctgaatagcggcgtcgaccatcatcatcatcattgagtttagccttagacatgactgttcctcagttcaagttgggcact
tacgagaagaccggtcttgctagattctaataagaggatgtcagaatgccatttgctgagagatgcaggcttcattttgata
ctttttatttgaacctatatagtataggattttttgtcattttgttctctcgtacgagcttgctcctgatcagcctatctcgcagct
gatgaatatcttgtgtaggggttgggaaaatcattcgagttgatgttttcttgggtatttccactcctcttcagagtacagaa
25 gattaagtgcagaccttctgttgcggatccccacacaccatagcttcaaatgttctactcctttttactcttccagattttctc
ggactccgcgcacatgccgtaccacttcaaacacccaagcacagcatactaaatttccctctttctcctctaggggtgtcgtta
attaccctgactaaagggttggaaaagaaaaagagaccgcctcgttcttttctcgtcgaaaaaggcaataaaaaattttatc
acgttcttttcttgaaatttttttagtttttctcttcagtgacctccattgatatttaagttaataaacggcttcaatttctcaa

gtttcagtttcatttttctgttctattacaacttttttacttctgttcattagaaagaaagcatagcaatctaataagggcgggtgt
tgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaacatggccaagtgaccagt
ccgttccgggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttctcccgggacttc
gtggaggacgacttcgccgggtgtggtccgggacgacgtgaccctgttcacgcgcgggtccaggaccaggtggtgccgg
5 acaacaccctggcctgggtgtggtgcgcggcctggacgagctgtacgccgagtggtcggaggctcgtgtccacgaacttc
cgggacgcctccggggccggccatgaccgagatcgccgagcagccgtggggggcgggagttcgccctgcgcgaccggg
ccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcccaggcctcgga
gatccgtccccccttttcttctgtcgaatcatgtaattagttatgtcacgcttacattcacgccctccccccacatccgctctaacc
gaaaaggaaggagtagacaacctgaagtctaggtccctatttttttatagttatgttagtattaagaacgttatttatattca
10 aatttttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggtttgggacgctcgaa
ggctttaattgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttg
ctggcgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca
ggactataaagataaccaggcggttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacct
gtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctcc
15 aagctgggctgtgtgcacgaaccccccttcagcccgaccgctgcgccttatccggttaactatcgtcttgagtccaacccgg
taagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcgggtgctacagag
ttcttgaagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcgga
aaaagagttggtagctcttgatccggcaacaaaccaccgctggttagcgggtggttttttgttgcaagcagcagattacgcg
cagaaaaaaaggatctcaagaagatccttgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgttaaggg
20 attttggtcatgcatgagatcagatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctgg
ctccgttgcaactccgaacgacctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaa
cgtctcttcccttctctctccttccaccgcccgttaccgtccctaggaaattttactctgctggagagcttcttctacggccccctt
gcagcaatgctcttccagcattacgttgccgggtaaaacggaggctcgtgtacccgacctagcagcccagggatggaaaag
tccggccgctcgtggcaataatagcggggcggacgcatgtcatgagattattggaaccaccagaatcgaatataaaaggc
25 gaacaccttcccaattttggtttctctgacccaaagactttaatttaattttgtccctatttcaatcaattgaacaactatttc
gaaacgaggaattcacgtggcccagccggcggcgtctcggtatcggtacc

SEQ ID NO: 19 pGAPADE1 glsII

tcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggttgcgtattggcgctcttcgcttctcgtcac
tgactcgtcgcctcggcgttcggctgcggcgagcgggtatcagctcactcaaaggcggtaatcgggtatccacagaatca
ggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcttgctggc
gttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact
5 ataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtccg
cctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccttcagcccagccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaaga
cacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcgggtgctacagagttcttg
aagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaag
10 agttggtagctcttgatccggcaaacaccaccgctggtagcgggtggtttttgtttgcaagcagcagattacgcgcagaa
aaaaaggatctcaagaagatccttgatctttctacgggggtctgacgctcagtggaacgaaaactcacgttaagggattttgg
tcatgagattatcaaaaaggatcttcacctagatccttttaataaaaaatgaagtttaaatcaatctaaagtatatatgagtaaa
cttggctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgcctgactcc
ccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcac
15 cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggctcctgcaactttatccgctcca
tccagcttattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttggtgacattgctacagg
catcgtgggtgtcacgctcgtcgtttgggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccat
gttggtcaaaaaagcggtagctccttcggctcctccgatcgttgctcagaagtaagttggccgcagtgttatcactcatggttatg
gcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattctgaga
20 atagtgtatgcggcgaccgagttgctcttggccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtg
ctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgtatgaaccactcgt
gcaccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaaacaggaaggcaaaatgccgcaaaaaa
gggaataagggcgacacggaaatgttgaatactcatactcttcttttcaatagctccaaggcaacaaattgactactcagac
cgacattcattcgttattgatttaaatcaacgataaacggaatggttacttgaatgatttcatttatgatcattgtttactaattacc
25 taaataggattttatatggaattggaagaataagggaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaag
cccatttcttgccagtaattgcttcataagcttcaatatacttttcttactcttgatagcaatttctgcatccatgggtacgccctctt
tgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgatttaccacttggttaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtagcaattcattgttttgcagtcctcaaatcgaatttcgt

atcagcaataatgatccccttcaaaagggcgaagtttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
aatgtctttacctacaatctcagcagcttggtcaatagagatgtttcatcgtgttcaccctgttcagctttcgttgaaggtgtgaa
aatcggagttggaaaggcgtcgtctcttgaaggttctcgttttcaaccttgactccatggacagttttgagttcttgactcttctc
catgcacttccagtgatgtaacctctgacaatggcttccaaaggtatcagctctgtgctttttactatcaaggatcgctccctctaat
5 tgagatttgatttttcttcagacagtttgatggtagtaaagcaaagacttccctgtcattagaagcaaccaaatgattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatcttcccttatcaggaataaccgtttgtcataatcacatc
gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttcctttggcgatt
aaaggtaggattccgtccagatcagtggtcacaatggacatacttggaaggatacagcaaagtgtgttggaagcgatgacac
atggaaaggaattttctgagtttctagagtagtatattggggcgggtgaaagttcagatgttaatgcttaatactcttatactcttc
10 aaagcgcceaagtgtttctgccaacctgactttttctgaataatgaatcgttcaagtggagtatttaaacatgattaagttacgt
gatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcgggtgcgggcctcttcgtattacgccagggcctcgaggcacaacgaacgtctcacttaatcttctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaaactcgaatgattttccaaacccctaccacaagatattcatcagctgcgagat
15 aggctgatcaggagcaagctcgtacgagaagaacaaaaatgacaaaaaaatcctatactatataggttacaataaaaaa
gtatcaaaaatgaagcctgcatctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggtcttctcgt
aagtgcccaacttgaactgaggaacagtcagtcttaaggctaaaactcaatgatgatgatgatggtcgacggcgctatic
agatcctctctgagatgagttttgttcgggccccaaaaataacttcccaatcttcagttatgtcaagcgatagggttaagaatgt
gatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccattttttgactttaatactgtctttcaacgttatgt
20 tgtgttgtaaattattctttgcgataatgatttttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcgggaa
tatgacttcgaacatttttaaggattgttttcgaaaatgaactgagtttctacgtactcacctctttggtagccaaaagtttctcca
tcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgttttcattaacattga
agatcttctatacttattcttcatagtgataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattcttt
ctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgtttct
25 gattgaccagggtccgtgacaggttgactaatagacctgaattactccagtaaaattggttatcgatatgatacaattcagcaa
attcagggtgttcaataaacattggattcattatcggaatccagtgacacttgatttatgaaacatggtgtataaggtaggtag
caggaaatatctcaattggataatatcacgtactatcgacttcaaaggttcattgaataagtatggttctcttctcttgggtgtctata
tgggcgtgtgctctaaaaaatgggtaccataagcccgttggtaccaacgtgcaatcaattcaggtgtaggatcctcagcaa

agccagctatgtcggctcctataaatggcataccagcaatgttgttgacagaacataggaatggaaatctttaagtaatccc
aattggccacattgtcaccagtcctatgtggcagcagtagcttgagagccggcaaaaaagccctgttagaaggaaaggac
gcttatcggatggtgaataaatcgattttattgcgtcgttaagtagcttcatgcactgafagaccatatgttatggacggatcttt
cctcaatgtaattgtcgtgaatcaaacttttggagctgtggtctctgggccatcgaaaatcgaaggctcgttcatatcattccaa
5 atgaataaattagttaaatcagccggcagatccataaacggttcgaaaaaggacttccaaacttttggccatatttgcttatggt
atcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttgactgctacattttcattaattacc
ctgtcactgatttcataatctttcttaaatgaggatcgattagtagcacaagatttctaccaactttttaatttgataacagcct
tttggattgggaaaggagtgctgctccaagtaaaatatttttgcgttcgtatactccaagtccaaccaataaaatcgttaag
gaatcatatgagcatccatctgagagtcactgtgagaacgtccatctcatcattataattccatctacattgatggtaccctata
10 gaggaatgggcggtaaaaaggggtctaccagtcaaactcggtaaatttgcattgatagttggaatatctggccccagggaca
tgactacatctatgacaccatttctggagatccaatgagtcacgttttatttttactgggtgcatactttatgtctaccaagtgtca
gctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggtggtaccg
atgttgactcaaagacatcaacgttgaaaagcctgtagggttctttccacctgaagtgtccatcagccttagcgacgtcgca
tgttccggtataccgtagacattagtagaaccatgaaagagaaatctagcgcaaccgattcaggcccaaaaggcatagag
15 tcatgctttgaatacaagaaattgtccttaaacatgttgaaagttgttcttctggcagcacgtgtgcgaagtttctgcttagttc
tatgatgttcaatgttcaggaaatttgcctcgtttacaataagtttcagcgcattttgccagtaaaactttcaattgaaaaggttcagc
aaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtttgaaagtgaacaaaatgaagatatttc
gaccagaatgagttcacagtttgttttgcttaaggaagtggaaattgtggaatactggtcctgttcgcctcctcttgaaatttctgt
cgaatgcgtacttccagggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttggttggtcattctcttttctcatt
20 atagtgaaccttactgagtgatcctgtaaaaagagagagagaatgggaactgaacggctatatcatcgccctccaatcttg
gtatagttttaattatggttagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagtg
atgagatttggaatatatttctgcataaacctgtttctatggcaaaaccagattgcgcacacttcttaatagatagtcggtaa
acgcatgcgaaaaagcggtaaagaagaccaattggcatacagaccatttcaaaggaccatctcgaggtaccgatccgag
acggccggctgggccacgtgaattcctcgtttcgaatatgttggtcaattgattgaaatagggacaaataaattaaattaaagt
25 ctttgggtcaggagaaacaaaattgggaaaggtgttcgccttttatattcgattctggtggttccaataatctcatgacatgcg
tccgcccgtattattgccagcgacggccgggactttccatccctgggctgctaggtcgggtacacgacctccgttttacc
gcaacgtaatgctgggaagagcattgctgcaagggggccgtagaagaagctctccagcagagtaaaatttctaggggacg
gtaacgggcggtggaaggagagagaagggaagagacgttctggttccattactccacatttaagttttaccccggaatt

tacgttgccagcaggtcgttcggagttgcaacggagccagatattcagagatggctacctgattggacgaggacaccaag
acatttctacaaaaagatctgatctca

SEQ ID NO: 20 pPICZAGLSII

5 cgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagttttagccttagac
atgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaataagaggatgtcagaatgccattt
gcctgagagatgcaggcttcattttgatactttttatttgaacctatatagtataggattttttgtcattttgtttcttcgtacga
gcttgctcctgatcagcctatctcgcagctgatgaatatcttgtaggggtttgggaaaatcattcgagttgatgttttcttg
gtatttcccactcctcttcagagtacagaagattaagtgagaccttcgtttgtcggatccccacacaccatagcttcaaatg
10 tttctactcctttttactcttcagattttctcggactccgcgcacgcggtaccacttcaaaacaccaagcacagcatactaa
atttccctcttttctcttaggggtgctgtaattaccgtactaaagggttggaaaagaaaaagagaccgcctcgtttcttttc
ttcgtcgaaaaaggcaataaaaaattttatcacgtttcttttcttgaaatttttttagtttttcttttcagtgcctccattgat
ttaagttaataaacggtcttcaatttctcaagtttcagtttcattttcttggtctattacaacttttttacttcttggtcattagaaagaaa
gcatagcaatctaataaggggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaagggtgag
15 gaactaaaccatggccaagttgaccagtgcggtccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctgga
ccgaccggctcgggttctccgggacttcgtggaggacgacttcgcccgtgtggtcgggacgacgtgacctgttcatca
gcgcggtccaggaccaggtggtgccggacaacacctggcctgggtgtggtgcgcggcctggacgagctgtacgccg
agtggtcggaggtcgtgtccacgaacttcgggacgcctccgggcccggccatgaccgagatcggcgagcagccgtggg
ggcgggagttcgccctgcgcgacctcgccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgac
20 ggcggcccacgggtcccaggcctcggagatccgtccccctttctttgtcgatatcatgtaattagttatgtcacgcttacatt
cacgccctccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctattttttttatagt
tatgttagtattaagaacgttatttatattcaaatttttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaacc
ttgcttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaag
gccaggaaccgtaaaaaggccgcgttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgc
25 tcaagtcagaggtggcgaaacccgacaggactataaagataaccaggcggttccccctggaagctccctcgtgcgctctct
gtccgacctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgtcacgctgtagg
tatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccagccgctgcgccttate
cggtaactatcgtcttgagtccaacccggttaagacacgacttatcgccactggcagcagccactggtaacaggattagcag

agcgaggatgtaggcgggtgctacagagttctgaagtgggtggcctaactacggctacactagaaggacagtatttggtatct
gcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaccaccgctggtagcgggtg
gtttttgttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatccttgatctttctacgggggtctgacgc
tcagtggaaacgaaaactcacgttaagggattttggcatgagatcagatctaacaatccaaagacgaaaggtgaatgaaacct
5 tttgccatccgacatccacaggtccattctcacacataagtgccaaacgcaacaggaggggatacactagcagcagaccg
ttgcaaacgcaggacctccactcctctctcctcaacacccacttttgccatcgaaaaaccagcccagttattgggcttgattg
gagctcgctcattccaattccttctattaggctactaacacatgactttattagcctgtctatcctggccccctggcgagggtc
atgtttgtttatttccgaatgcaacaagctccgcattacacccgaacatcactccagatgagggctttctgagtgtgggtcaa
atagtttcatgttcccaaagtggcccaaaactgacagtttaaacgctgtcttggaaacctaatatgacaaaagcgtgatctcatcc
10 aagatgaactaagtttggttcgttgaaatgctaacggccagttgggtcaaaaagaaacttccaaaagtcggcataccgtttgtct
tgtttggtattgattgacgaatgctcaaaaataatctcattaatgcttagcgcagttctctatcgttctgaaccccggtgcacct
gtgccgaaacgcaaagtggggaacacccgcttttggtatgattatgcattgtctccacattgtatgcttccaagattctgggtg
gaatactgctgatagcctaacgttcatgatcaaaatttaactgttctaaccctacttgacagcaatatataaacagaaggaag
ctgccctgtcttaaacctttttttatcatcattattagcttactttcataattgcgactgggtccaattgacaagcttttgattttaacg
15 actttaacgacaacttgagaagatcaaaaaacaactaattatcgaaacgaggaattcacgtggcccagccggcgtctcg
gateggtagctcgagatgggtccttttgaaatggctcgtatgccaatgggtctctttaccgcttttcgcatgcgtttaccgactat
ctattaaagaagtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaatctcatcactgctatt
acaaagtggacgccgagttctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattgg
agggcgatgatatagccgttcagttcccattctctctctctttttacaggatcactcagtaagggttactataaatgagaaagag
20 agaatgccaccaacagcagcggtttggtgatctcttcacaacgggtcaatgagacctggaagtacgcattcgacaagaaatt
tcaagaggaggcgaacaggaccagtattccacaattccacttccttaagcaaaaacaaactgtgaactcattctggctcgaaa
atatcttcattttgtcactttcaaactccactgcagacacatttcattctcgaaacgggtgatgtatccgtagaaatcttgctgaac
ctttcaattgaaagtttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttctgaacattgaacatcatagaact
aagcaggaaaacttcgcacacgtgctgccagaagaacaactttcaacatgtttaaggacaatttctgtattcaaagcatga
25 ctctatgcctttggggcctgaatcggttgcgctagatttctctttcatgggttctactaatgtctacgggtataccggaacatgcga
cgctcgctaaggctgatggacacttcaggtggaaaggaaccctacaggcctttcaacgttgatgtctttgagtacaacatcggt
accagccaaccaatgtacgggtcgtatccattcatgtttcatcttcgtccacatctatctttgggtcaatgcagctgacacttg
ggtagacataaagtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatgggtgtcatagatgtagtcatgt

ccctggggccagatattccaactatcattgacaaatttaccgatttgactggtagacccttttaccgcccatttcctctatagg
 taccatcaatgtagatggaattataatgatgagatggacgttcacagtggactctcagatggatgctcatatgattccttacg
 attttattgggtggacttggagtatacgaacgacaaaaatattttacttgggaagcagcactcctttcccaatccaaaaaggctg
 ttatccaaattaaaaaagttgggtagaaatcttgcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgacagggt
 5 aattaatgaaaatgtagcagtcaggatcacatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgat
 accataagcaaatatggccaaaagatttgggaagtccttttcgaacggtttatggatctgccggctgatttaactaatttattcatt
 tggaaatgatatgaacgagccttcgattttcgatggcccagagaccacagctccaaaagatttgattcacgacaattacattga
 ggaaagatccgtccataacatatgtgtctatcagtgcatgaagctacttacgacgcaataaaatcgatttattcaccatccgat
 aagcgtcctttccttctaacaagggcctttttgccggctctcaacgtactgctgccacatggactggtgacaatgtggccaatt
 10 gggattacttaaagatttccattcctatggttctgtcaacaacattgctgggtatgccatttataggagccgacatagctggcttt
 gctgaggatcctacacctgaattgattgcacgttggtagcaagcgggcttatggtagccatttttagagcacacgcccata
 gacaccaagagaagagaaccatacttattcaatgaaccttgaagtcgatagtagtgatattatccaattgagatatttctgc
 tactaccttatacaccatgtttcataaatcaagtgtcactggatttccgataatgaatecaatgtttattgaacacctgaattgc
 tgaattgtatcatatcgataaccaattttactggagtaattcagggtctattagtcacacctgtcacggagcctgggtcaatcagaaa
 15 cggaaatggttttccacccgggtatatctatgaattcgcacttttacactctttataaacaatgggtactgattgatagaaaaga
 atatttctgcaccattggataaaattccattatttattgaaggcgggtcacattatcactatgaaagataagtatagaagatcttcaa
 tggtaatgaaaaacgatccatagtgaatagttatagcccctgataccgagggacgagccgttggagatctttatgttgatgatg
 gagaaacttttggctaccaaagaggtgagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtc
 atattcccagagaatttgacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataatt
 20 acaacacaacataacgttgaaagacagtattaaagtcaaaaaaatggcgaagaaagttcattgccgactagatcgtcatat
 gagaatgataataagatcaccattctaacctatcgcttgacataactgaagattgggaagttattttggggcc

SEQ ID NO: 21 pPICADE1glsII

aaattcctcgtttcgaataattagttgtttttgatcttctcaagttgtcgttaaaagtcgttaaaatcaaaagcttgcattggaac
 25 cagtcgcaattatgaaagtaagctaataatgatgataaaaaaaagggttaagacagggcagcttccttctgtttatatattgct
 gtcaagtaggggttagaacagttaaattttgatcatgaacgttaggctatcagcagttatcccaccagaatcttgggaagcatac
 aatgtggagacaatgcataatcatccaaaagcgggtgtttccccatttgcgtttcggcacaggtgcaccgggggttcagaag
 cgatagagagactgcgctaagcattaatgagattattttgagcattcgtcaatcaataccaacaagacaaacgggtatgccg

acttttgaagtttcttttgaccaactggccgtagcatttcaacgaaccaaacttagttcatcttgatgagatcacgcttttgc
atattaggtccaagacagcggttaaactgtcagtttgggccatttggggaacatgaaactatttgacccacactcagaaag
ccctcatctggagtgatgttcgggtgtaatgcggagcctgttgcatcggaaataaacaacatgaacctgccaggggggc
caggatagacaggctaataaagtcagtgtgtagtagcctaataagaaggaattggaatgagcgagctccaatcaagcccaat
5 aactgggctggttttcgtatggcaaaagtgggtgttgaggagaagaggagtggaggtcctgcgtttgcaacggtctgctgct
agtgtatcccctcctgttgcgtttggcacttatgtgtgagaatggacctgtggatgtcggatggcaaaaaggtttcattcaacct
ttcgtctttggatgttgcgaccggctgcattaatgaatcggccaaacgcgcggggagaggcggtttgcgtattgggcgctctt
ccgcttcctcgtcactgactcgtgcgctcggctcgttcggctgcggcgagcggtatcagctcactcaaaggcggtataacg
gttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaa
10 ggccgcgttgctggcgttttccataggtcgcgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataaccaggcggtttcccctggaagctccctcgtgcgctctcctgttccgacctgccgctt
accggataacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtag
gtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccagccgctgcgccttatccggttaactatcgtcttga
gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg
15 gtgctacagagttctgaagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagcca
gttaccttcggaaaaagagttggttagctcttgatccggcaacaaaccaccgctggttagcggtggtttttgtttgcaagcag
cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacgggggtctgacgctcagtggaacgaaaactc
acgttaagggttttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtttaaatcaatcta
aagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgctatttcgttcac
20 catagttgcctgactccccgtcgtgtagataactacgatacgggagggccttaccatctggccccagtgctgcaatgataccg
cgagaccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggctcctg
caactttatccgctccatccagttctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgtt
gttgccattgctacaggcatcgtgggtgtcacgctcgtcgtttggtatggcttcattcagctccggttccaacgatcaaggcga
gttacatgatccccatgttgtgcaaaaaagcggttagctccttcggctcctccgatcgttgtcagaagtaagtggccgcagtg
25 ttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtagtactca
accaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcggcggtcaatacgggataataccgcgccacata
gcagaactttaaaagtgtcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagtt
cgatgtaaccactcgtgcaccaactgatcttcagcatctttactttaccagcggtttctgggtgagcaaaaacaggaaggc

aaaatgccgcaaaaaaggggaataagggcgacacggaaatgttgaatactcactcttcttttcaatagctccaaggcaac
aaattgactactcagaccgacattcattcgttattgattttaaataacgataaacggaatggttacttgaatgatttcactttatga
tcattgtttactaattacctaataaggattttatatggaattggaagaataagggaaatttcagatgtctgaaaaaggcgaggag
ggtactaatcattcaagcccatttcttgccagtaattgcttcataagcttcaatatacttttcttactcttgatagcaattctgcac
5 catggctacgccctctttgccattcaatccgttggccgtcaaccaatctctgagaaactgcttactgtaactctcttgcgatttac
ccacttggttaagtcttttgattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgtttcgtccag
tccaaattcgaatttcgtatcagcaataatgatcccttcaaaggcggaagttttgcagcagaatacaactcgaccgccttg
acagcgaccttctcacaatgtctttacctacaatctcagcagcttgttcaatagagatgtttcatcgtgttcacctgttcagct
ttcgttgaagggtgtgaaaatcggagttggaaaggcgctcgtctcttgaaggttctcgtttcaaccttgactccatggacagtttt
10 tgagttcttgactctttccatgcacttcagtgatgtaacctctgacaatggcttcaaagggtatcagttctgtctttttactatca
aggatcgtccctctaattgagatttgatttttcttcagacagtttgatggtagtaaagcaaagacttcttgcattagaagcaa
ccaatgattctttatgtagggtgccaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaatac
cgttgcataatcacatcgtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctc
taacctttcctttggcgattaaaggtaggattccgtccagatcagtggtcacaatggacatacttgaaggatacagcaaagtg
15 tgttgaagcgtatgacacatggaaaggaattttcaggttcttagagtagtatattggggcgggtgaaagttcagatgttlaatg
cttaatactcttatactcttcaaagcgcccaagtgtttctgccaacctgactttttctgaataatgaatcgttcaagtggagtattta
aaccatgattaagttacgtgatttggcactggataaggctgaaaaatatccgtattcataaacgattattggtaaaagttacaaa
ataccactaattacggagaagcttagtaacagttatcatctcttggtcgattaacgcttacaatttcattcgcattcaggctgc
gcaactgttgggaagggcgatcgggtcggggcctcttcgtattacgccagggcctcgaggcacaacgaacgtctcactta
20 atcttctgtactctgaagaggagtgggaaataccaagaaaaacatcaaactcgaatgattttccaaaccctaccacaagat
attcatcagctgcgagataggctgatcaggagcaagctcgtacgagaagaacaaaatgacaaaaaaatccitactatat
aggttacaataaaaaagtatcaaaaatgaagcctgcattctcaggcaaatggcattctgacatcctcttgattagaatctag
caagaccggtcttctcgttaagtgcccaactgaactgaggaacagtcattgtctaaaggctacaaactcaatgatgatgatg
atggtcgacggcgctattcagatcctctctgagatgagttttgttcgggccccaaaataacttccaatcttcagttatgtcaa
25 gcgataggttaagaatggtgatcttattatcattctcatatgacgatctagtcggcaatgaactttctcgccattttttgacttta
atactgtctttcaacgttatgttgtgtgttaaattattctttgcgataatgatttttcaatattggtattcctcaaagtattgtggtgaat
gcctgtcaaattctcgggaatatgacttcgaacatttttaaggtattgtttcgaaaatgaactgagtttctacgtactcacctcttt
ggtagccaaaagtttctccatcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatag

gatcgttttcattaacattgaagatcttctatacttattctttcatagtgataatgtgaccgccttcaataaataatggaatttatcca
atgggtgcagaaatattctttctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggt
gggaaaaccatttccgtttctgattgaccagggtccgtgacagggttgactaataagacctgaattactccagtaaaattggttat
cgatatgatacaattcagcaaattcagggtgttcaataaacattggattcattatcggaaatccagtgcacttgatttatgaaac
5 atgggtgtataaggtaggtagcaggaaatctcaattggataaatcacgtactatcgacttcaaagggttcattgaataagtatg
gttctcttctcttgggtgtctatatgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattca
gggtgtaggatcctcagcaaagccagctatgtcgggtcctataaatggcataaccagcaatgttggttgacagaacctaggaat
ggaaatctttaagtaatcccaattggccacattgtcaccagtcctatgtggcagcagtagcttgagagccggcaaaaaagcc
cttgtagaaggaaaggacgcttatcggtatgggtgaataaatcgattttattgctcgttaagtagcttcatgcactgatagaccat
10 atatgttatggacggatctttcctcaatgtaattgtcgtgaatcaaacttttggagctgtggtctctgggccatcgaaaatcgaa
ggctcgttcatatcattccaaatgaataaattagttaaatcagccggcagatccataaacggttcgaaaaaggacttccaaatc
tttggccatatttgcttatggtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttga
ctgctacattttcattaattacctgtcactgatttcataatctttcttlaaatgaggatcgattagtagcacaagatttctaccaac
tttttaatttggataacagcccttttggattgggaaaggagtgtgcttccaagtaaaatatttttgcgttcgtatactccaagtc
15 caaccaataaaaatcgtaaggaaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccat
ctacattgatgtaccctatagaggaaatgggcggtaaaaagggtctaccagtcaaactcggtaaatttgtcaatgatagttgg
aatatctggccccaggacatgactacatctatgacaccatttccggagatccaatgagtcacgttttattttactgggtgtcata
ctttatgtctaccaagtgtcagctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccg
tacattggttggctgggtaccgatgttgactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtcca
20 tcagccttagcgacgtcgcgtgttccggtataccgtagacattagtagaaccatgaaagagaaatctagcgcaaccgattc
aggccccaaggcatagagtcgttgaatacaagaaattgtccttaaacatgttgaaagtgtttctctggcagcacgtgt
gcgaagtttctctgcttagttctatgatgttcaatgttcaggaaatttgcgtcttacaataagtttcagcgcattttgccagtaaa
ctttcaattgaaaagggtcagcaaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtggagtgtgaaag
tgacaaaaatgaagatatttgcaccagaatgagttcacagtttgttttgcttaagggaagtgaattgtggaatactggctcctgt
25 tcgcctcctcttgaaatttctgtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttg
gttggcatttctcttttctcatttatagtgaaccttactgagtatcctgtaaaaagagagagagaatgggaactgaacggcta
tatcatcgccctccaatcttgggtatagtttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcg
tccactttagaatagcagtgatgagatttggcaatatttctgcataaacctgtttctatggcaaaaccagattgcgcacactt

ctttaatagatagtcggtaaacgcatgcgaaaaagcggtaaagaagaccaattggcatacagagccatttcaaaaggaccatc
tcg

SEQ ID NO: 22 pYPT1ZAGLSII

5 cgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagttttagccttagac
atgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaataagaggatgtcagaatgccattt
gcctgagagatgcaggcttcattttgatactttttatttgaacctatatagtataggattttttgtcattttgtttcttcgtacga
gcttgctcctgatcagcctatctcgcagctgatgaatatcttgtaggggttgggaaaatcattcgagtttgatgttttcttg
gtatttcccactcctcttcagagtacagaagattaagtgcagaccttcgttgtagggatccccacacaccatagcttcaaatg
10 tttctactcctttttactcttcagattttctcggactccgcgcacgcggtaccacttcaaacacccaagcacagcatactaa
attttccctctttctcctctaggggtgctgtaattaccggtactaaagggttggaaaagaaaaagagaccgctcgtttcttttc
ttcgtcgaaaaaggcaataaaaaattttatcacgtttcttttctgaaatttttttagtttttctcttcagtacctccattgatat
ttaagttaataaacggtcttcaatttctcaagtttcagtttcattttcttggtctattacaactttttacttctgttcattagaaagaaa
gcatagcaatctaataagggggcggtgtgacaattaatcatcggcatagtatatcggcatagtataatacgacaagggtgag
15 gaactaaaccatggccaagtgaccagtgccgttcgggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctgga
ccgaccggctcgggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcattca
gcgcggtccaggaccaggtggtgccggacaacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccg
agtggctcggaggtcgtgtccacgaacttcgggacgcctccgggcccggccatgaccgagatcggcgagcagccgtggg
ggcgggagttcgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgac
20 ggcgggcccacgggtcccaggcctcgagatccgtcccccttttcttgcgatatcatgtaattagttatgtcacgcttacatt
cacgcctccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttttttatagt
tatgtagtattaagaacgttatttatatttcaaatttttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaacc
ttgcttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaag
gccaggaaccgtaaaaaggccgcgttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgc
25 tcaagtcagaggtggcgaaacccgacaggactataaagataccaggcggttccccctggaagctccctcgtgcgctctcct
gttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtagg
tatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccagaccgctgcgccttacc
cggttaactatcgtcttgagtccaaccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag

agcgaggatgtaggcgggtgctacagagttcttgaagtgggtggcctaactacggctacactagaaggacagtatttggtatct
gcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaaccaccgctggtagcgggtg
gttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacgggggtctgacgc
tcagtggaaacgaaaactcacgttaagggattttggtcatgagatcagatctatgatgagtcacaatctgcttcacagacgagt
5 acaaggacaggcaaaaggaattggaagaagttgctaaccaataatgagcaagttctatggagctgctgggtgagctcctg
gtggagctcctgggtggcttcctggagggttcctggcggagctggcgcagctggcgggtgccccagggtggtgctgcccc
ggcggagacagcggaccaaccgtggaagaagtcgattaagcaattcaacggataaattctggtaatatataaacgtgaat
aggaaattaaggaaattttggatctaataatgtgctgtatgccgacatcgggcacgtagattgtatagtatcgtgacactata
ataagccagccaaaaccctaaaccagttgccctccactaattagtgtactaccaatcttgccctcttcgggtgtctttataag
10 gacagattcacaagctctgttgcccaatacacacatacacacagagataatagcagtcgaattcacgtggcccagccggc
cgtctcggatcggtagctccttgagatggctctttgaaatggctcgtatgccaatggctctctttaccgcttttcgcatgcgttac
cgactatctattaaagaagtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcac
tgctattacaaagtggacgccgagctctattgcacacgacatccttagagaatgtgcttcattgctaccataattaaaactataccaa
gattggagggcgatgatatagccgttcagttccattctctctctctttttacaggatcactcagtaaggttactataaatgag
15 aaagagagaatgccaccaacagcagcgggtttgttgatctcttcacaacgggtcaatgagacctggaagtacgcattcgaca
agaaatttcaagaggaggcgaacaggaccagttccacaattccacttccttaagcaaaaacaaactgtgaactcattctgg
tcgaaaatatcttcattttgtcactttcaaactccactgcagacacatttcattcttcgaaacgggtgatgtatccgtagaaatctttg
ctgaacctttcaattgaaagttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttctgaacattgaacatcat
agaactaagcaggaaaacttcgcacacgtgctgccagaagaacaactttcaacatgtttaaggacaatttctgtattcaaa
20 gcatgactctatgcctttggggcctgaatcgggtgcgctagatttctctttcatgggttctactaatgtctacgggtataccggaac
atgcgacgtcgctaaggctgatggacattcaggtggaaaggaaccctacaggcttttcaacgttgatgtctttgagtacaac
atcgggtaccagccaaccaatgtacgggtcgtatccattcatgttttcattcttcgtccacatctatcttttgggtcaatgcagctga
cacttgggtagacataaagtatgacaccagtaaaaataaaacgatgactcattggatctccgaaaatgggtgcatagatgtag
tcatgtccctggggccagatattccaactatcattgacaaatttaccgatttgactggtagaccctttttaccgcccatttcctcta
25 tagggtaccatcaatgtagatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattcc
ttacgattttatttgggttgacttggagtatacgaacgacaaaaaatattttacttggaaagcagcactcctttcccaatccaaaa
ggctgttatccaaattaaaaaagttgggtagaaatcttgcgtactaatcgatcctcatttaagaaagattatgaaatcagtgac
agggtaatatgaaaatgtagcagtcaggatcacatggaaatgactatgtaggtcattgctggccaggtaattctatatgg

attgataccataagcaaatatggccaaaagatttgaagtccttttcgaacggtttatggatctgccggctgatttaactaattta
ttcatttggatgatatgaacgagccttcgatttcgatggcccagagaccacagctccaaaagatttgattcacgacaattac
attgaggaaagatccgtccataacatatatgggtctatcagtgcataagctacttacgacgcaataaaatcgattattcaccat
ccgataagcgtcctttccttctaacaagggccttttttgcgggtctcaacgtactgctgccacatggactggtgacaatgtggc
5 caattgggattacttaagatttccattcctatgggtctgtcaaacaacattgctggatgccatttataggagccgacatagctg
gctttgctgaggatcctacacctgaattgattgcacgttggtagcaagcgggcttatggtagccatttttagagcacacgccc
atatagacaccaagagaagagaaccatacttattcaatgaacctttgaagtcgatagtacgtgatattatccaattgagatatt
cctgctacctaccttatacaccatgtttcataaatcaagtgctactggatttccgataatgaatccaatgtttattgaacacctga
atttgctgaattgtatcatatcgataaccaattttactggagtaattcaggtctattagtcacaaacctgtcacggagcctggtaac
10 agaaacggaaatggtttcccacccgggtatattctatgaattcgcactctttacactctttataaacaatggtagctgattgataga
aaagaatatttctgcaccattggataaaattccattatttattgaaggcgggtcacattatcactatgaaagataagtatagaagat
cttcaatgttaatgaaaaacgatccatatgtaatagttatagccctgataccgaggagcagagccgttggagatctttatgttga
tgatggagaaacttttggctaccaaaagaggtagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcg
aagtcataattcccgagaatttgacaggcattcaccacaatactttgaggaataccaatattgaaaaatcattatcgcaaagaat
15 aatttaacacacataacgttgaagacagtaftaaagtcaaaaaaatggcgaagaaagttcattgccgactagatcgctc
atatgagaatgataataagatcaccattcttaacctatcgcttgacataactgaagattgggaagtatttttgggcc

SEQ ID NO: 23 pYPT1ADE1glsII

gtcgaccggctgcattaatgaatcggccaacgcgcggggagaggcggttgcgtattgggcgctcttccgcttcctcgctca
20 ctgactcgctgcgctcggctcgttcggctgcggcgagcgggtatcagctcactcaaaggcggtaatacggttatccacagaatc
aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctgg
cgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagga
ctataaagataccaggcggtttccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtcc
gcctttctcccttcgggaagcgtggcgcttttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagc
25 tgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagccaaccggtaag
acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctt
gaagtggtaggcttaactacggctacactagaaggacagtatttgggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttggtagctcttgatccggcaaacaaccaccgctggtagcggtgggtttttgtttgcaagcagcagattacgcgcaga

aaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaaacgaaaactcacgttaagggattttg
gtcatgagattatcaaaaaggatcttcacctagatccttttaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaa
acttgggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgcctgactc
ccgctcgtgtagataactacgatacgggagggccttaccatctggccccagtgtgcaatgataccgcgagaccacgctca
5 ccggctccagattatcagcaataaaccagccagccggaagggccgagcgcagaagtggctcctgcaactttatccgcctcc
atccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttggtgccattgctacag
gcatcgtgggtgtcacgctcgtcgtttgggtatggcttcattcagctccgggttcccaacgatcaaggcgagttacatgatcccca
tgttgtgcaaaaaagcggtagctccttcgggtcctccgatcgtgtcagaagtaagtggccgcagtggtatcactcatggttat
ggcagcactgcataattcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattctgag
10 aatagtgtatgcgggcgaccgagttgctcttggccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagt
gctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaaccactcg
tgcaccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaa
agggataagggcgacacggaaatgttgaatactcatactcttcttttcaatagctccaaggcaacaaattgactactcaga
ccgacattcattcgttattgattttaaatcaacgataaacggaatggttacttgaatgatttcactttatgatcattgtttactaattac
15 ctaaataaggattttatatggaattggaagaataagggaaattcagatgtctgaaaaaggcgaggagggtactaatcattcaa
gccatttcttgccagtaattgcttcataagcttcaatatacttttcttactcttgatagcaatttctgcatccatggctacgccctct
ttgccattcaatccgttggccgtcaaccaatctctgagaaactgcttatcgtaactctcttgcgattaccacttggttaagtctttt
gattccaaaatctagaagaatctggagttaaaacttcatctactagtaccaattcattgttttctccagttccaaattcgaatttctg
atcagcaataatgatccccttcaaaaggcggaagtttttgcagcagaatacaactcgaccgccttgacagcgaccttctcac
20 aaatgtctttacctacaatctcagcagcttgttcaatagagatgttttcacgtgttcaccctgttcagctttcgttgaagggtgtgaa
aatcggagttggaaaggcgtcgtctcttgaagggtctcgttttcaaccttgactccatggacagttttgagttcttgtactctttc
catgcacttccagtgatgtaacctctgacaatggcttcaaagggtatcagttctgtgcttttttactatcaaggatcgtccctctaat
tgagatttgatttttcttcagacagttttgatggtagtaaagcaaagacttcttgcattagaagcaaccaaattgattctttatgta
gggtgccaaaaaatcaaaccagaaaactgagagctgagtcaaaatctttcccttatcaggaataccgtttgtcataatcacatc
25 gtaagcggagatacggtcagttgcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttcctttggcgatt
aaaggtaggattccgtccagatcagttgcacaatggacatacttgaaggatacagcaaaagtgtgttggaagcgatgacac
atggaaaggaattttcaggtttcctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttc
aaagcgcccaagtgtttctgccaacctgactttttctgaataatgaatcgttcaagtggagttattaaacctatgattaagtacgt

gatttggcactggataaggtcgaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggag
aagcttagtaacagttatcatctcttggtcgattaacgcttacaatttccattcgccattcaggctgcgcaactgttgggaaggg
cgatcgggtgcgggcctcttcgctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaaga
ggagtgggaaataccaagaaaaacatcaaactcgaatgatttcccaaaccctaccacaagatattcatcagctgcgagat
5 aggctgatcaggagcaagctcgtacgagaagaaacaaaatgacaaaaaaatcctatactatataggttacaaataaaaa
gtatcaaaaatgaagcctgcatctctcaggcaaattggcattctgacatcctcttgattagaatctagcaagaccggctctctcgt
aagtgcccaacttgaactgaggaacagtcattgtctaaggctacaaactcaatgatgatgatgatggtcgacggcgctatt
cagatcctctcttgagatgagttttgttcgggccccaaaaataacttccaatcttcagttatgtcaagcgatagggttaagaatgg
tgatcttattatcattctcatatgacgatctagtcggcaatgaactttcttcgccatttttttgactttaatactgtctttcaacgttatg
10 ttgtgttgtaaattattctttgcgataatgatttttcaatattggatttctcaaagtattgtggtgaatgcctgtcaaattctcgga
atatgacttcgaacatttttaaggtattgtttcgaaaatgaactgagtttctacgtactcacctctttggtagccaaaagttctcc
atcatcaacataaagatctccaacggctcgtccctcggtatcaggggctataactattacatatggatcgttttcattaacattg
aagatcttctatacttattctttcatagtataatgtgaccgccttcaataaataatggaattttatccaatggtgcagaaatattctt
tctatcaaatcagtaccattgtttataaaagagtgtaaagatgcgaattcatagaatataccgggtgggaaaaccatttccgttc
15 tgattgaccagggtccgtgacaggttgactaatagacctgaattactccagtaaaattggttatcgatatgatacaattcagca
aattcagggtgttcaataaacattggattcattatcggaatccagtacacttgatttatgaaacatggtgtataaggtaggta
gcaggaaatatctcaattggataatatcacgtactatcgacttcaaagggttcattgaataagtatggttctcttcttgggtgtctat
atgggcgtgtgctctaaaaaatgggtaccataagcccgcttggtaccaacgtgcaatcaattcagggtgtaggatcctcagca
aagccagctatgtcggctcctataaatggcataccagcaatgttgtttgacagaaccataggaatggaaatctttaagtaatcc
20 caattggccacattgtcaccagtccatgtggcagcagtagcttgagagccggcaaaaaagcccttgtagaaggaaagga
cgcttatcggtatggtgaataaatcgattttattgcgtcgttaagtagcttcatgcactgatagaccatataatgttatggacggatctt
tctcaatgtaattgtcgtgaatcaaactttttggagctgtggtctctgggccatcgaaaatgaaggctcgttcatatcattcca
aatgaataaattagttaaatcagccggcagatccataaacggttcgaaaaaggacttccaaatcttttggccatatttgcttatg
gtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccattgtgatccttgactgctacattttcattaatta
25 ccctgtcactgatttcataatctttctttaaattgaggatcgattagtagcacaagatttctaccaactttttaattggataacagc
cttttggattgggaaaggagtgtgcttccaagtaaaatatttttgcgttcgtatactccaagtccaaccaataaaaatcgtaa
ggaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatcattataattccatctacattgatggtaccctat
agaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaattgtcaatgatagttggaatatctggccccaggggac

atgactacatctatgacaccattttcggagatccaatgagtcacggtttatctttactggtgtcatactttatgtctacccaagtgtc
agctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgggatcgaaccgtacattggttggtggtacc
gatgttgactcaaagacatcaacgttgaaaagcctgtagggttcctttccacctgaagtgtccatcagccttagcgacgtcgc
atgttccggtataccgtagacattagtagaaccatgaaagagaaatctagcgcaaccgattcaggcccaaaggcataga
5 gtcattgcttgaatacaagaaattgtccttaacatgttgaaagtgtttcttctggcagcacgtgtgcgaagtttctgcttagtt
ctatgatgttcaatgttcaggaaattttgctcgtttacaataagtgttcagcgcattttgccagtaaactttcaattgaaaagggttcag
caaagatttctacggatacatcacggtttcgaagatgaaatgtgtctgcagtggagtttgaaagtgacaaaaatgaagatatttt
cgaccagaatgagttcacagtttgttttgcttaaggaagtgggaattgtggaatactggtcctgttcgcctcctcttgaaatttctt
gtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaaaccgctgctgttggttggtcattctctctttctcat
10 ttatagtgaaccttactgagtgatcctgtaaaaagagagagagaatgggaactgaacggctatatcatcgccctccaatctt
ggtatagttttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaatagactcggcgtccactttgtaatagcagt
gatgagatttggaataatttctgcataaacctgtttctatggcaaaaccagattgcgcacacttcttaatagatagtcggtta
aacgcatgcgaaaaagcggtaaagaagaccaattggcatacgagccatttcaaaaggaccatctcgaggtaccgatccga
gacggccgggtggggccacgtgaattcgactgctattatctctgtgtgtatgtgtgtattgggcaacaagagctgtgaatctgt
15 ccttataaaagacacccgaagaggcaagattgggtagtagactaattagtggagggcaactggttaggggttttggtggc
ttattatagtgtcagcgatactatacaatctacgatgccgatgtcggcatacagcacattattagatccaaaatttccttaatttc
ctattcacgttatataattaaccagaatttatccgttgaattgcttaatcgacttcttccacgggttggtccgctgtctccgcctggg
gcagcaccacctggggcaccgccagctgcgccagctccgccagggaacctccagggaagccaccaggagctccacc
aggagctccaccagcagctccatagaacttgetcattattgggttagcaacttcttccaattcctttgcctgtccttgactcgtc
20 tgtggaagcagattgtgactcatcatagatctgatctcat

SEQ ID NO: 24 pGAPZAgIsIIHDEL

tcgagatggtccttttgaaatggctcgtatgccaaattggtcttctttaccgctttttcgcattgcgtttaccgactatctattaagaa
gtgtgcgcaatctgggttttgccatagaaacagggtttatgcagaaaatattgccaaatctcatcactgctattacaaagtga
25 cgccgagtctattgcacacgatcctttagagaatgtgcttcatgctaccataattaaaactataccaagattggagggcgatga
tatagccgttcagttccattctctctctctttttacaggatcactcagtaagggttactataaatgagaaagagagaatgccaa
ccaacagcagcgggttggtgatctcttcacaacgggttcaatgagacctggaagtacgcattcgacaagaaattcaagagga
ggcgaacaggaccagtattccacaattccacttccctaagcaaaaacaaactgtgaactcattctggtcgaaaatatcttcatt

ttgtcactttcaaactccactgcagacacatttcattcttcgaaacggatgatccgtagaaatctttgctgaacctttcaattg
aaagttactggcaaaatgcgctgaaacttattgtaaacgagcaaaatttctgaacattgaacatcatagaactaagcagga
aaacttcgcacacgtgctgccagaagaacaactttcaacatgtttaaggacaatttctgtattcaagcatgactctatgcctt
tggggcctgaatcgggtgcgctagatttctcttcattgggttctactaatgtctacgggtataccggaacatgcgacgtcgctaag
5 gctgatggacacttcaggtggaaaggaaccctacaggctttcaacgttgatgtctttgagtacaacatcgggtaccagccaac
caatgtacgggtcgatcccattcatgttttcattcttcgtccacatctatctttgggtcaatgcagctgacacttgggtagacataa
agtatgacaccagtaaaaaataaaacgatgactcattggatctccgaaaatgggtgcatagatgtagtcatgtccctggggcca
gatattccaactatcattgacaaatttaccgatttgactggtagacccttttaccgcccatttctctatagggtaccatcaatgta
gatggaattataatgatgagatggacgttctcacagtggactctcagatggatgctcatatgattccttacgattttatttggttg
10 acttgagatatacgaacgacaaaaaatatttacttgaagcagcactcctttcccaatccaaaaaggctgttatccaaattaa
aaagtgggtagaaatcttgcgtactaatcgatcctcatttaaagaaagattatgaaatcagtgcagggtaattaatgaaaat
gtagcagtcaaggatcacaatggaaatgactatgtaggtcattgctggccaggtaattctatatggattgataccataagcaaa
tatggccaaaagatttgaagtccttttcgaacgggttatggatctgccggctgatttaactaatttattcatttgaatgatatga
acgagccttcgatttcgatggcccagagaccacagctccaaaagatttgattcacgacaattacattgaggaaagatccgtc
15 cataacatatatggctatcagtgcataagctacttacgacgcaataaaatcgatttattcaccatccgataagcgtcctttcct
tctaacaagggttttttgcgggtctcaacgtactgctgccacatggactgggtgacaatgtggccaattgggattacttaag
atttcattcctatggttctgtcaacaacattgctgggtatgccatttataggagccgacatagctggctttgctgaggatcctac
acctgaattgattgcacgttgggtaccaagcgggcttatggtacccatttttagagcacacgcccatatagacaccaagagaa
gagaaccatacttattcaatgaaccttgaagtcgatagctgatattatccaattgagataattcctgctacctaccttataca
20 ccatgtttcataaatcaagtgtcactggatttccgataatgaatccaatgtttattgaacaccctgaatttgcgaattgtatcatat
cgataaccaattttactggagtaattcaggtctattagtaaaacctgtcacggagcctgggtcaatcagaaacggaaatggtttc
ccaccgggtatattctatgaattcgcattctttacactctttataaacaatgggtactgatttgatagaaaagaataatttctgcacat
tggtataaaattcattatttattgaaggcgggtcacattatcactatgaaagataagtatagaagatcttcaatgttaatgaaaaac
gatccatataatagttatagccctgataccgaggggacgagccgttgagatctttatgttgatgatggagaaacttttggt
25 accaaagaggtgagtacgtagaaactcagttcattttcgaaaacaataccttaaaaaatgttcgaagtcatttcccgagaattt
gacaggcattcaccacaatactttgaggaataccaatattgaaaaaatcattatcgcaaagaataattacaacacaacataac
gttgaaagacagtattaaagtcaaaaaaatggcgaagaaagttcattgccgactagatcgatcatatgagaatgataataaga
tcaccattcttaacctatcgcttgacataactgaagattgggaagtatttttggggccgaacaaaaactcatctcagaagagg

atctgaatagcgccgtcgaccacgacgaactgtgagtttagccttagacatgactgttcctcagttcaagttgggcacttacg
agaagaccggtcttgctagattctaataagaggatgtcagaatgccattgcctgagagatgcaggcttcattttgatactttt
ttatttgtaacctatatagtataggattttttgtcattttgtttcttcgtacgagcttgctcctgatcagcctatctcgcagctgat
gaatatcttgtaggggttgggaaaatcattcgagtttgatgttttcttggtatttccactcctcttcagagtacagaagatt
5 aagtgagaccttcgttggtgcggatccccacacaccatagcttcaaaatgtttctactcctttttactcttcagattttctcgga
ctccgcgcacgcgtaccacttcaaaacacccaagcacagcatactaaatttccctctttctcctctagggtgtcgtaatta
cccgtactaaagggttgaaaagaaaaagagaccgcctcgtttcttttcttcgtcgaaaaaggcaataaaaattttatcacg
tttcttttcttgaaatttttttagtttttctcttcagtacctccattgatatttaagtaataaacggtcttcaatttctcaagttc
agtttcattttcttggtctattacaacttttttacttcttggtcattagaaagaaagcatagcaatctaataaggcggtgttgac
10 aattaatcatcggcatagtatatcggcatagtataatacgacaagggtgaggaactaaacctaggccaagtgaccagtgccgt
tccggtgtcaccgcgcgcgacgtcgccggagcggctcgagttctggaccgaccggctcgggttctccgggacttcgtgg
aggacgacttcgccggtgtggtccgggacgacgtgacctgttcacgcgcggtccaggaccaggtggtgccggacaa
cacctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaacttcggg
gacgcctccgggcccggccatgaccgagatcggcgagcagccgtggggggcgggagttcgccctgcgcgacccggccg
15 gcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggccacgggtcccaggcctcggagat
ccgtcccccttttcttgatcatgtaattagttatgtcacgcttacattcacgcctccccccacatccgctctaaccgaa
aaggaaggagttagacaacctgaagtctaggtccctattttttatagttatgtagtattaagaacgttatttatatttcaaatt
tttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaagggtttgggacgctcgaaggct
ttaatttgcaagctggagaccaacatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggccgcgttgctgg
20 cgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagga
ctataaagataccaggcggttccccctggaagctccctcgtgcgctctcctgtccgacctgccgcttaccggatacctgtcc
gcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagc
tgggctgtgtgcacgaacccccgttcagcccagccgctgcgccttatccggtaactatcgtcttgagtccaacccggttaag
acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgtacagagttctt
25 gaagtggtagcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttgtagctcttgatccggcaaaacaaaccaccgctggttagcggtggtttttgttgcaagcagcagattacgcgcaga
aaaaaaggatctcaagaagatccttgatctttctacggggtctgacgctcagtggaaacgaaaactcacgttaagggtttg
gtcatgcatgagatcagatctttttgtagaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgt

tgcaactccgaacgacctgctggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctc
ttcccttctctctctccaccgcccgttacgctccctaggaaattttactctgctggagagcttcttctacggcccccttgacgc
aatgctcttcccagcattacgttgcgggtaaaacggaggctcgtgtacccgacctagcagcccagggatggaaaagtcccg
gccgtcgtggcaataatagcgggcggacgcatgtcatgagattattggaaccaccagaatcgaatataaaaggcgaac
5 accttcccaattttggttctctgacccaaagactttaatttaatttattgtccctatttcaatcaattgaacaactatttcgaaa
cgaggaattcacgtggcccagccggccgtctcggtacgtacc

SEQ ID NO: 25 pGAPADE1glsIIHDEL

cgtactcacctctttggtagccaaaagtttctccatcatcaacataaagatctccaacggctcgtccctcgggtatcaggggctat
10 aactattacatatggatcgttttcattaacattgaagatcttctatacttattcttcatagtataatgtgaccgccttcaataaataa
tggaattttatccaatggtgcagaaatattctttctatcaaatacgtaccattgtttataaaagagtgtaaagatgcgaattcatag
aatataccgggtgggaaaaccatttccgtttctgattgaccaggctccgtgacagggttgactaatagacctgaattactccag
taaaattggttatcgatatgatacaattcagcaaattcagggtgttcaataaacattggattcattatcggaaatccagtacact
tgatttatgaaacatggtgtataaggtaggttagcaggaaatatctcaattggataatatcacgtactatcgacttcaaagggttcat
15 tgaataagtatggttcttctcttgggtgtctatatgggcgtgtgctctaaaaaatgggtaccataagcccgcttgggtaccaacgt
gcaatcaattcagggtgtaggacctcagcaaagccagctatgtcggctcctataaatggcataaccagcaatgttggttgacag
aaccataggaatggaaatctttaagtaatcccaattggccacattgtcaccagtccatgtggcagcagtacgttgagagccg
gcaaaaaaagcccttgttagaaggaaaggacgcttatcggatggtgaataaatcgattttattgcgtcgtagtagcttcatgc
actgatagaccatatatgttatggacggatcttctcctcaatgtaattgtcgtgaatcaaacttttggagctgtggtctctgggcca
20 tcgaaaatcgaaggctcgttcatatcattccaaatgaataaattagttaaatcagccggcagatccataaaccgttcgaaaaag
gacttccaaatcttttggccatatttgcttatggtatcaatccatatagaattacctggccagcaatgacctacatagtcatttccat
tgtgatccttgactgctacattttcattaattaccctgtcactgatttcataatcttctttaaataaggatcgattagtagcacaaga
tttctaccaactttttaatttgataacagccttttggattgggaaaggagtgtgcttccaagtaaaatatttttgcgttcgta
tactccaagtccaaccaataaaaatcgtaaggaaatcatatgagcatccatctgagagtccactgtgagaacgtccatctcatc
25 attataattccatctacattgatgttacccctatagaggaaatgggcggtaaaaagggtctaccagtcaaatcggtaaatttgta
atgatagttggaatatctggccccaggacatgactacatctatgacaccattttcggagatccaatgagtcacgttttattttta
ctggtgtcactttatgtctaccaagtgtcagctgcattgacccaaaagatagatgtggacgaagatgaaaacatgaatgg
gatcgaaccgtacattggttggtgtggtaccgatgtgtactcaaagacatcaacgttgaaaagcctgtagggttcccttccacc

tgaagtgtccatcagccttagcgacgtcgcatgttccggtataccgtagacattagtagaaccatgaaagagaaatctagc
gcaaccgattcaggcccaaaggcatagagtcagcttgaatacaagaaattgtccttaacatgttgaaagttgtttctctg
gcagcacgtgtgcgaagtttctgcttagttctatgatgttcaatgttcaggaaatttgctcgttacaataagtttcagcgcatt
ttgccagtaaactttcaattgaaaagggttcagcaaagatttctacggatacatcaccgtttcgaagatgaaatgtgtctgcagtg
5 gagtttgaaagtgacaaaaatgaagatatcttcgaccagaatgagttcacagtttgttttgcttaaggaagtggaaattgtggaat
actggtcctgttcgcctcctcttgaaatttctgtcgaatgcgtacttccaggtctcattgaaccgttgtgaagagatcaacaac
cgctgctgttggttggcattctctctttctcatttatagtgaaacttactgagtgatcctgtaaaaaagagagagagaatgggaa
ctgaacggctatatcatcgccctccaatcttggtatagtttaattatggtagcatgaagcacattctctaaaggatcgtgtgcaa
tagactcggcgtccactttgtaatagcagtgatgagatttggaatatcttctgcataaacctgtttctatggcaaaaccagat
10 tgcgcacacttcttaatagatagtcggtaaacgcatgcgaaaaagcggtaaagaagaccaattggcatacagagccatttca
aaaggaccatctcgaggtaccgatccgagacggccggctggggccacgtgaattcctcgtttcgaaatagttgttcaattgatt
gaaatagggacaaataaattaaatttaaagtctttgggtcaggagaaaccaaattgggaaagggtgttcgccttttatattcgat
tctggtggtttcaataatctcatgacatgcgtccgcccgctattattgccagcgacggccgggacttttccatccctgggctg
ctaggtcgggtacacgacctccgttttaccgcaacgtaatgctgggaagagcattgctgcaagggggccgtagaagaag
15 ctctccagcagagtaaaatttctagggacggtaacggggcgggtggaaggagagagaagggaagagacgtttctggttcca
ttactccacatttaagttttaccccggaagaattttacgttgccagcaggtcgttcggagttgcaacggagccagatatttcagag
atggctacctgattggacgaggacaccaagacatttctacaaaaagatctgatctcatcgaccggctgcattaatgaatcgg
ccaacgcgcggggagagggcggttgcgtattgggcgctcttccgcttctcgtcactgactcgtcgctcggtcgttcgg
ctgcggcgagcggtatcagctcactcaaaggcggttaatacggttatccacagaatcaggggataacgcaggaaagaacat
20 gtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcggttttccataggctccgccccctg
acgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcggttccccct
ggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggc
gctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgtccaagctgggctgtgtgcacgaacccccct
tcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggttaagacacgacttatcgccactggcagca
25 gccactggtaacaggattagcagagcgaggtatgtaggcgggtgctacagagttcttgaagtgggtggcctaactacggctac
actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa
acaaaccaccgctggtagcggtgggtttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctt
tgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggttttgggtcatgagattatcaaaaaggatctt

cacctagatccttttaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttgggtctgacagttaccaatgctta
atcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatac
gggagggccttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataa
accagccagccggaagggccgagcgcagaagtggctctgcaactttatccgcctccatccagtctattaartgttgccggg
5 aagctagagtaagtagttcggcagttaatagtttgcgcaacggttgccattgctacaggcatcgtgggtgcacgctcgtcgtt
tggtatggcttcattcagctccgggtcccaacgatcaaggcgagttacatgatcccccattgttgcaaaaaagcggtagctc
cttcggctcctccgatcgttgctcagaagtaagtggccgcagtggtatcactcatggttatggcagcactgcataattctcttactg
tcatgccatccgtaagatgcttttctgtgactgggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagtt
gctcttggccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacgttcttcg
10 gggcgaaaactctcaaggatcttaccgctgttgagatccagttcgtatgaaccactcgtgcaccaactgatcttcagcatct
tttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataaggggcgacacggaa
atgttgaatactcatactcttcttttcaatagctccaaggcaacaattgactactcagaccgacattcattcgttattgatttaa
atcaacgataaacggaatggttacttgaatgattcactttatgatcattgttactaattacctaataaggattttatatggaattgg
aagaataagggaatttcagatgtctgaaaaaggcgaggagggtactaatcattcaagcccatttcttgccagtaattgcttca
15 taagcttcaatacttttcttactcttgatagcaattctgcatccatggctacgccctcttggccattcaatccgttgccgctcaa
ccaatctctgagaaactgcttategtaactctcttgcgatttaccacttggttaagtcttttgattccaaaatctagaagaatctgg
agttaaaacttcatctactagtagtaaccattcattgtttcgtccagtccaaattcgaatttcgtatcagcaataatgatcccttcaa
aggcggaagtttttgagcagaatacaactcgaccgccttgacagcgaccttctcacaatgtctttacctacaatctcagca
gcttgttcaatagagatgtttcatcgtgttcaccctgttcagctttcgttgaagggtgtgaaaatcggagttggaaaggcgtcgtc
20 ctcttgaagggttctcgtttcaaccttgactccatggacagttttgagttctgtactctttccatgcacttccagtgtatgaacctc
tgacaatggcttccaaagggtatcagtcgtgtgttttactatcaaggatcgtccctctaattgagatttgatttttcttcagacagt
tttgatggtagtaaagcaaagacttccctgtcattagaagcaaccaaatgattctttatgtagggtgccaaaaaatcaaaccag
aaaactgagagctgagtcaaaatcttcccttatcaggaataccggttgataatcacatcgtaagcggagatacgggtcagtt
gcgacgaacagcaagttgttctcatcgactgcataaatgtctctaacctttccttggcgattaaaggtaggattccgtccagat
25 cagtgttcacaatggacatacttgaaggatacagcaaagtgtgttgaagcgatgacacatggaaaggaattttcaggtt
cctagagtagtatattggggcggtgaaagttcagatgtttaatgcttaatactcttatactcttcaaagcgcccaagtgttctgc
caacctgactttttctgaataatgaatcgttcaagtggagttatctaaccatgattaagttacgtgatttggcactggataaggtc
gaaaaatatccgtattcataaacgattattggtaaaagttacaaaataccactaattacggagaagcttagtaacagttatcatct

cttggtcgattaacgcttacaatttcattcgccattcaggctgcgcaactgttgggaagggcgatcgggtgcgggcctcttcg
ctattacgccagggcctcgaggcacaaacgaacgtctcacttaatcttctgtactctgaagaggagtgggaaataccaagaa
aaacatcaaactcgaatgatttcccaaaccctaccacaagatattcatcagctgcgagataggctgatcaggagcaagct
cgtacgagaagaacaaaatgacaaaaaaatcctatactatataggttacaataaaaaagtatcaaaaatgaagcctgca
5 tctctcaggcaaatggcattctgacatcctcttgattagaatctagcaagaccggctcttctcgtaagtgcccaactgaactgag
gaacagtcattgtctaaggctaaaactcacagttcgtcgtggctgacggcgctattcagatcctctctgagatgagttttgttc
gggcccacaaaataacttcccaatcttcagttatgtcaagcgatagggttaagaatggtgatcttattatcattctcatatgacgatc
tagtcggcaatgaactttcttcgccattttttgactttaatactgtctttcaacgttatgttggtgtgtgtaaattattctttgcgataatg
atctttcaatattggtattcctcaaagtattgtggtgaatgcctgtcaaattctcggaatatgacttcgaacatttttaaggtattg
10 ttttcgaaaatgaactgagtttcta

SEQ ID NO: 26

GACGAGATCTTTTTTTCAGACCATATGACCGG

15 SEQ ID NO: 27

GCGGAATTCTTTTCTCAGTTGATTTGTTTGT

SEQ ID NO: 28

20 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

25 atggccctcttctcagtaagagactgttgagattaccgtcattgcagggtcggttattgttctcctcctaacattgaattccaac
agtagaactcagcaatatattccgagttccatctccgctgcattgattttacctcaggatctatatcccctgaacaacaagtcatt
ctctgaggaaaatgatgctaaaaaattagagcaaagtgtctgaattcagaggcaagcgaagactccgaagccatggatga
agaatccaaggctctgaaagctgccgctgaaaaggcagatgccccgatcgacactaaaacaacctatggattatatcactcc
atcttttgctaacaagctggtaagccaaaagcttggtacgtcactttggtgagaaacaaggagttgaaaggttgctaagctc
30 cattaaatatgtggaaaacaaaattaacaagaaattcccatatccttgggttttctaaacgatgaaccttttactgaagaattca

aggaagcagtcaccaaagctgtttctccgaagttaagtttggtattttgcceaaggaacattggtcatatcctgaatggattaat
caaaccaaggctgctgaaattcgtgcagatgctgccaccaatacatatacgggtggctccgaatcttatagacacatgtgtcg
ttaccaatctgggttttctggagacatgaattattagaagagtacgattgggtactggcgtgtggaaccagacatcaagttatac
tgtgatattaattacgacgttttaagtggatgcaagaaaacgaaaaagtttacggctttaccgtttctattcatgaatatgaagtg
5 acgatcccaacactatggcaaagctccatggatttcatcaaaaagaaccccgaaacttagatgaaaacaacctgatgagttt
tcttcgaacgataatggtaaaacatacaatctgtgccatttctggtcaaactttgaaattgcaaacttgaatttgtggagggtcac
cagcctacagagagtattttgacactttggatcatcaagggtggatttttctacgaaagatggggcgatgctcccggttcattctatt
gctgctgctttgttttgcceaaggataaaatccattattttcagacattgggtaccatcatccaccttatgataactgccattgg
acaaggagggtctataacagtaacaactgtgaatgtgaccaaggtaatgattcactttccaagggtactcttgtggttaaggaat
10 attatgatgctcaagggttggttaaagccaaaaaactggaaaaaattccgtgagtagaaatcttgaacatactgtttctttgttt
gactttatactttctatttatattttattttataactggtaagtacacataggactgcgtatcaaacatataagtgaggcaatccac
atttttttaagattcgaatattttattctcattagcgtattccgagaatagttcgaaaaaataaggtatatcaagagttttaca
agtgagaggaaagaggaataagctataagcaaaaaagcgtaaaaaaattagctgaagacatagaactatggatgttctca
aagaggtgtgtcactagaccaagataaattgaccagctgaaggaaacgagccgagataaaacaaatgaaacggatgatc
15 ctttgaaaactatttgaaggattgtaaatttaaagcgccttcaaacaagatcagtcaccatttgctaaactaaatcattacag
gaaactcattctaacaatgaagcggctattaatataattatctcctcaattgattgattacttaaccgaattcactaatagggtatcaa
attacacacaagatttagacttcattaaaaaaaagtccaatgaattacagtcattgctcgaatacaactccactaaactggcaca
tatctctcctatggttaatgatttgatgattcctcctgaactcattgatgacatcattaaagggaagatcaatgaaagctggcagg
ataatataacattcatagcagataaagaagaaatttataacaagtataggccaataatctcgatcaagacaacaaggacgca
20 gaaaattcagcaatgctagcaccaaaggattttgataagttatgtcaactcctggacatcctaaaaaatgttattctaga

SEQ ID NO: 31 Human mannosidase II

aaaatgaagttaagccgccagttcacctgttcggcagtgcgatcttctgtgtggtgattttctcgctctacctgatgctggacc
ggggtcacttagactaccccaggaacccgcgcgcgagggctccttcctcagggccagctctcaatgttgcaagaaaaa
25 atagaccatttggagcgtttgctagctgagaataatgagatcatctcaaatttagagactcagtcattatgagtgagtcgt
ggaggatgggtccgaaaagttcacaagcaatttcagccaagggtgctggctcacatcttctgccctcacaattatccctctcagt
tgacactgcagactgtctgtttgcttcacaaagtgaagtcacaattcagatgtgcagatgttggtggttacagtctaattcttt
tgacaatccagatggtggagtttgaagcaaggatttgacattacttatgaatctaataatgggacactgaaccccttcaagt

ctttgtggtgcctcattcccataacgacccagggttggtgaagactttcaatgactacttttagagacaagactcagtatatttttaa
taacatgggtcctaaagctgaaagaagactcacggaggaagttatttgggtctgagatctcttacctttcaaagtgggtgggatatt
atagatattcagaagaaggatgctgttaaaagttaatagaaaatggtcagcttgaaattgtgacaggtggctgggttatgcct
gatgaagctactccacattatttgccttaattgatcaactaattgaaggacatcagtggtggaaaataatagggagtgaac
5 ctcgggtccggctgggctattgatcccttggacactaccaacaatggcttatcttctaaaccgtgctggactttctcacatgctt
atccagagagttcattatgcagttaaaaaacactttgcactgcataaaacattggagtttttggagacagaattgggatctgg
gatctgtcacagatattttatgccacatgatgcccttctacagctatgacatccctcacacttgggacctgatcctaaaatatgc
tgccagtttgattttaaacgtcttctggaggcagatttgggtgtccctggggagtccccccagaaacaatacatcctggaaat
gtccaaagcagggctcggatgctactagatcagtaccgaaagaagtcaaagcttttgcaccaaagttctcctggctccact
10 aggagatgattccgctactgtgaatacacggaatgggatttacagtttaagaattatcagcagcttttgattatatgaattctca
gtccaagtttaaagttaagatacagtttggaaactttatcagattttttgatgcgctggataaagcagatgaaactcagagagac
aagggccaatcgatgttccctgttttaagtggagatttttacttatgccgatcgagatgatcattactggagtggctattttaca
tccagacccttttacaacgaatggacagaatcatggaatctcatttaagggctgctgaaattctttactatttcgccctgagac
aagctcacaatacaagataaataatttctctcatcatcactttacacggcactgacagaagccagaaggaatttgggactgt
15 ttcaacatcatgatgctatcacaggaactgcaaaagactgggtggttggattatggtaccagacttttcttcgttaatggtt
tggaagaagataattggaaattctgcatttctcttattgggaaggacaaactcacatacgactcttactctctgataccttctg
gagatggatttgaaacaaaaatcacaagattctctgccacaaaaaatataaaggctgagtgccggagccaaggtaccttg
tggtctataatcctttagaacaagaccgaatctcgttgggtctcagtctatgtgagttccccgacagtgcaagtgttctctgcttca
ggaaaacctgtggaagtcaagtcagcgcagtttgggatacagcaaatactatttcagaaacagcctatgagatctcttttga
20 gcacataaccgccattgggactgaaagtgtataagattttggaatcagcaagttcaaattcacatttagctgattatgtcttgtat
aagaataaagtagaagatagcggaaatttcaccataaagaatatgataaatactgaagaaggtataacactagagaactcctt
tgtttacttcggtttgatcaaaactggacttatgaagcaaattgatgactaaagaagatggtaaacaccatgaagtaaattgtcaa
tttcatgggtatggaaccacaattaaaagagacaaaagtgggtgcctacctcttcttacctgatggtaatgccaagccttatgttta
cacaacaccgcccttgcagagtgcacatggaaggatttattcgggaagtgacttgctttttgaccatgttactcatagagtc
25 cgactataccacatacagggaaatagaaggacagtctgtggaagtttccaatattgtggacatccgaaaagtatataaccgtga
gattgcaatgaaaatttctctgatataaaaagccaaaatagattttatactgacctaaatgggtaccagattcaacctagaatga
cactgagcaaattgcctcttcaagcaaattgtctatcccatgaccacaatggcctatatccaggatgccaaacatcgtttgacac
tgctctctgctcagtcattaggggtttcgagtttgaatagtgggtcagattgaagttatcatggatcgaagactcatgcaagatga

taatcgtggccttgagcaaggtatccaggataacaagattacagctaattctatttogaataactactagaaaaagaagtgtgt
 taatacggagaagaagaagaagtcggtcagttatccttctctccttagccacataacttcttctcatgaatcatccagtcattc
 caatggcaaataagttctcctcacctacccttgagctgcaaggtgaattctctccattacagtcattcttgccttgtagacattcatc
 tggtaatttgagaacaatacagtcacaaggtgggcaatgggcactccaatgaggcagccttgatcctccacagaaaagggtt
 5 tgattgtcggttctctagcaaaaggcacagggtgtttgttctactactcagggaagatattggtacagaaaacttttaacaag
 ttattgtcgaaagttctcacaccttcatcactatccttgatgcattcacctcccggcactcagaatataagtgagatcaacttgag
 tccaatggaaatcagcacattccgaatccagttgaggtgaacctgacttcacattggattgagaatcattggcctttataccttt
 ctgggttgacgtgcaataaagaagcacattattttagcttctggctactgtgagaacatgaattctgtgattctgtgggtttttctt
 ttttcttttaccagtacagtaaga

10 SEQ ID NO: 32 sense primer

CGCGGATCCATGGCCAAAAAGTTCACAAAGCAATTTC

SEQ ID NO: 33 antisense primer

GTGTCCCATTTCATTAGATTC

SEQ ID NO: 34 human GlcNAc-transferase I

15 ATGCTGAAGAAGCAGTCTGCAGGGCTTGTGCTGTGGGGCGCTATCCTCTTT
 GTGGCCTGGAATGCCCTGCTGCTCCTCTTCTTCTGGACGCGCCCAGCACCT
 GGCAGGCCACCCTCAGTCAGCGCTCTCGATGGCGACCCCGCCAGCCTCAC
 CCGGGAAGTGATTTCGCTGGCCCAAGACGCCGAGGTGGAGCTGGAGCGGC
 AGCGTGGGCTGCTGCAGCAGATCGGGGATGCCCTGTCGAGCCAGCGGGGG
 20 AGGGTGCCACCGCGGGCCCCTCCCGCCCAGCCGCGTGTGCCTGTGACCCCC
 GCGCCGGCGGTGATTCCCATCCTGGTCATCGCCTGTGACCGCAGCACTGTT
 CGGCGCTGCCTGGACAAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTC
 CCCATCATCGTTAGCCAGGACTGCGGGCACGAGGAGACGGCCCAGGCCAT
 CGCCTCCTACGGCAGCGCGGTACGCACATCCGGCAGCCCGACCTGAGCA
 25 GCATTGCGGTGCCGCCGGACCACCGCAAGTTCCAGGGCTACTACAAGATC
 GCGCGCCACTACCGCTGGGCGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTC
 CCCGCGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCCCCGGACTTCTT
 CGAGTACTTTCGGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTG

GTGCGTCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCA
GCAGGCCTGAGCTGCTCTACCGCACCGACTTTTTCCCTGGCCTGGGCTGGC
TGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCC
TTCTGGGACGACTGGATGCGGGCGGCCGGAGCAGCGGCAGGGGGCGGGCCTG
5 CATACGCCCTGAGATCTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGA
GCCACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGCTGAACCAGC
AGTTTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCT
ATGACCGAGATTTCTCTCGCCCGCGTCTACGGTGCTCCCCAGCTGCAGGTGG
AGAAAGTGAGGACCAATGACCGGAAGGAGCTGGGGGAGGTGCGGGTGCA
10 GTATACGGGCAGGGACAGCTTCAAGGCTTTCGCCAAGGCTCTGGGTGTCA
TGGATGACCTTAAGTCGGGGGGTTCCGAGAGCTGGCTACCGGGGTATTGTC
ACCTTCCAGTTCCGGGGGCCGCGTGTCCACCTGGCGCCCCCACCAGCGTGG
GAGGGCTATGATCCTAGCTGGAATATGCTGAAGAAGCAGTCTGCAGGGCT
TGTGCTGTGGGGCGCTATCCTCTTTGTGGCCTGGAATGCCCTGCTGCTCCTC
15 TTCTTCTGGACGCGCCCAGCACCTGGCAGGCCACCCTCAGTCAGCGCTCTC
GATGGCGACCCCGCCAGCCTCACCCGGGAAGTGATTCGCCTGGCCCAAGA
CGCCGAGGTGGAGCTGGAGCGGCAGCGTGGGCTGCTGCAGCAGATCGGGG
ATGCCCTGTGCGAGCCAGCGGGGGAGGGTGCCCACCGCGGGCCCCTCCCGCC
CAGCCGCGTGTGCCTGTGACCCCGCGCCGGCGGTGATTCCCATCCTGGTC
20 ATCGCCTGTGACCGCAGCACTGTTTCGGCGCTGCCTGGACAAGCTGCTGCAT
TATCGGCCCTCGGCTGAGCTCTTCCCCATCATCGTTAGCCAGGACTGCGGG
CACGAGGAGACGGCCCAGGCCATCGCCTCCTACGGCAGCGCGGTACGCA
CATCCGGCAGCCCGACCTGAGCAGCATTGCGGTGCCGCCGGACCACCGCA
AGTTCCAGGGCTACTACAAGATCGCGCGCCACTACCGCTGGGCGCTGGGC
25 CAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGAC
CTGGAGGTGGCCCCGGACTTCTTCGAGTACTTTCGGGCCACCTATCCGCTG
CTGAAGGCCGACCCCTCCCTGTGGTGCGTCTCGGCCTGGAATGACAACGG
CAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCACCG

ACTTTTTCCCTGGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGC
TGGAGCCCAAGTGGCCAAAGGCCTTCTGGGACGACTGGATGCGGCGGCCG
GAGCAGCGGCAGGGGCGGGCCTGCATACGCCCTGAGATCTCAAGAACGAT
GACCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTTCTTTGACCAGCACCT
5 CAAGTTTATCAAGCTGAACCAGCAGTTTGTGCACTTCACCCAGCTGGACCT
GTCTTACCTGCAGCGGGAGGCCTATGACCGAGATTTCTCGCCCGCGTCTA
CGGTGCTCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGACCGGAAGG
AGCTGGGGGAGGTGCGGGTGCAGTATACGGGCAGGGACAGCTTCAAGGCT
TTCGCCAAGGCTCTGGGTGTCATGGATGACCTTAAGTCGGGGGGTTCCGAGA
10 GCTGGCTACCGGGGTATTGTCACCTTCCAGTTCCGGGGCCGCGTGTCCAC
CTGGCGCCCCCACCAGCGTGGGAGGGCTATGATCCTAGCTGGAAT

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

MALFLSKRLLRFTVIAGAVIVLLLTLNSNSRTQQYIPSSISAADFDTSGSISPEQQVISEENDAKKL
15 EQSALNSEASEDSEAMDEESKALKAAAEKADAP

SEQ ID NO: 36 pPIC6AKrecoGnTI

GAAATTTTTTTTTTTAGTTTTTTTCTCTTTCAGTGACCTCCATTGATATTAA
GTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCT
20 ATTACAACTTTTTTTACTTCTTGTTTCATTAGAAAGAAAGCATAGCAATCTA
ATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATA
GTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGCCTTTGTCTCA
AGAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCC
CCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCC
25 GCATCTTCACTGGTGTCAATGTATATCATTTTTACTGGGGGACCTTGTGCAG
AACTCGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTT
GTATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTGAGCCCCTGCGGA
CGGTGCCGACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTG

AAGGACAGTGATGGACAGCCGACGGCAGTTGGGATTCGTGAATTGCTGCC
CTCTGGTTATGTGTGGGAGGGCTAAGCACTTCGTGGCCGAGGAGCAGGAC
TGACACGTCCGACGGCGGCCCCACGGGTCCCAGGCCTCGGAGATCCGTCCC
CCTTTTCCTTTGTCGATATCATGTAATTAGTTATGTCACGCTTACATTCACG
5 CCCTCCCCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTG
AAGTCTAGGTCCCTATTTATTTTTTTTATAGTTATGTTAGTATTAAGAACGTT
ATTTATATTTCAAATTTTTCTTTTTTTTTCTGTACAGACGCGTGTACGCATGT
AACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCT
TTAATTTGCAAGCTGGAGACCAACATGTGAGCAAAAGGCCAGCAAAAGGC
10 CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCC
CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC
TTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
15 GGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA
GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
20 GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
GTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATCAGATCTAA
CATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCAC
25 AGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTA
GCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACC
CACTTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCG
CTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTG

TCTATCCTGGCCCCCCTGGCGAGGTTTCATGTTTGTTTATTTCCGAATGCAAC
AAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGT
GGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAACACTGACAGTTTAAA
CGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAA
5 CTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTT
CCAAAAGTCGGCATAACCGTTTGTCTTGTTTGGTATTGATTGACGAATGCTC
AAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCC
CGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTTGGATG
ATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACT
10 GCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTT
GACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTT
TATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTTCCAATTGACAAG
CTTTTGATTTTAAACGACTTTTAAACGACAACCTTGAGAAGATCAAAAAACAAC
TAATTATTCGAAACGAGGAATTATCCATATTCGCAAGCAGTTCCACTCGAA
15 AGCATGGCCCTCTTTCTCAGTAAGAGACTGTTGAGATTTACCGTCATTGCA
GGTGCGGTTATTGTTCTCCTCCTAACATTGAATTCCAACAGTAGAACTCAG
CAATATATTCCGAGTTCCATCTCCGCTGCATTTGATTTTACCTCAGGATCTA
TATCCCCTGAACAACAAGTCATCTCTGAGGAAAATGATGCTAAAAAATTA
GAGCAAAGTGCTCTGAATTCAGAGGCAAGCGAAGACTCCGAAGCCATGGA
20 TGAAGAATCCAAGGCTCTGAAAGCTGCCGCTGAAAAGGCAGATGCCCCGC
CGGCGGTGATTCCCATCCTGGTCATCGCCTGTGACCGCAGCACTGTTTCGGC
GCTGCCTGGACAAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTCCCCA
TCATCGTTAGCCAGGACTGCGGGCACGAGGAGACGGCCCAGGCCATCGCC
TCCTACGGCAGCGCGGTACGCACATCCGGCAGCCCGACCTGAGCAGCAT
25 TGCGGTGCCGCCGGACCACCGCAAGTTCCAGGGCTACTACAAGATCGCGC
GCCACTACCGCTGGGCGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTCCCCG
CGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCCCCGGACTTCTTCGAG
TACTTTCGGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTGC

GTCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAG
GCCTGAGCTGCTCTACCGCACCGACTTTTTCCCTGGCCTGGGCTGGCTGCT
GTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCCTTCT
GGGACGACTGGATGCGGGCGGCCGGAGCAGCGGCAGGGGGCGGGCCTGCAT
5 ACGCCCTGAGATCTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGAGCC
ACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGCTGAACCAGCAGT
TTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCTATG
ACCGAGATTTCTCTCGCCCGCGTCTACGGTGCTCCCCAGCTGCAGGTGGAGA
AAGTGAGGACCAATGACCGGAAGGAGCTGGGGGAGGGTGCGGGTGCAGTA
10 TACGGGCAGGGACAGCTTCAAGGCTTTCGCCAAGGCTCTGGGTGTCATGG
ATGACCTTAAGTCGGGGGTTCGAGAGCTGGCTACCGGGGTATTGTCACCT
TCCAGTTCCGGGGGCCCGGTGTCCACCTGGCGCCCCCACCAGCTGGGAG
GGCTATGATCCTAGCTGGAATTAGCACCTGTCGACTGGAGACCTGCAGGC
ATGCAAGCTTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGAC
15 ATGACTGTTCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCT
AGATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGG
CTTCATTTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTT
GTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAG
CTGATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTTCGAGTTTGATGT
20 TTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGAC
CTTCGTTTGTGCGGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCC
TTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCA
AAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGTG
TCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTC
25 GTTTCTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTT
TTCTT

SEQ ID NO: 37 pGAPKreManII

ATCCATATTCGCAAGCAGTTCCACTCGAAAGCATGGCCCTCTTTCTCAGTA
AGAGACTGTTGAGATTTACCGTCATTGCAGGTGCGGTTATTGTTCTCCTCC
TAACATTGAATTCCAACAGTAGAACTCAGCAATATATTCCGAGTTCCATCT
CCGCTGCATTTGATTTTACCTCAGGATCTATATCCCCTGAACAACAAGTCA
5 TCTCTGAGGAAAATGATGCTAAAAAATTAGAGCAAAGTGCTCTGAATTCA
GAGGCAAGCGAAGACTCCGAAGCCATGGATGAAGAATCCAAGGCTCTGA
AAGCTGCCGCTGAAAAGGCAGATGCCCCGCCAAAAAGTTCACAAAGCAAT
TTCAGCCAAGGTGCTGGCTCACATCTTCTGCCCTCACAATTATCCCTCTCA
GTTGACACTGCAGACTGTCTGTTTGCTTCACAAAGTGGAAGTCACAATTCA
10 GATGTGCAGATGTTGGATGTTTACAGTCTAATTTCTTTTGACAATCCAGAT
GGTGGAGTTTGGAAGCAAGGATTTGACATTACTTATGAATCTAATGAATG
GGACACTGAACCCCTTCAAGTCTTTGTGGTGCCCTATTCCCATAACGACCC
AGGTTGGTTGAAGACTTTCAATGACTACTTTAGAGACAAGACTCAGTATAT
TTTTAATAACATGGTCCTAAAGCTGAAAGAAGACTCACGGAGGAAGTTTA
15 TTTGGTCTGAGATCTCTTACCTTTCAAAGTGGTGGGATATTATAGATATTC
AGAAGAAGGATGCTGTAAAAGTTTAATAGAAAATGGTCAGCTTGAAATT
GTGACAGGTGGCTGGGTTATGCCTGATGAAGCTACTCCACATTATTTTGCC
TTAATTGATCAACTAATTGAAGGACATCAGTGGCTGGAAAATAATATAGG
AGTGAAACCTCGGTCCGGCTGGGCTATTGATCCCTTTGGACACTCACCAAC
20 AATGGCTTATCTTCTAAACCGTGCTGGACTTTCTCACATGCTTATCCAGAG
AGTTCATTATGCAGTTAAAAAACACTTTGCACTGCATAAAACATTGGAGTT
TTTTTGGAGACAGAATTGGGATCTGGGATCTGTACAGATATTTTATGCCA
CATGATGCCCTTCTACAGCTATGACATCCCTCACACTTGTGGACCTGATCC
TAAAATATGCTGCCAGTTTGATTTTAAACGTCTTCCTGGAGGCAGATTTGG
25 TTGTCCCTGGGGAGTCCCCCAGAAACAATACATCCTGGAAATGTCCAAA
GCAGGGCTCGGATGCTACTAGATCAGTACCGAAAGAAGTCAAAGCTTTTT
CGTACCAAAGTTCTCCTGGCTCCACTAGGAGATGATTTCCGCTACTGTGAA
TACACGGAATGGGATTTACAGTTTAAGAATTATCAGCAGCTTTTTGATTAT

ATGAATTCTCAGTCCAAGTTTAAAGTTAAGATACAGTTTGGAACTTTATCA
GATTTTTTTTGATGCGCTGGATAAAGCAGATGAACTCAGAGAGACAAGGG
CCAGTCGATGTTCCCTGTTTTAAGTGGAGATTTTTTCACTTATGCCGATCGA
GATGATCATTACTGGAGTGGCTATTTTACATCCAGACCCTTTTACAAACGA
5 ATGGACAGAATCATGGAATCTCATTTAAGGGCTGCTGAAATTCTTTACTAT
TTCGCCCTGAGACAAGCTCACAAATACAAGATAAATAAATTTCTCTCATCA
TCACTTTACACGGCACTGACAGAAGCCAGAAGGAATTTGGGACTGTTTCA
ACATCATGATGCTATCACAGGAACTGCAAAGACTGGGTGGTTGTGGATT
ATGGTACCAGACTTTTTTCATTCGTTAATGGTTTTTGGAGAAGATAATTGGAA
10 ATTCTGCATTTCTTCTTATTTTGAAGGACAAACTCACATACGACTCTTACTC
TCCTGATACCTTCCTGGAGATGGATTGAAACAAAAATCACAAGATTCTCT
GCCACAAAAAATATAATAAGGCTGAGTGCGGAGCCAAGGTACCTTGTGG
TCTATAATCCTTTAGAACAAGACCGAATCTCGTTGGTCTCAGTCTATGTGA
GTTCCCCGACAGTGCAAGTGTTCTCTGCTTCAGGAAAACCTGTGGAAGTTC
15 AAGTCAGCGCAGTTTGGGATACAGCAAATACTATTTTCAGAAACAGCCTAT
GAGATCTCTTTTCGAGCACATATACCGCCATTGGGACTGAAAGTGTATAAG
ATTTTGGAAATCAGCAAGTTCAAATTCACATTTAGCTGATTATGTCTTGTAT
AAGAATAAAGTAGAAGATAGCGGAATTTTCACCATAAAGAATATGATAAA
TACTGAAGAAGGTATAAÇACTAGAGAACTCCTTTGTTTTACTTCGGTTTGA
20 TCAAACCTGGACTTATGAAGCAAATGATGACTAAAGAAGATGGTAAACACC
ATGAAGTAAATGTGCAATTTTCATGGTATGGAACCACAATTAAGAGAC
AAAAGTGGTGCCTACCTCTTCTTACCTGATGGTAATGCCAAGCCTTATGTT
TACACAACACCGCCCTTTGTCAGAGTGACACATGGAAGGATTTATTCGGA
AGTGACTTGCTTTTTTTGACCATGTTACTCATAGAGTCCGACTATACCACAT
25 ACAGGGAATAGAAGGACAGTCTGTGGAAGTTTCCAATATTGTGGACATCC
GAAAAGTATATAACCGTGAGATTGCAATGAAAATTTCTTCTGATATAAAA
AGCCAAAATAGATTTTATACTGACCTAAATGGGTACCAGATTCAACCTAG
AATGACACTGAGCAAATTGCCTCTTCAAGCAAATGTCTATCCCATGACCAC

AATGGCCTATATCCAGGATGCCAAACATCGTTTGACACTGCTCTCTGCTCA
GTCTTTAGGGGTTTCGAGTTTGAATAGTGGTCAGATTGAAGTTATCATGGA
TCGAAGACTCATGCAAGATGATAATCGTGGCCTTGAGCAAGGTATCCAGG
ATAACAAGATTACAGCTAATCTATTTTGAATACTACTAGAAAAAAGAAGT
5 GCTGTTAATACGGAAGAAGAAAAGAAGTCGGTCAGTTATCCTTCTCTCCTT
AGCCACATAACTTCTTCTCTCATGAATCATCCAGTCATTCCAATGGCAAAT
AAGTTCTCACCTACCCTTGAGCTGCAAGGTGAATTCTCTCCATTACAGTCA
TCTTTGCCTTGTGACATTCATCTGGTTAATTTGAGAACAAATACAGTCAAAG
GTGGGCAATGGGCACTCCAATGAGGCAGCCTTGATCCTCCACAGAAAAGG
10 GTTTGATTGTCGGTTCTCTAGCAAAGGCACAGGGCTGTTTTGTTCTACTACT
CAGGGAAAGATATTGGTACAGAACTTTTAAACAAGTTTATTGTGCGAAAG
TCTCACACCTTCATCACTATCCTTGATGCATTCACCTCCCGGCACTCAGAAT
ATAAGTGAGATCAACTTGAGTCCAATGGAAATCAGCACATTCCGAATCCA
GTTGAGGTGAACCTGACTTTCACATTTGGATTGAGAATCATTGGCTTTTAT
15 ACCTTTCTTGGTTTGACGTGCAATAAAGAAGCACATTATTTTAGCTTCTGG
CTACTGTGAGAACATGAATTCTGTGATTCTGTGGGGTTTTTTCTTTTTCTTTT
ACCAGTACAGTAAGAAAAAAAAAAAAAAAAAAAAAACTCGAGCCGCGGC
GGCCGCCAGCTTGGGCCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAA
TAGCGCCGTCGACCATCATCATCATCATTGAGTTTTAGCCTTAGACAT
20 GACTGTTCCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAG
ATTCTAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTT
CATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTGTC
ATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTG
ATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTGAGTTTGATGTTTT
25 TCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGACCTT
CGTTTGTGCGGATCCATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGA
GCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC

TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTG
TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
5 GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
10 CGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT
CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCA
CCTAGATCCTTTTAAATTA AAAAATGAAGTTTTAAATCAATCTAAAGTATAT
ATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA
15 TCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGT
GTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA
TGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC
20 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGG
TCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT
TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
25 TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCG
GCGACCGAGTTGCTCTTGCCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACT

CGTGCACCCAACTGATCTTCAGCATCTTTTACTTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGAC
ACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATGCTTCTAGTGGTAG
GAATTAATTCTGTACCGGTTTACAGAAGGACGACTCTTGATGCGCCAACCA
5 CAGTGACAATAGACATAGAGGAAATGACAAAAGGATTATGCGAGGATGCT
GCTGGAGACGATTCAAAGTTTAGTTTAGAAAGGTCCTCCATTTATGCTGAT
AGAATACTAGATAACCGTGAACCTTTGTCTCAGGAGATCCGCATCAGACGA
AGGATGTTCCGACCTGCAAATAATCGAAGAAGAGACCCCTAGGCCGTACG
ACAGTCAGTTAGTAGATATTTATACCATTCTGCGAGAAGGTCCCCCAGTAT
10 AGATAAGAGAGTTTCAATTGTATGTACTATAACAGTTTATTGTATCATTAA
CAAACCTCAGTATAAAACCTTCAAGACTCTAGCTTTTCATTCAGTGCATCCA
ATTGCTTTAAACAGCAGTCTTAGAAGTTCCTCCGGTGGCAGTTCTTTTTTC
AACACTGGCTTCAAAGTCAAACGTTGAAGCCACATCAGCCTCAAAACGGG
AGTCAATGGATTTCAATTGTTCGAGGGACAACCTGATCAATACCAGAAAGG
15 TTCAACTCCTCGGCTTGTCTGACACATTCACCAGAAATGTGGTGAGTTTCT
CTGAATGGAACCTCCCCTTCTAACTAAATAGTCGGCAAGATCTGTAGCCAGC
ATATCCATAGTTAGAGCATTCTTCATTCGTTCCGGCATCAATGTTCAAGGTA
GAAACTACACCGGATGCTATCAAAATCGAGTGCTCTACAGTGATTAGAGT
ATCAAATAAAGGCTCCTTATCCTCTTGCAATATCTTTGTTATAGGTTGACGG
20 AATGGACTTAATAGACATGAGGAAACCAGCCAAGGCCCCAAAACATCTAC
CAGATTTACCCCTCAATAACTCCAAAGAGTCTGGGTTTTTTTTTTTGAGGCA
TCAGAGAAGATCCAGTAGAATAAGCATCTGCCAACTTGATAAATCCAAAC
TCTCCAGTGGAGTAAATGATCAAATCTTCTGAGAATCGAGAAATATGATTC
ATAAACAACGAAGACCAGAACATGGTTTTCGACTACAAAATCTCTGTCTGA
25 AACAGCGGCCAAAGAATTACCAATAACAGAATCAAACCCTAATCTCTCAG
CAATGTATTCACGATCAATTCCATAAGGATGACCAGCCAAAGCTCCAGCTC
CCAATGGGGATTTGTTCAACCTTTTAACGATTTGATTCAGTCTCTCATAATC
TTCAGTGAAATAGGTAGCATACATGCTCAACCAGTGAGACCATCTGATTG

GTTGAGCTCTTTGCAAGTGAGTATAACCGGGCATCAAGACGTCTATTTTCCT
G TTCAGCTCTCTTGATGATTACTTGAATGAACTGCTTCAGATAGTCAGCCA
ACTGAGTTAGATTGTCTCTGACATACAACCGCATATCAGTGGCAACTTGAT
CATTCTAGACCTTCCGGTATGAACCTTACCAGAGATTCCACGACCAATCA
5 ACTCACCCAAGCGACGTTTATTAGCAGTGTGAATATCCTCATCCCCTGGCT
TCTCAACAACTTCCCTTCTGCCCCTCTGCTTCAATCAATTTGAGACCAC
GATGAATCTCACTTAGTTCGTCTTTAGTTAGCAAATTAATTTTCTCCAGGCC
CTCAGTGTAACCTTTTGTTCCTTCTAAATCCACCTTGTACATTTTCTTGTCG
TAAGGTAAGGAAGCGTTATACAAATCCATCAAGGGGGTCAGTAGCCCCAGT
10 AAACCTGCCACCCACAGTTTAAGTCTTTCTTCTTGATTTCGACATAGATAG
CTGGTAATAAGTTTAGAACAAAAGGAAAGAGAAGGTAGAATATAGGTGA
AAGAATTGCCAAAAGAGGCAGCGCGGTAAAAAGAAAAATGAATCATCGA
AATTAGTCTAGATTCTCGTATTTCTGTTGGTTATTGTATTATCTAATCAGGG
TAAACACCTAAGGCTTACATCTCCAGTGAAACGTGGGTATAATGAACA
15 GTCTTTCACAGTGAATCTGTCGCACAACCATGCTAAGATACGTTCCGTTCC
TAGACCGTAACCACCGTGAGGGCAGATTAACGCTTACAATTTCCATTTCGCC
ATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGGGCACTAGTGAGATCAGATCTTTTTTTGTAGAAATGTCTTGG
TGTCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAA
20 CTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAACTTAAAT
GTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCCTTCCACCGC
CCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGC
CCCCTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGT
CGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTG
25 GCAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGA
ATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGGTTTCTCCTGACCCA
AAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACAACCTAT
TTCGAAACGAGGAATTCACGTGGCCC

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

MALFLSKRLLRFTVIAGAVIVLLLTLNSNSRTQQYIPSSIAAFDFTSGSISPEQQV
ISEENDAKKLEQSALNSEASEDSEAMDEESKALKAAAEKADAPIDTKTTMDYI
5 TPSFANKAGKPKACYVTLVRNKELKGLLSSIKYVENKINKKFPYPWVFLNDEP
FTEEFKEAVTKAVSSEVKFGILPKEHWSYPEWINQTKAAEIRADAATKYTYGG
SESYRHMCRYQSGFFWRHELLEEDWYWRVEPDIKLYCDINYDVFKWMQEN
EKVYGFTVSIHEYEVTIPTLWQTSMDFIKKNPEYLDENNLMSFLSNDNGKTYN
LCHFWSNFEIANLNLWRSPAYREYFDTLDHQGGFFYERWGDAPVHSIAAALF
10 LPKDKIH YFSDIGYHHPPYDNCPLDKEVYNSNNCECDQGNDFTFQGYSCGKE
YYDAQGLVKPKNWKKFRE

SEQ ID NO: 39 primer Kai5

AAggAgTTAgACAACCTgAAgTCTA
15

SEQ ID NO: 40 primer Kai7

AAggAgTAgAAACATTTTgAAgCTA

SEQ ID NO: 41 primer Kai10.

20 CCATggCgAAggCAgATggCAgTT

SEQ ID NO: 42 primer Kai11

AgATCTTTAgTCCTTCCAACCTTCCTT

25 SEQ ID NO: 43 primer Kai13

ATCTAAGCTATATTCgCCgTTTCTgTCATT

SEQ ID NO: 44 primer Kai14

CgATTATggATgTTAgATCTgATCTCATgA

SEQ ID NO: 45 pZ5'PpOCHITrunc

tcattgagatcagatctaacaatccaTAATCGATCTAAgctatattcgccgtttctgtcatttgcgtttgtacggaccct
5 cacaacaattatcatctccaaaaatagactatgatccattgacgctccgatcacttgattgaagactttggaagctccttcaca
gttgagtccaggcaccgtagaagataatcttgaagacaattggagtttcattttccttaccgcagttacgaacctttcccca
catatttggcaaacgtggaaagtcttccctctgatagttcctttccgaaaaacttcaaagacttaggtgaaagtggctgcaa
ggccccaaattatgatcattttgtgatacccgatgatgcagcatgggaacttattcaccatgaatacgaacgtgtaccagaag
tcttggaaagcttttgattttaacgactttaacgacaacttgagaagatcaaaaaacaactaattattcgcgaaacgaggaattc
10 acgtggcccagccggcgtctcggtacgtcgagccgcggcgccagcttctagagaacaaaaactcatctc
agaagaggatctgaatagcgcgtcgaccatcatcatcatcatcattgagttgtagccttagacatgactgttcctcagttcaa
gttgggcacttacgagaagaccggtcttctagattctaataagaggatgtcagaatgccatttgcctgagagatgcaggct
tcatttttgatactttttatttgaacctatatagtaggattttttgtcattttgttcttctcgtacgagcttgctcctgatcagcct
atctcgagctgatgaatatcttggtaggggttgggaaaatcattcgagttgatgttttcttggtatttccactcctcttca
15 gagtacagaagattaagtgagaccttcgttgtcggtatccccacacaccatagcttcaaatgtttctactcctttttactctt
ccagattttctcgactccgcgcacgtcgccgtaccacttcaaacacccaagcacagcactaaattttcctcttttcttcteta
gggtgtcgttaattaccggtactaaagggttggaaaagaaaaagagaccgcctcgtttcttttcttctcgtaaaaaaggcaat
aaaaattttatcacgtttcttttcttgaaatttttttagtttttcttctcagtgacctccattgatatttaagttaataaacgggtct
tcaatttctcaagtttcagtttcatttttcttcttattacaacttttttacttcttgttcattagaaagaaagcatagcaatctaacta
20 aggggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgaacagggtgaggaactaaaccatggcca
agttgaccagtgcggttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttc
tcccgggacttctgtggaggacgacttcgccggtgtggtccgggacgacgtgacctgttcatcagcgcggtccaggacca
ggtggtgccggacaacacctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggagggtcgtg
tccacgaacttccgggacgcctccgggcccggccatgaccgagatcggcgagcagccgtgggggcccggagttcgccctg
25 cgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggccacgggtcc
caggcctcggagatccgtcccccttttcttctgtcgatatcatgtaattagttatgtcacgcttacattcacgccctccccccaca
tccgctctaaccgaaaaggaaggagtttagacaacctgaagtctaggtccctattttttatagttatgttagtattaagaacg
ttatttatatttcaaatttttcttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggttttg

gacgctcgaaggcttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcggttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgctt
accggatacctgtccgcctttctcccttcgggaagcggtggcgctttctcaatgctcacgctgtaggtatctcagttcgggttag
5 gtcgttcgctccaagctgggctgtgtgcacgaacccccggtcagcccgaccgctgcgccttatccggtaactatcgtcttga
gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg
gtgctacagagttcttgaagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagcca
gttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggttagcggtggtttttgtttgcaagcag
cagattacgcgcagaaaaaaaggatctcaagaagatccttgatctttctacggggcttgacgctcagtggaacgaaaactc
10 acgttaagggttttgg

SEQ ID NO: 46 pZMFManHDEL5'PpOCH1Mut

ggtcgactctagaatggcgaaggcacatggcagtttgctctagtataatcctcacatccaccagaaggattacttctacat
15 ggctatattcgccgtttctgtcatttgcgtttgtacggaccctcacacaattatcatctccaaaaatagactatgatccattgac
gctccgatcacttgatttgaagactttggaagctccttcacagttgagtcagggcaccgtagaagataatcttgaagacaatt
ggagtttcatttcttaccgcagttacgaacctttccccaacataattggcaaacgtggaaagtttctccctctgatagttcctt
ccgaaaaactcaaagacttaggtgaaagttggctgcaaaggtccccaattatgatcatttgtgatacccgatgatgcagc
atgggaacttattcaccatgaatacgaacgtgtaccagaagcttgggaagctttgattttaacgactttaacgacaacttgag
20 aagatcaaaaaacaactaattattcgcgaaacgaggaattcacgtggcccagccggcgtctcggatcggtacctcgagcc
gcggcgccgcccagctttctagagaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatc
atcattgagttgtagccttagacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaac
aagaggatgtcagaatgccattgcctgagagatgcaggcttcattttgatactttttatttgtaacctatatagtataggatttt
tttgcattttgttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttggtagggggttgggaaa
25 atcattcgagttgatgttttcttggtatttccactcctcttcagagtacagaagattaagtgagaccttcgtttgtcggatcctt
ttttagaaaatgtcttggtgtcctcgtccaatcaggtagccatctctgaaatatctggctccgttgcaactccgaacgacctgt
ggcaacgtaaaattctccggggtaaaacttaaatgtggagtaatggaaccagaaacgtctcttcccttctctctccttccaccg
cccgttaccgtccctaggaaatttactctgctggagagcttcttctacggcccccttgacgaatgctcttccagcattacgtt

gcgggtaaacggaggctgtgtacccgacctagcagcccagggatggaaaagtcggccgctgctggcaataatagcg
ggcggacgcatgtcatgagattattggaaccaccagaatcgaatataaaaggcgaacaccttcccaattttggtttctctg
acccaaagactttaatttaatttattgtccctatttcaatcaattgaacaactatttcgcgaaacgatgagatttcttcaatttta
ctgctgttttattcgcagcatcctccgcattagctgctccagtcaacactacaacagaagatgaaacggcacaattccggct
5 gaagctgtcatcggttactcagattagaaggggatttcgatgttgctgtttgccattttcaacagcacaataacgggttatt
gtttataaatactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaagagaggctgaagctgaattcg
ccacaaaacgtggatctcccaaccctacgagggcggcagcagtcaaggccgcatccagacgtcgtggaacgcttacca
ccattttgcctttcccatgacgacctccaccgggtcagcaacagctttgatgatgagagaaacggctggggctcgtcggca
atcgatggcttggaacacggctatcctcatgggggatgccgacattgtgaacacgatccttcagtatgtaccgcagatcaactt
10 caccacgactgcggttgccaaccaaggatcctccgtgttcgagaccaacattcggctacctcggctggcctgctttctgcctatg
acctgttgcgaggtcctttcagctccttggcgacaaaccagacccctggtaaacagccttctgaggcaggctcaaactggc
caacggcctcaagggtgcgttcaccactcccagcgggtgtcccgaccctaccgtcttctcaaccctactgtccggagaagt
ggtgcatctagcaacaacgtcgtgaaattggaagcctgggtgctcgagtggacacgggtgagcgacctgacgggaaaccc
gcagtatgccagcttgcgcagaagggcgagtcgtatctctgaatcaaagggaagcccggaggcatggcctggcctg
15 attggaacgtttgtcagcacgagcaacgggtacctttcaggatagcagcggcagctggctccggcctcatggacagcttctacg
agtacctgatcaagatgtacctgtacgacctgggttgcgtttgcactacaaggatcgctgggtccttgggtccgactcgacc
attgggcatctcggctctcaccgctcgacgcgcaaggacttgaccttttgccttcgtacaacggacagtctacgtcgccaaa
ctcaggacatttgccagttttggcgggtggcaacttcatttgggaggcattctcctgaacgagcaaaagtacattgactttgg
aatcaagcttgccagctcgtactttggcaçgtacaccagacggcttctggaatcgccccgaaggcttcgctgggtgga
20 cagcgtgacgggcgccggcggtcgcgccctcgtcccagtcgggttctactcgtcggcaggattctgggtgacggca
ccgtattacatcctgcggccggagacgctggagagcttgactacgcataccgcgtcacgggcgactccaagtggcagga
cctggcgtgggaagcgttgagtgccattgaggacgcatgccgcgccggcagcgcgtactcgtccatcaacgacgtgacg
caggccaacggcgggggtgcctctgacgatatggagagcttctggtttgccgaggcgtcaagtatgcttacctgatctttg
cggaggagtcggatgtgcaggtgcaggccaccggcgggaacaaattgtctttaacacggaggcgcacccctttagcatc
25 cgttcacatcacgacggggcggccaccttgctcacgacgagttgtaatctagggcggccgacgtttctagaacaaaaa
ctcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttagccttagacatgactgttcctc
agttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgctgagagat
gcaggcttcatttttgatactttttatttgaacctatatagtataggattttttgtcattttgtttctctcgtacgagcttgctcctga

tcagcctatctcgcagctgatgaatatcttgtaggggttgggaaaatcattcgagttgatgttttcttggtattcccactc
ctcttcagagtacagaagattaagtgagaccttcgttgtgcggatccccacacaccatagcttcaaaatgttctactcctttt
tactctccagatttctcggactccgcgcacgcgcgtaccacttcaaaacaccaagcacagcactaaatttccctctttct
tctctaggggtgtcgttaattacctactaaagggttggaaaagaaaaagagaccgcctcgtttcttttcttcgtcgaaaaa
5 ggcaataaaaaattttatcacgtttcttttcttgaaatttttttagtttttctttcagtgcctccattgatatttaagttaataa
acggcttcaatttctcaagttcagtttcaatttcttggtctattacaacttttttacttctgttcattagaaagaaagcatagcaat
ctaataagggcggtgttgacaattaatcatcggcatagtatatcggcatagtataatacgacaaggtgaggaactaaacca
tgccaagtgtaccagtgccgttccgggtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctc
gggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgacctgttcacgcgcggtcca
10 ggaccaggtggtgccggacaacacctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcgga
ggtcgtgtccacgaacttccgggacgcctccggggccggccatgaccgagatcggcgagcagccgtgggggcccggagtt
cgccctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggccca
cgggtcccaggcctcggagatccgtcccccttttctttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctcc
ccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttttttatagttatgttagtatt
15 aagaacgttatattatattcaaatTTTTTTTTTctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaa
ggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaaggccaggaacc
gtaaaaaggccgcttgctggcgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcaga
ggtggcgaaacccgacaggactataaagataaccaggcggttccccctggaagctccctcgtgcgctctctgttccgacct
gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttttcaatgctcacgctgtaggtatctcagttcg
20 gtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccagccgtgcgccttatccggtaactatc
gtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatg
taggcggtgctacagagttcttgaagtggtaggcttaactacggctacactagaaggacagtatttggtatctgcgctctgctg
aagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaaccaccgctggtagcgggtgtttttgtttgc
aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacgggggtctgacgctcagtggaaacg
25 aaaactcacgttaagggttttggtcatgc

PCT\EP03\06711
Flanders Interuniversity Institute for Biotechnology

98936/m8
25.05.2004

CLAIM SET

1. A strain of a methylotrophic yeast for producing proteins having a mammalian-like glycosylation pattern, wherein said strain expresses a GlcNAc-transferase I or a functional part thereof.
2. The strain of claim 1, wherein said GlcNAc-transferase I or said functional part thereof is localized at a site in the secretory pathway of said strain, and wherein the glycoproteins at said site have the oligosaccharide structure of $\text{Man}_5\text{GlcNAc}_2$.
3. The strain of claim 2, wherein said site is the Golgi apparatus.
4. The strain according to any one of claims 1-3, which produces proteins having $\text{GlcNAcMan}_5\text{GlcNAc}_2$.
5. The strain according to any one of claims 1-4, wherein said methylotrophic yeast is *Pichia*.
6. The strain according to any one of claims 1-5, wherein said strain also expresses an α -1,2-mannosidase or a functional part thereof.
7. The strain of claim 6, wherein said α -1,2-mannosidase or said functional part thereof is localized in the endoplasmic reticulum (ER) of said strain.
8. The strain according to any one of claims 1-7, wherein said strain is incapable of making a functional protein encoded by the OCH1 gene.

9. The strain according to any one of claims 1-8, where said strain also expresses a mannosidase II or a functional part thereof.

10. The strain according to claim 9, wherein said mannosidase II or said functional part thereof is localized in the Golgi apparatus of said strain.

11. The strain according to claim 10, wherein said strain produces glycoproteins having an oligosaccharide structure with fewer than five mannose residues.

12. A strain of a methylotrophic yeast for producing glycoproteins having a mammalian-like glycosylation pattern, wherein said strain expresses a mannosidase II or a functional part thereof.

13. The strain of claim 12, wherein said mannosidase II or said functional part thereof is localized in the Golgi apparatus of said strain.

14. The strain according to claim 12 or 13, wherein said strain produces glycoproteins having an oligosaccharide structure with fewer than five mannose residues.

15. The strain according to any one of claims 12-14, wherein said methylotrophic yeast is *Pichia*.

16. The strain according to any one of claims 12-15, wherein said strain also expresses at least one of (a) an α -1,2-mannosidase or a functional part thereof; or a GlcNAc-transferase I or a functional part thereof.

17. The strain according to claim 16, wherein said α -1,2-mannosidase or said functional part thereof is localized in

the ER, and said GlcNAc-transferase I or said functional part thereof is localized in the Golgi apparatus.

18. The strain according to any one of claims 12-17, wherein said strain is incapable of making a functional protein encoded by the OCH1 gene.

19. A vector for expressing a GlcNAc-transferase I or a functional part thereof in a methylotrophic yeast strain, comprising a promoter functional in said methylotrophic yeast strain, operably linked to a nucleotide sequence coding for said GlcNAc-transferase I or said functional part, wherein said GlcNAc-transferase I or said functional part comprises a genetically engineered localization signal for a site in the secretory pathway of said strain, wherein the glycoproteins at said site have the oligosaccharide structure of $\text{Man}_5\text{GlcNAc}_2$.

20. The vector of claim 19, wherein the said methylotrophic strain is *Pichia*.

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

22. The vector of claim 19, wherein said localization signal is a Golgi localization signal.

23. The vector according to any one of claims 19-22, further comprising a nucleotide sequence coding for an α -1,2-mannosidase or a functional part thereof.

24. A vector for expressing a mannosidase II or a functional part thereof in a methylotrophic yeast strain, comprising a promoter functional in said methylotrophic yeast strain, operably linked to a nucleotide sequence coding for said

mannosidase II or said functional part, wherein said mannosidase II or said functional part comprises a Golgi localization signal.

25. The vector of claim 24, wherein the said methylotrophic strain is *Pichia*.

26. The vector of claim 24, wherein said mannosidase II is of a mammalian origin.

27. The vector according to any one of claims 24-26, further comprising a nucleotide sequence coding for an α -1,2-mannosidase or a functional part thereof.

28. Use of the strain according to any one of claims 1-18 for producing a glycoprotein with a mammalian-like glycosylation pattern.

29. Use of the vector according to any one of claims 19-27 for creating a methylotrophic yeast strain that produces a glycoprotein with a mammalian-like glycosylation pattern.

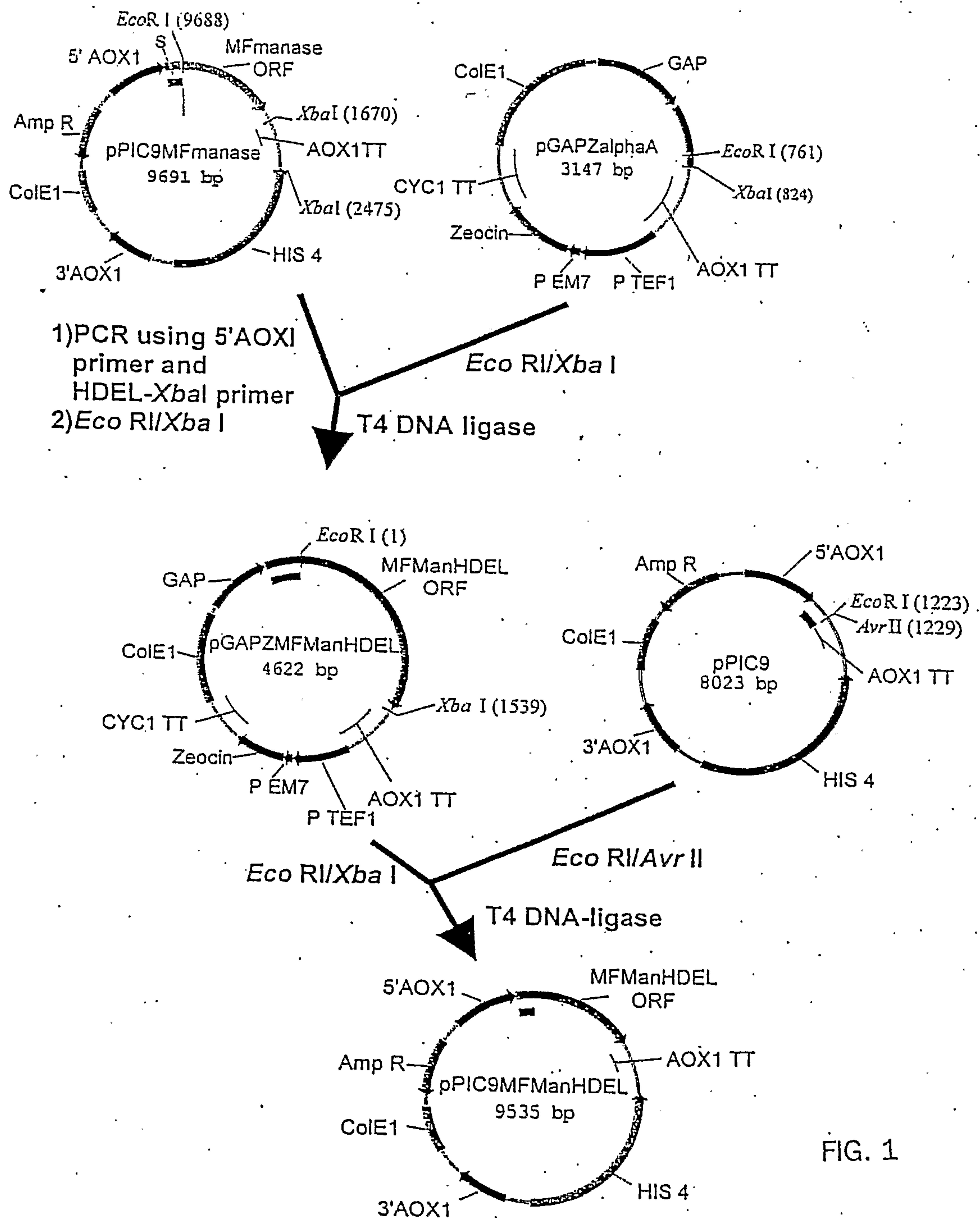
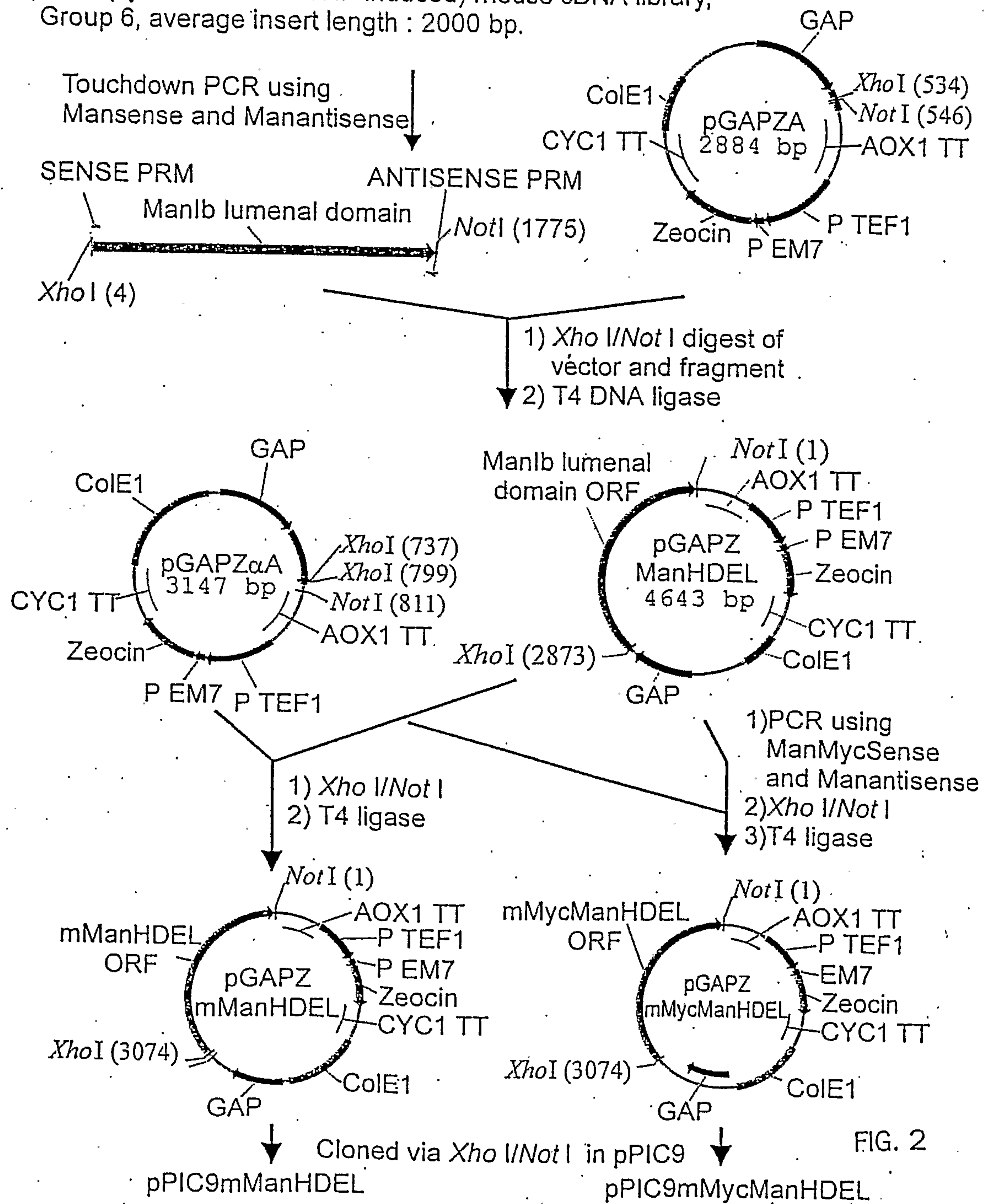


FIG. 1

L929 (cycloheximide+ TNF induced) mouse cDNA library,
Group 6, average insert length : 2000 bp.



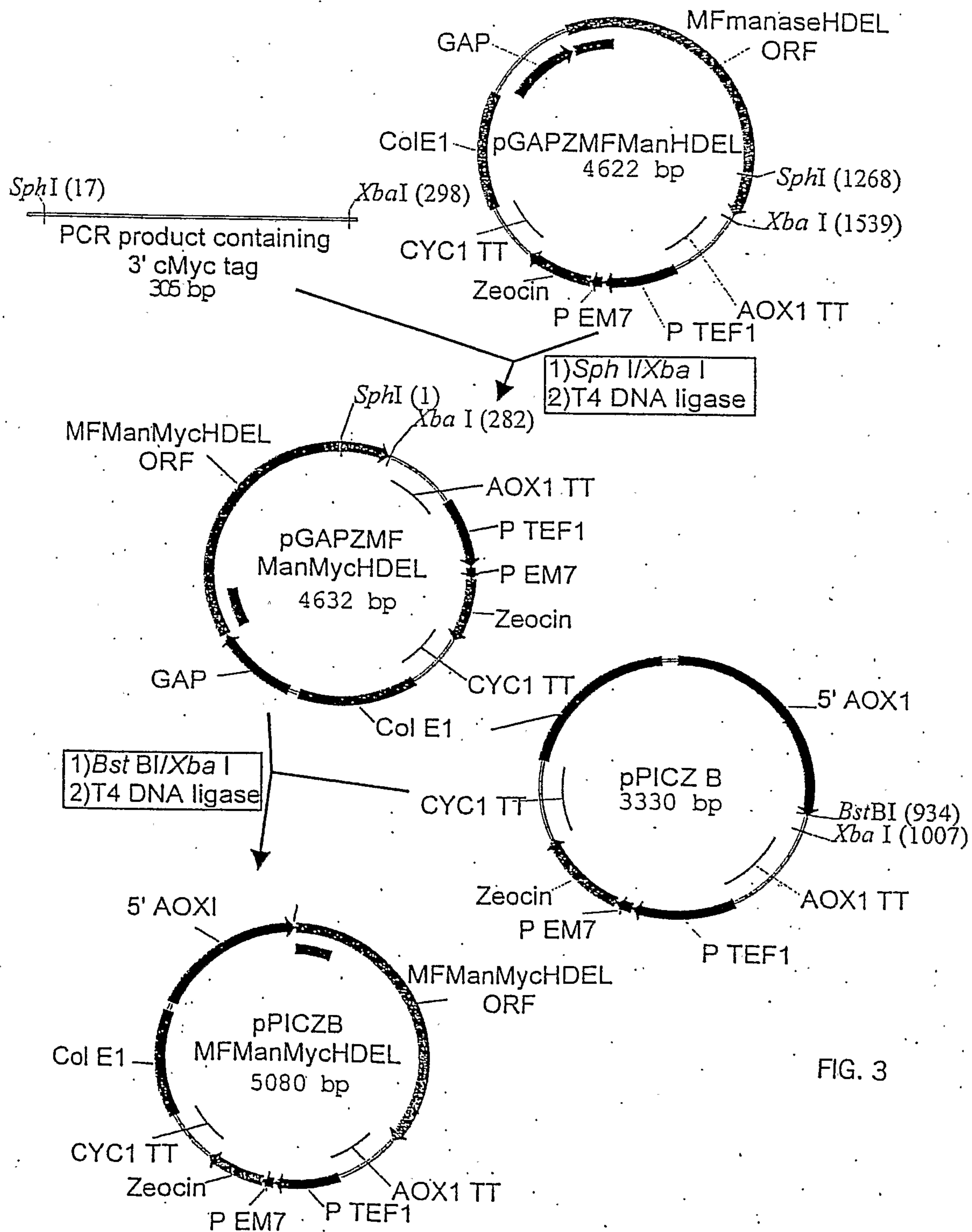
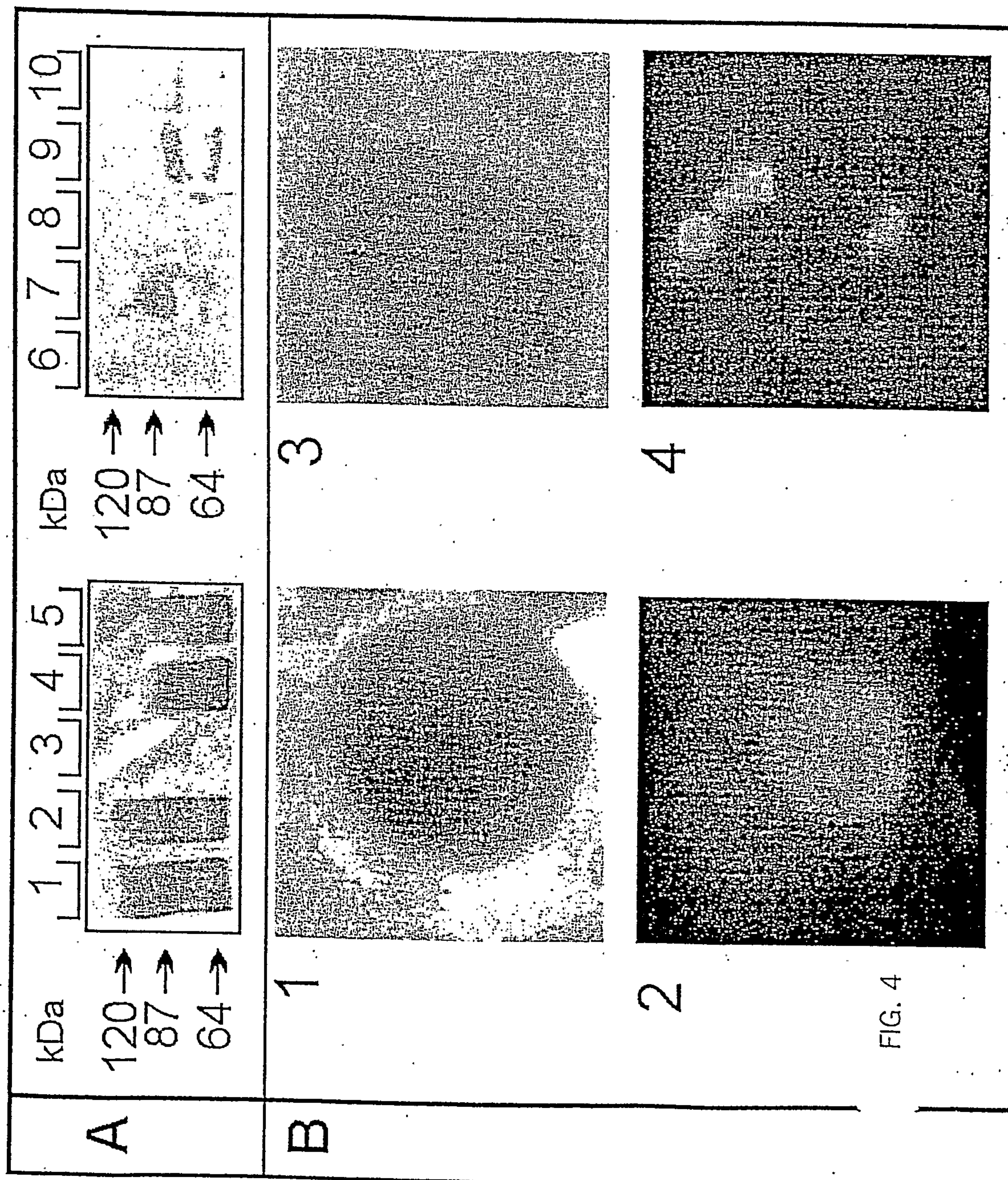


FIG. 3



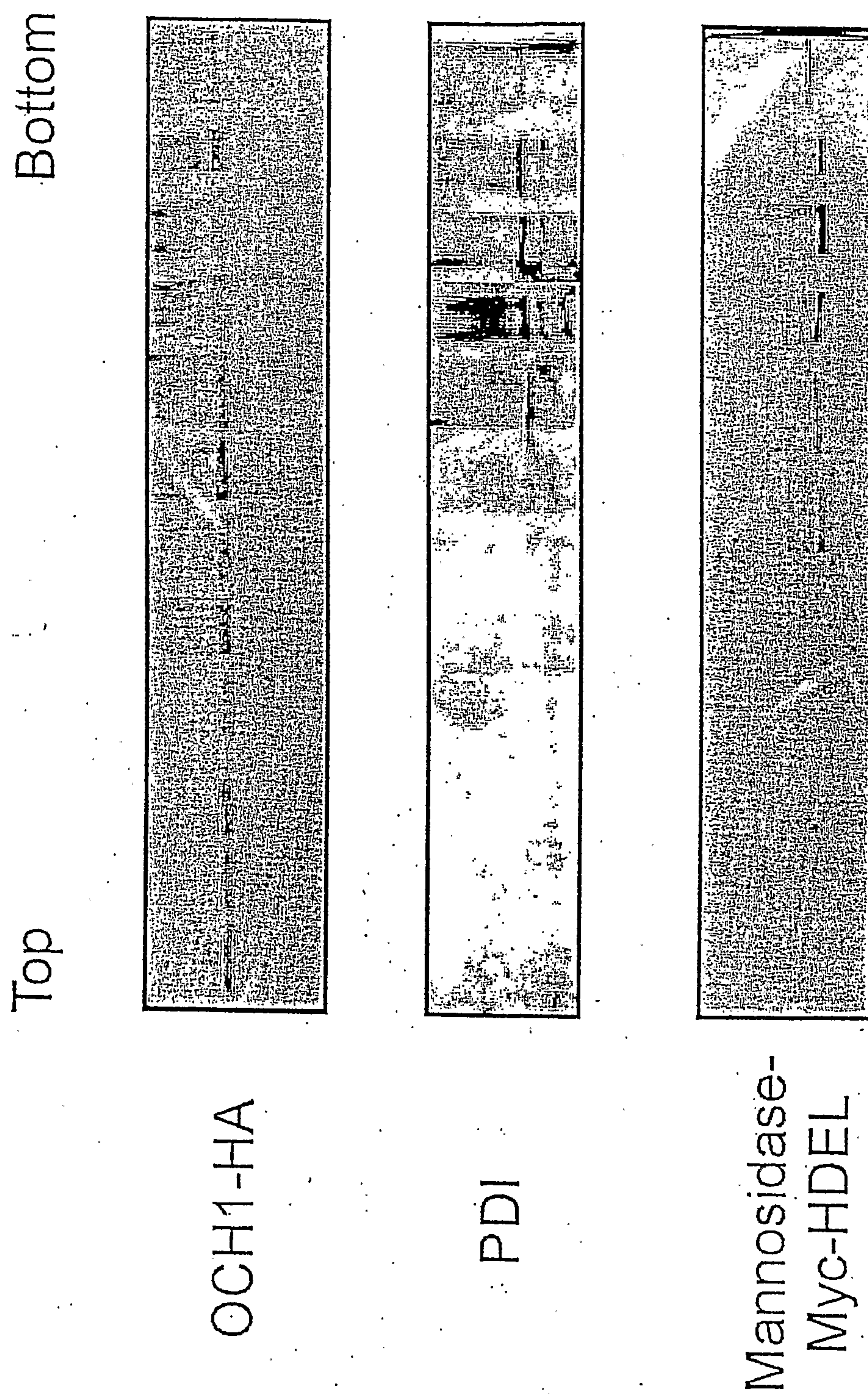
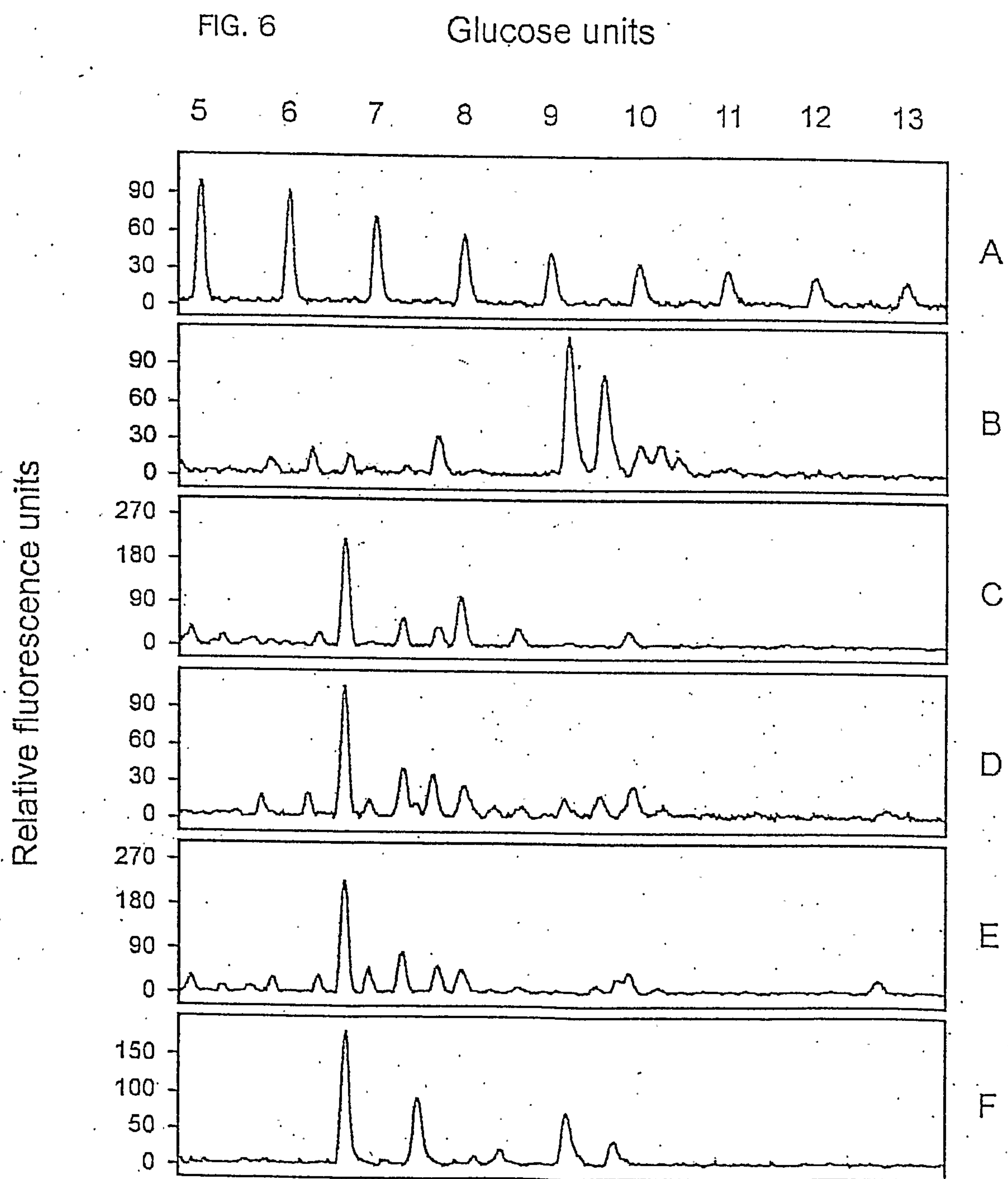
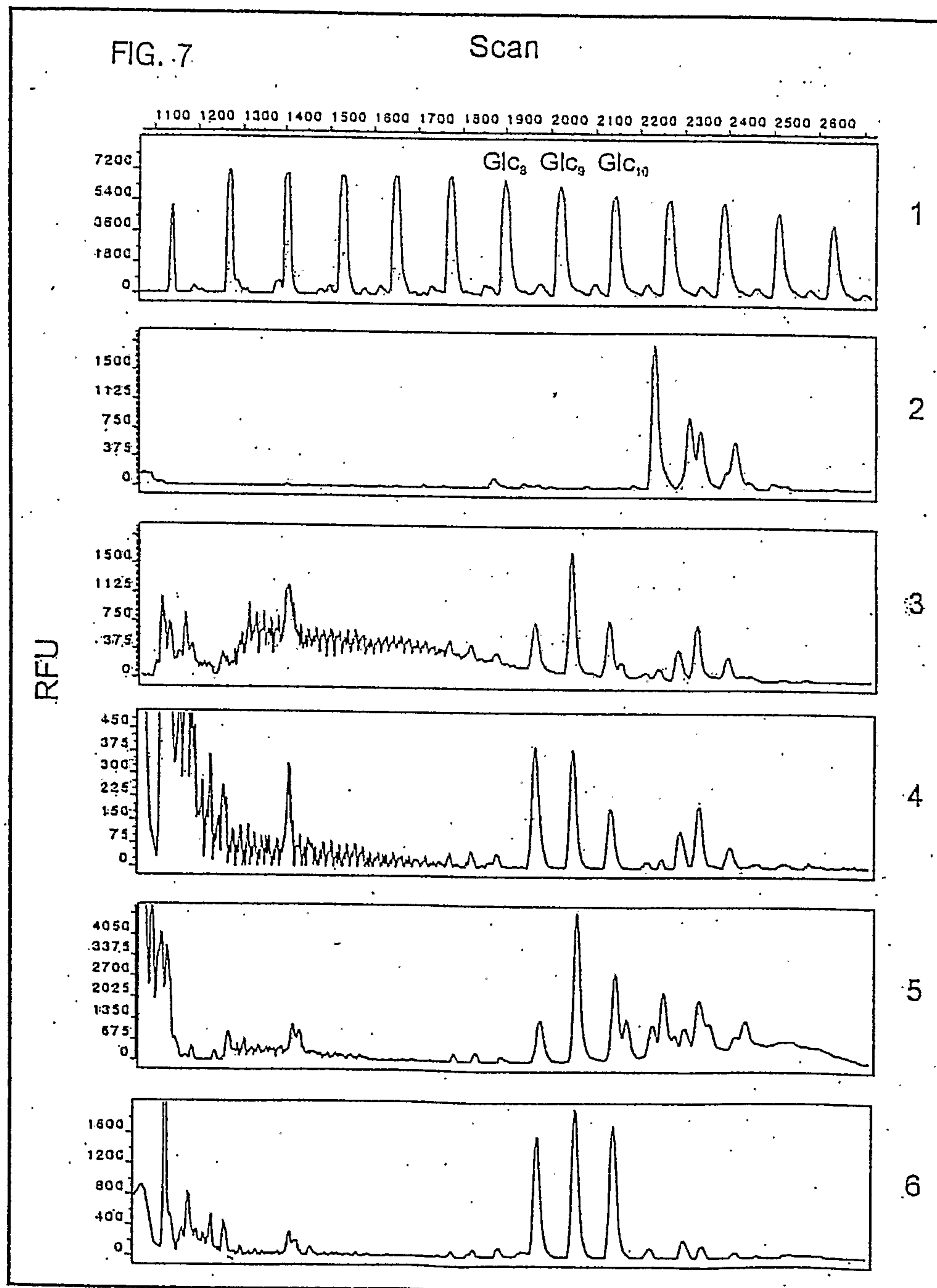


FIG. 5





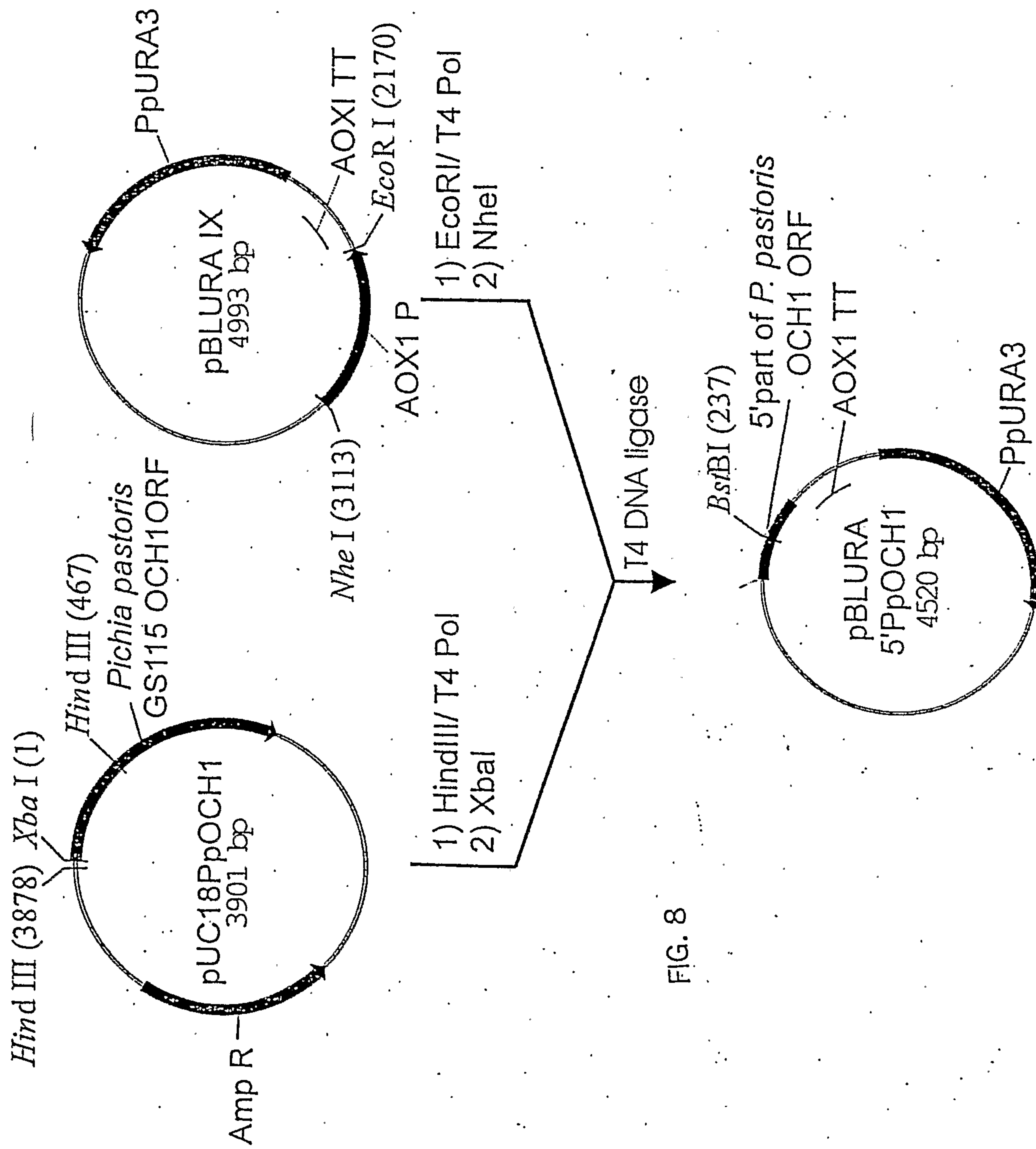


FIG. 8

Pichia OCH1 Genomic fragment

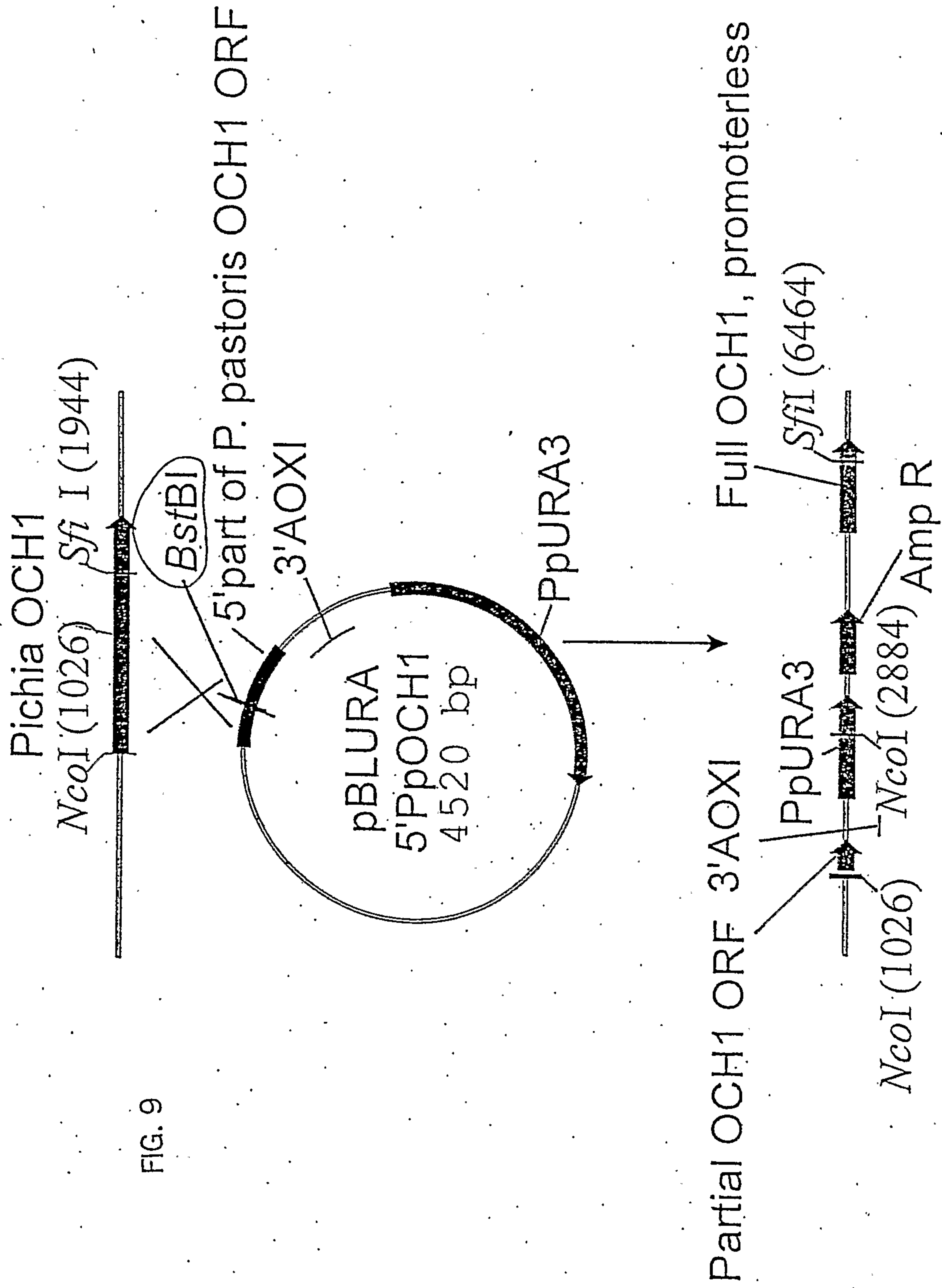


FIG. 10

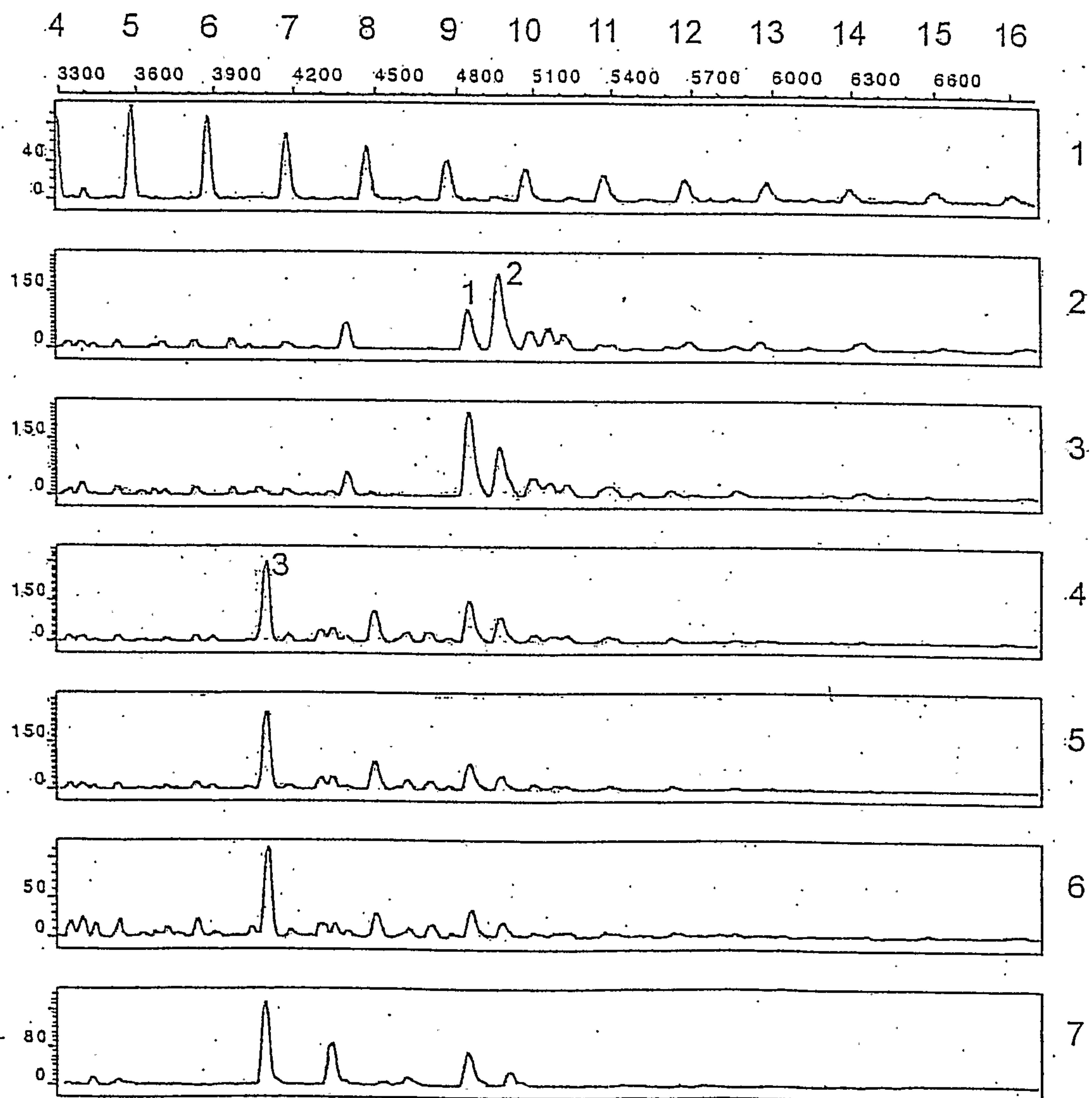
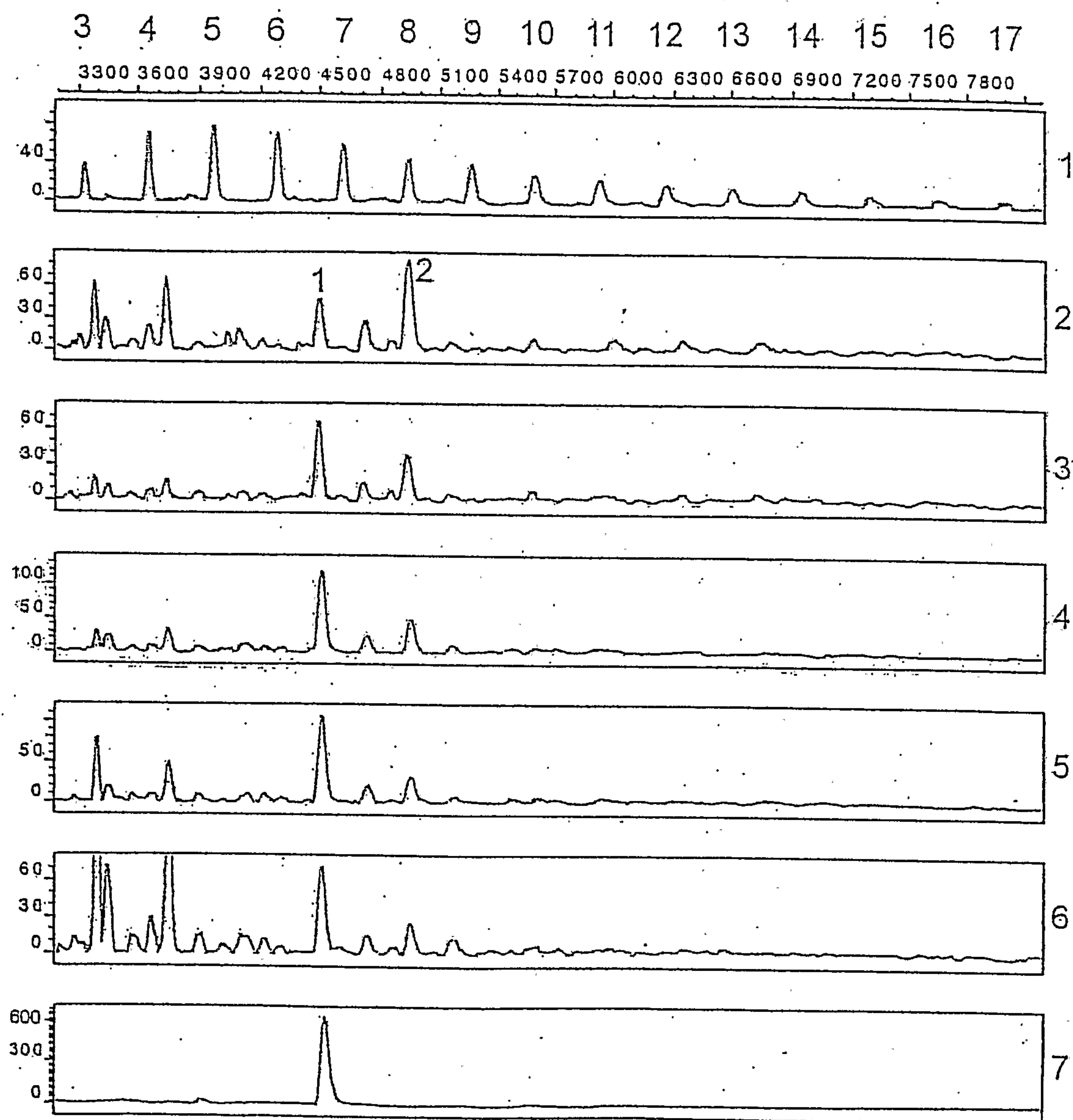


FIG. 11



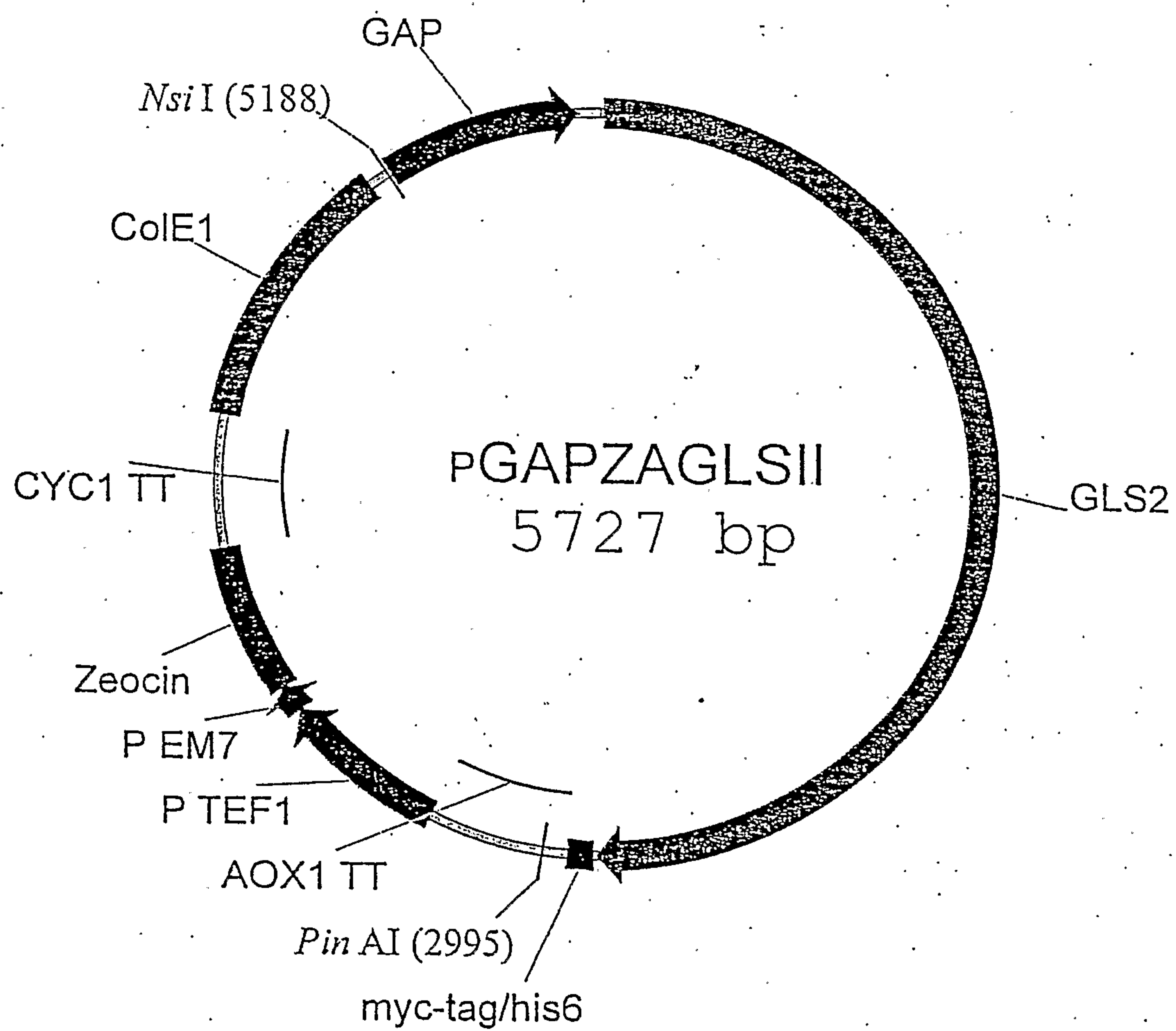


FIG. 12

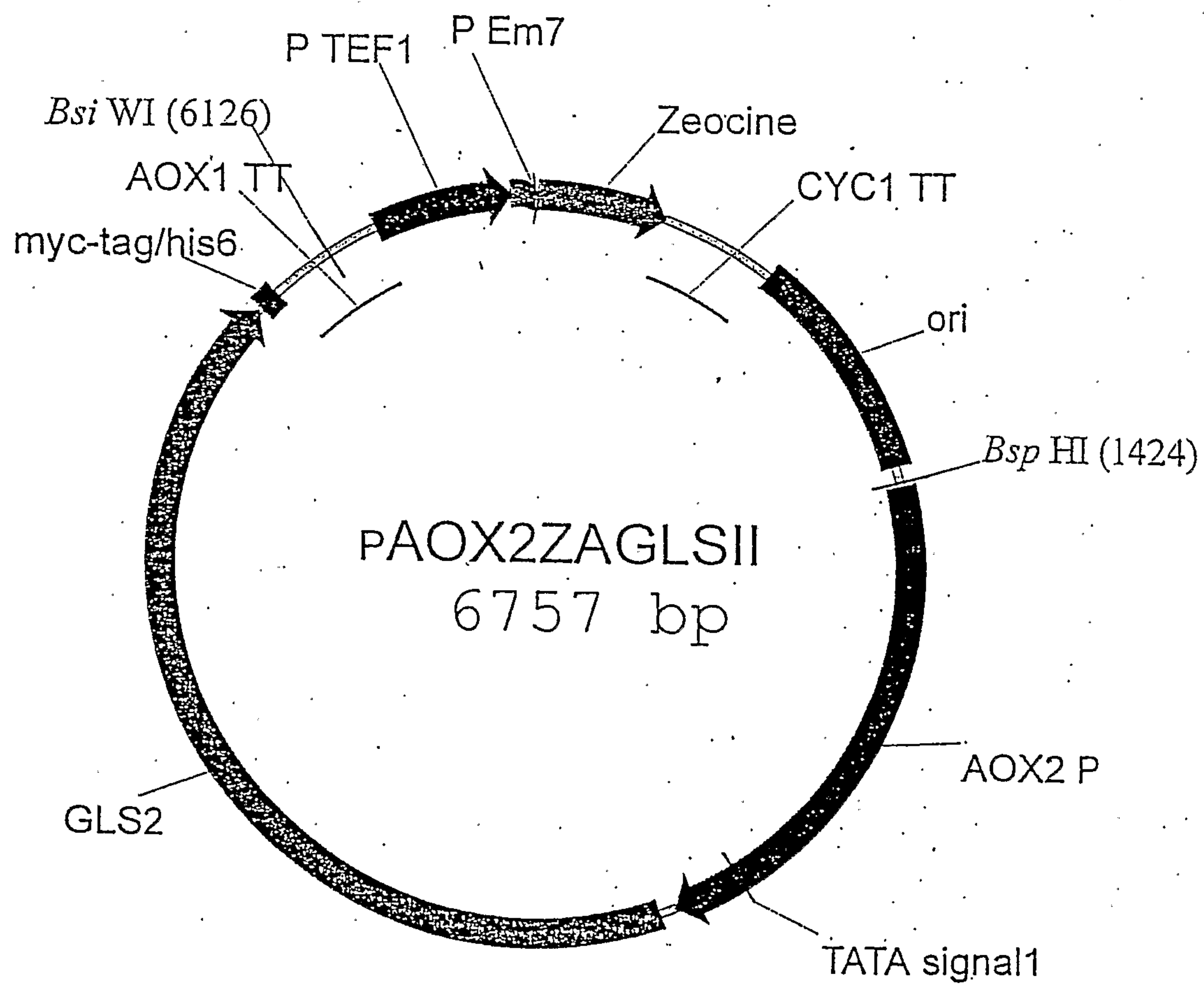


FIG.13

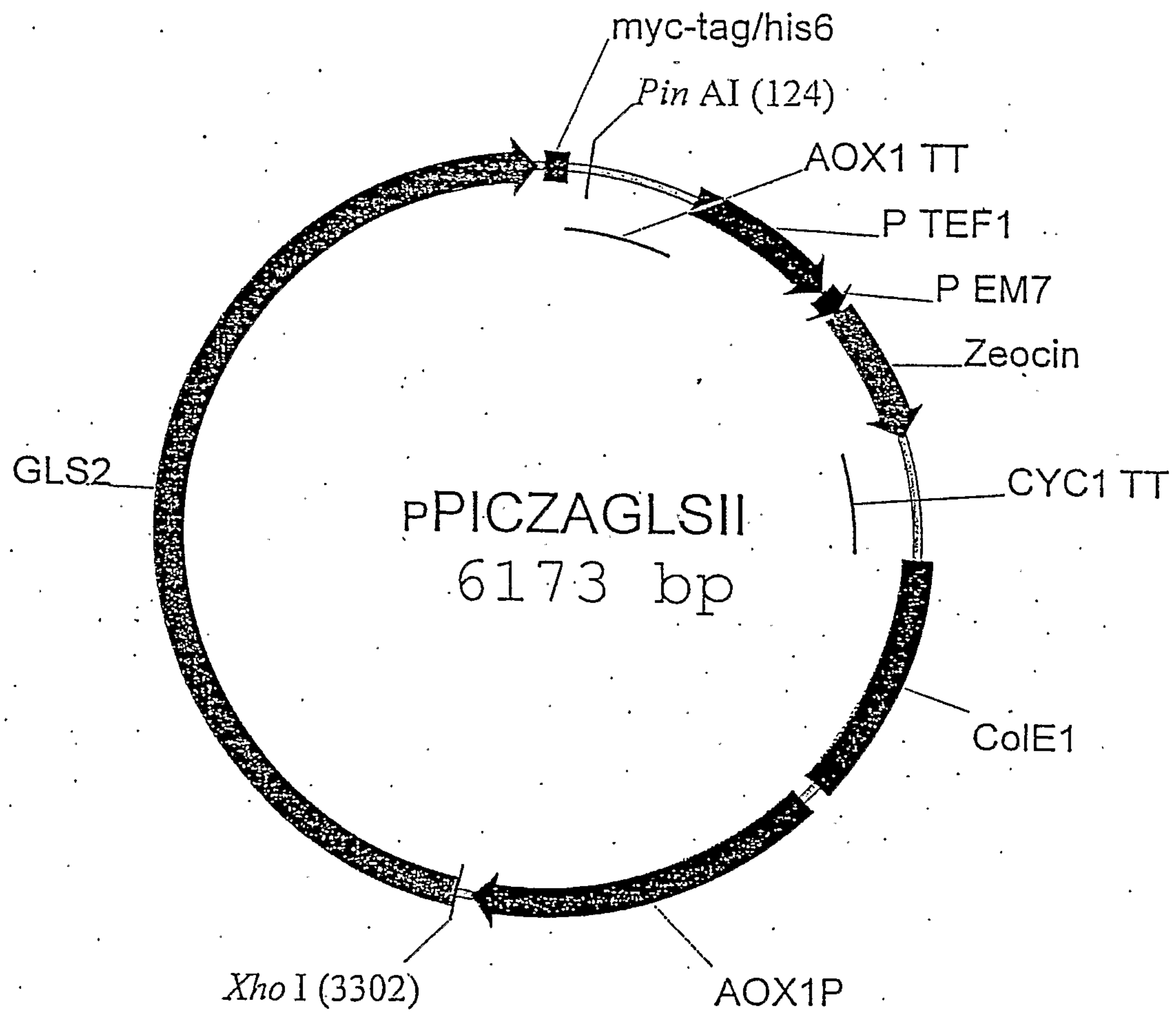


FIG. 14

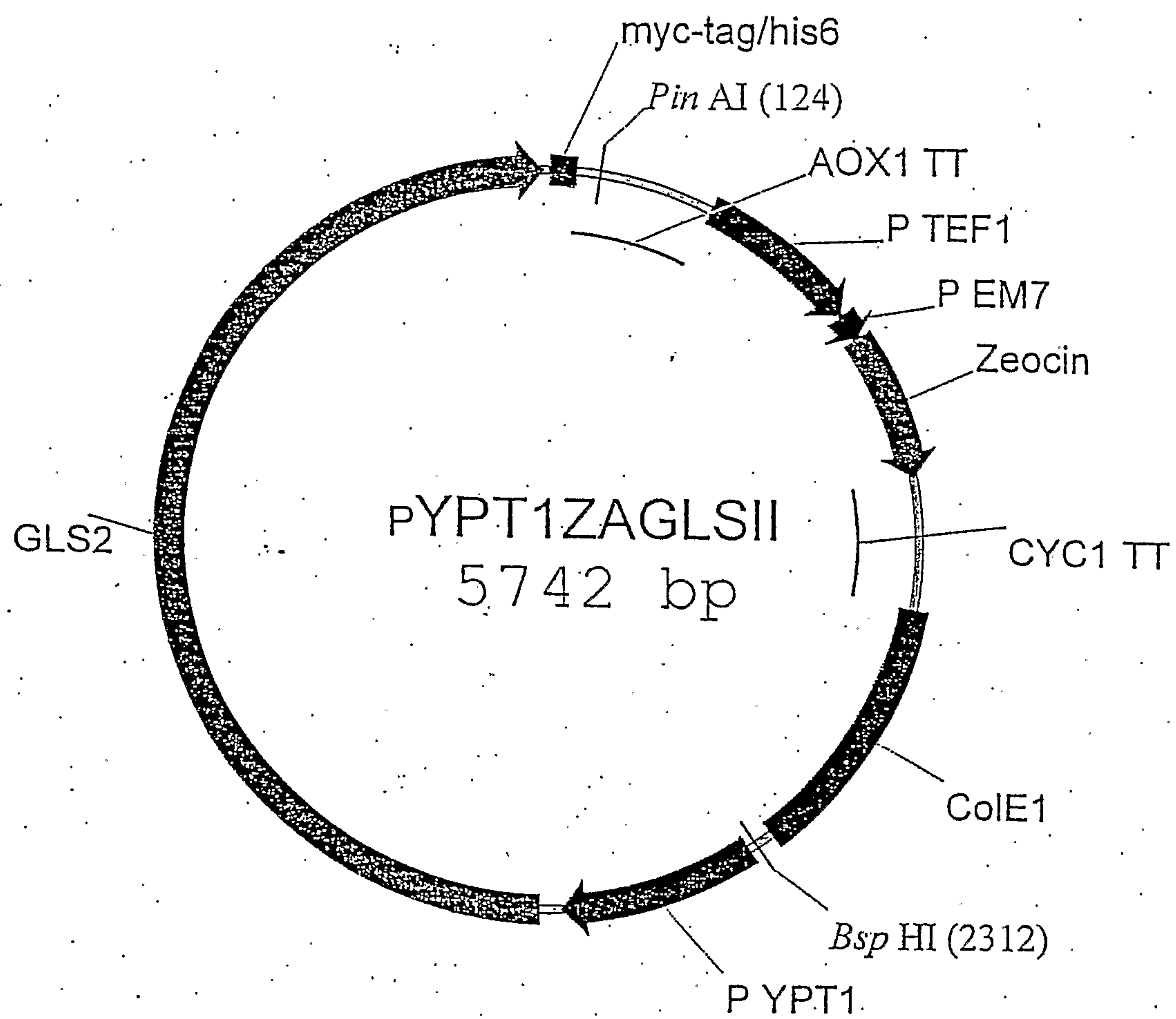


FIG. 15

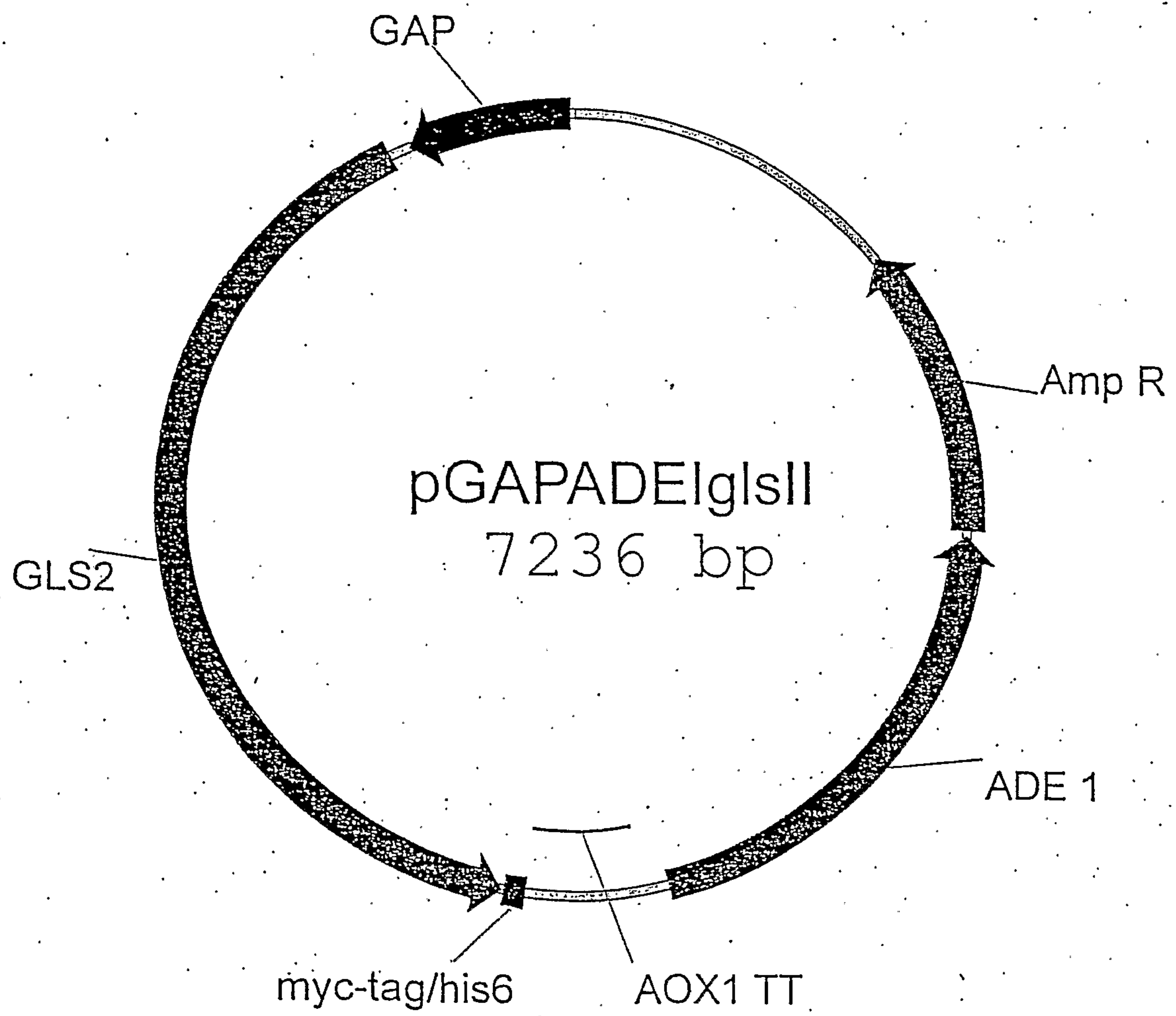


FIG. 16

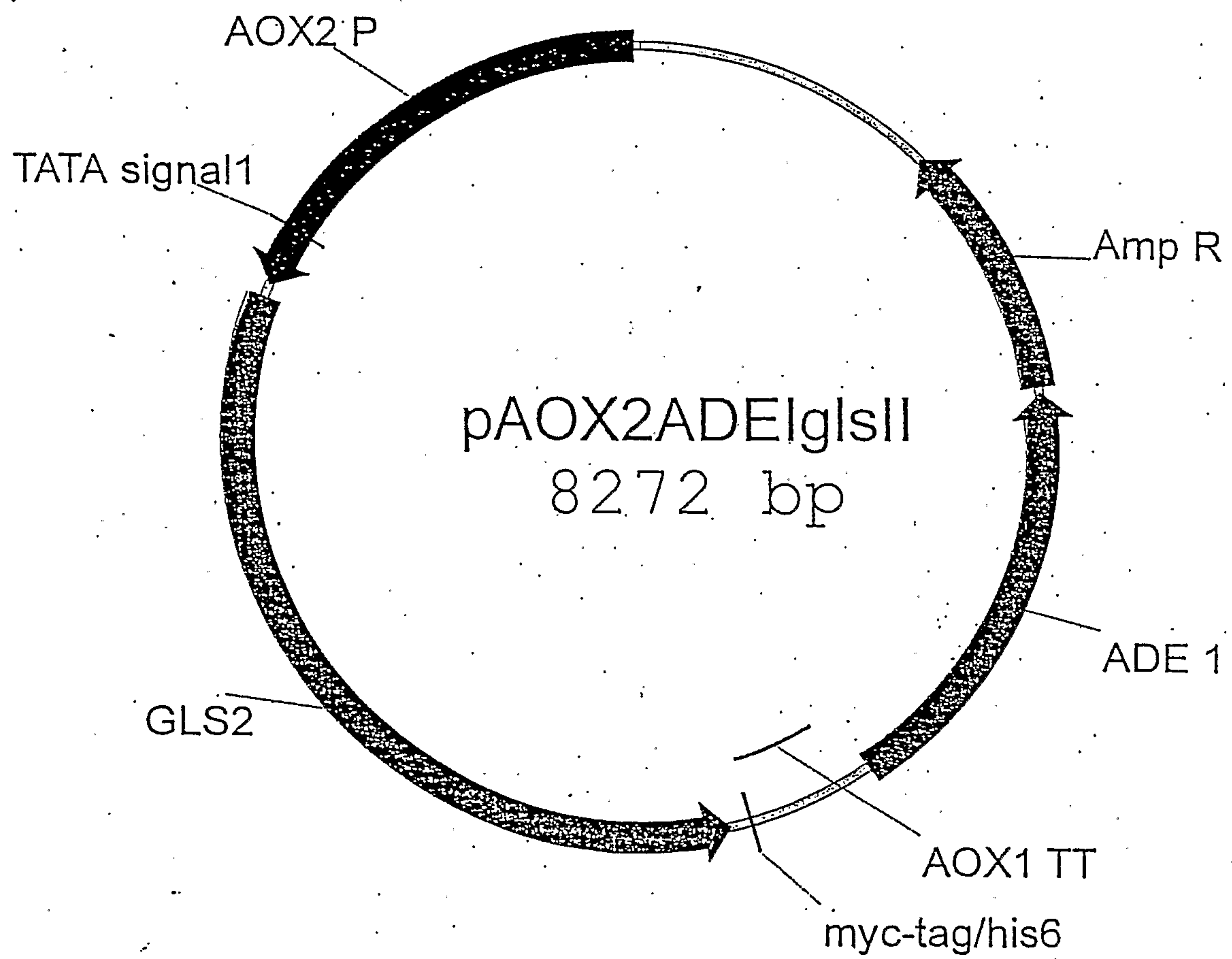


FIG. 17

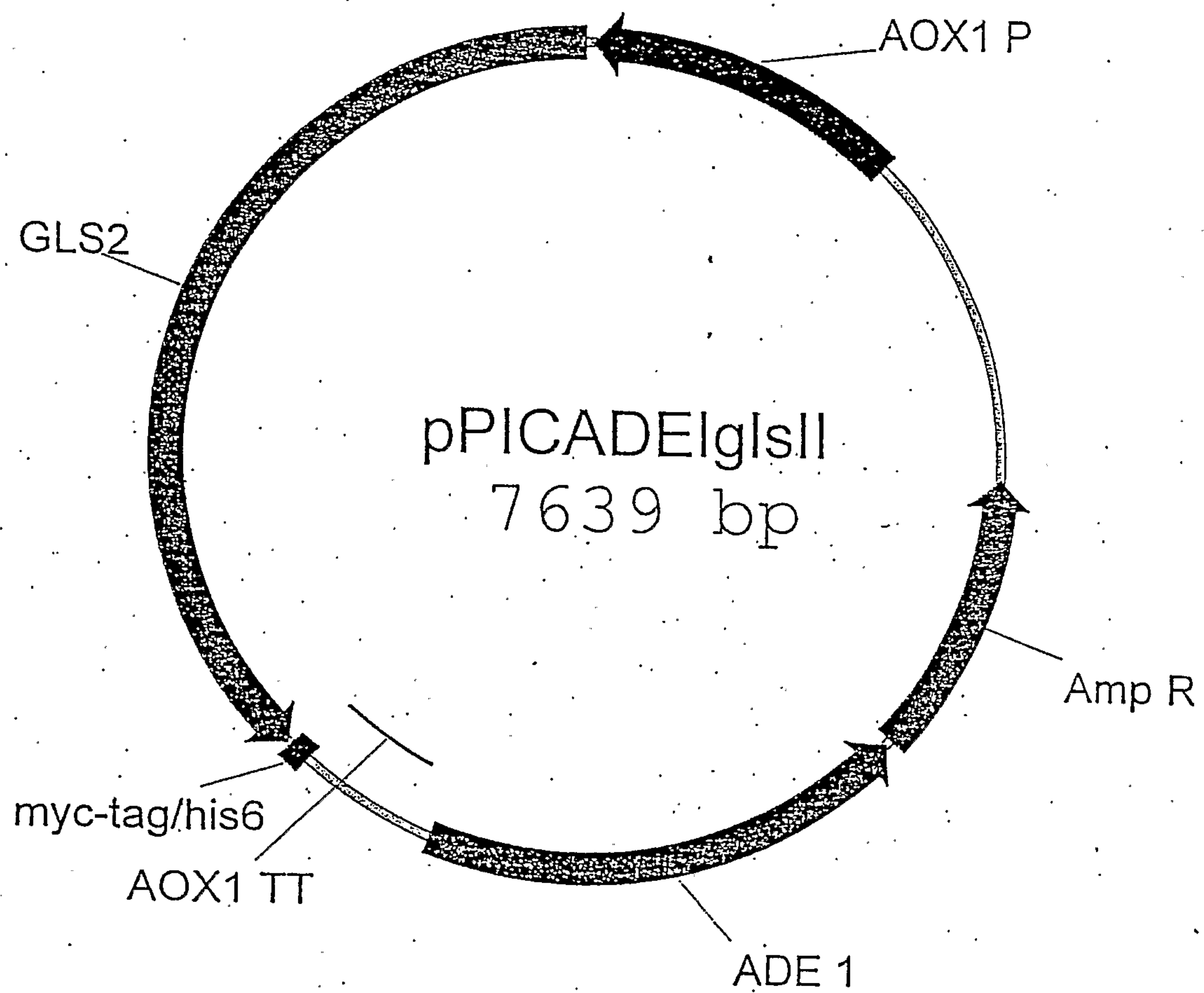


FIG. 18

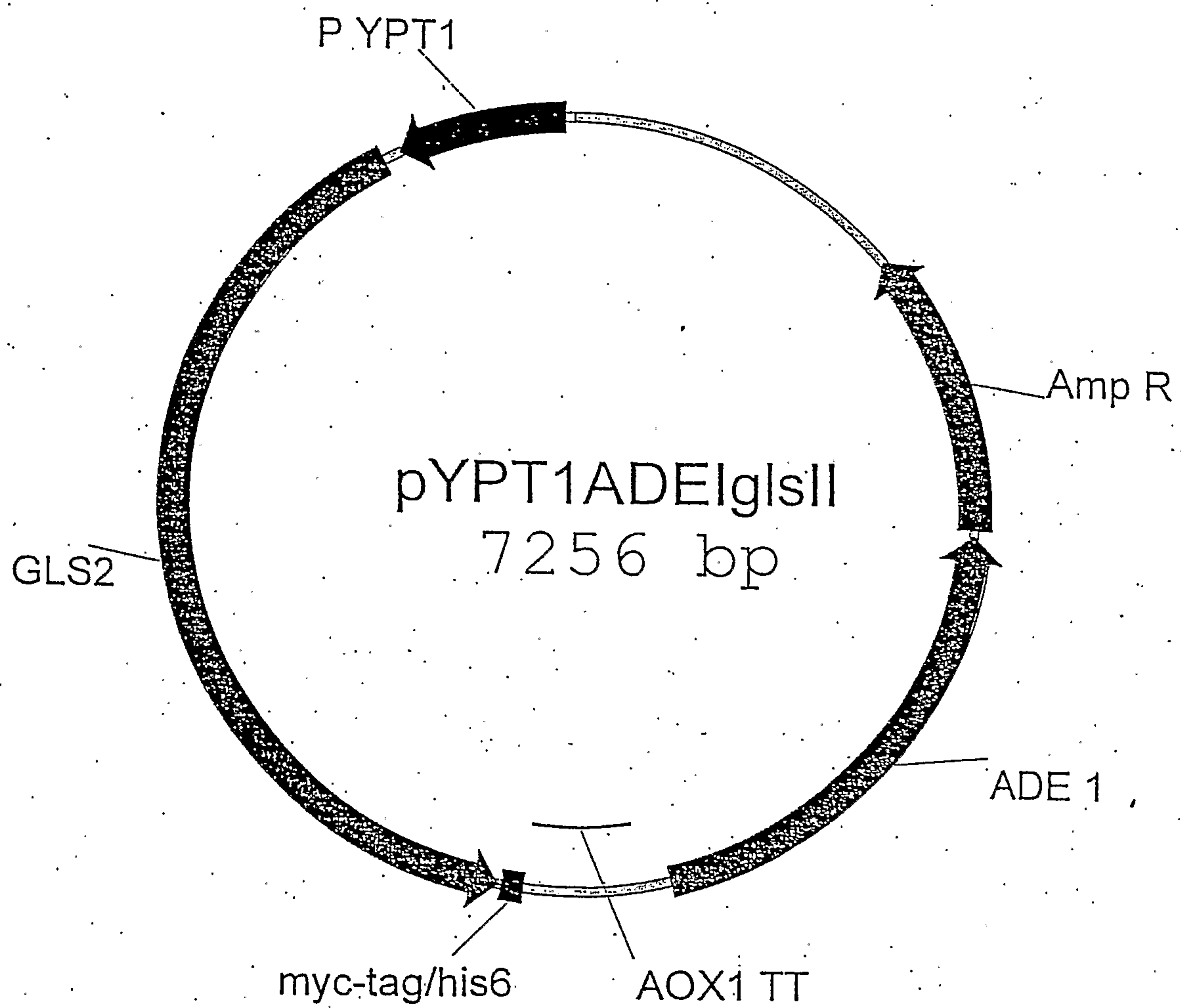


FIG. 19

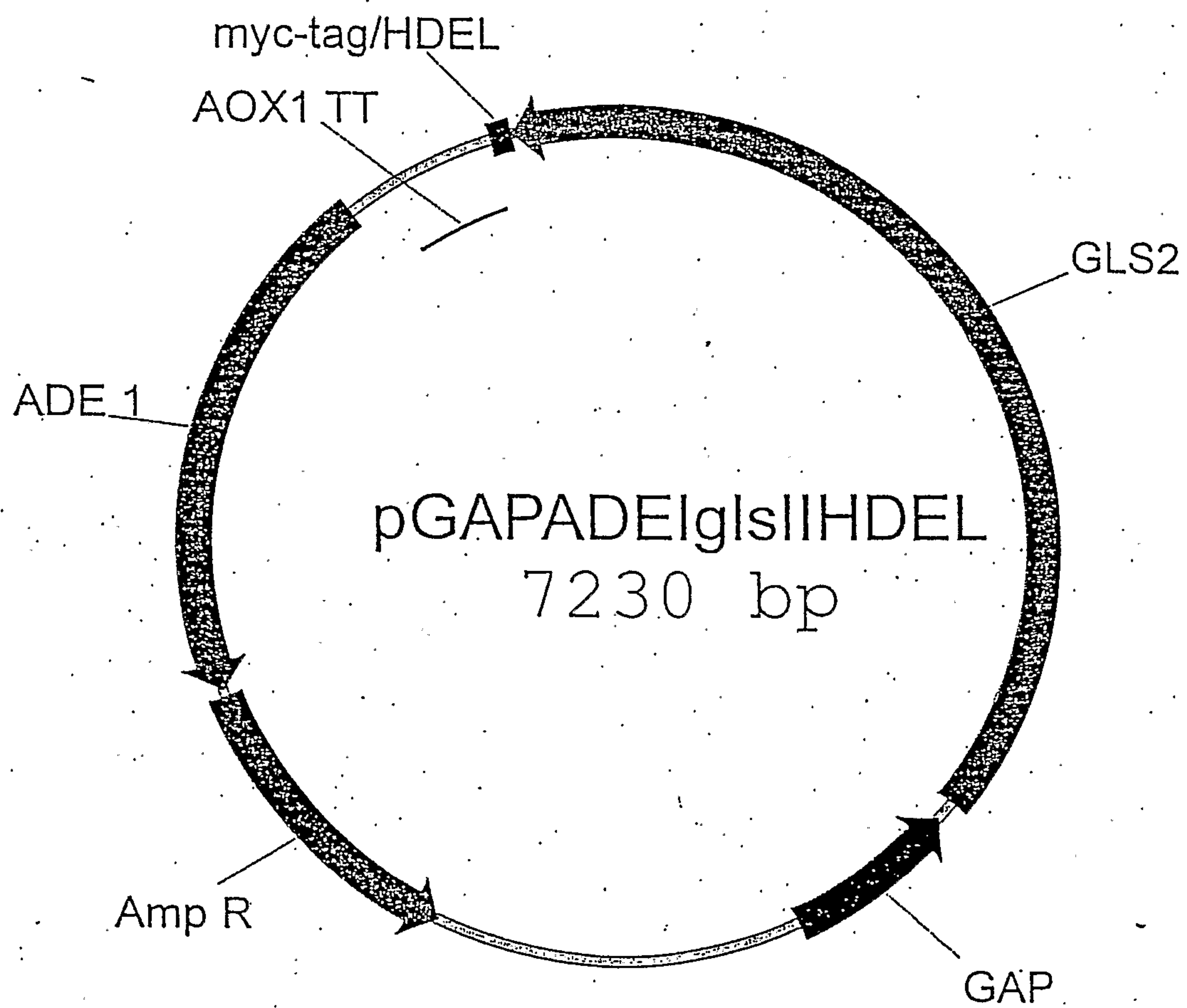


FIG. 20

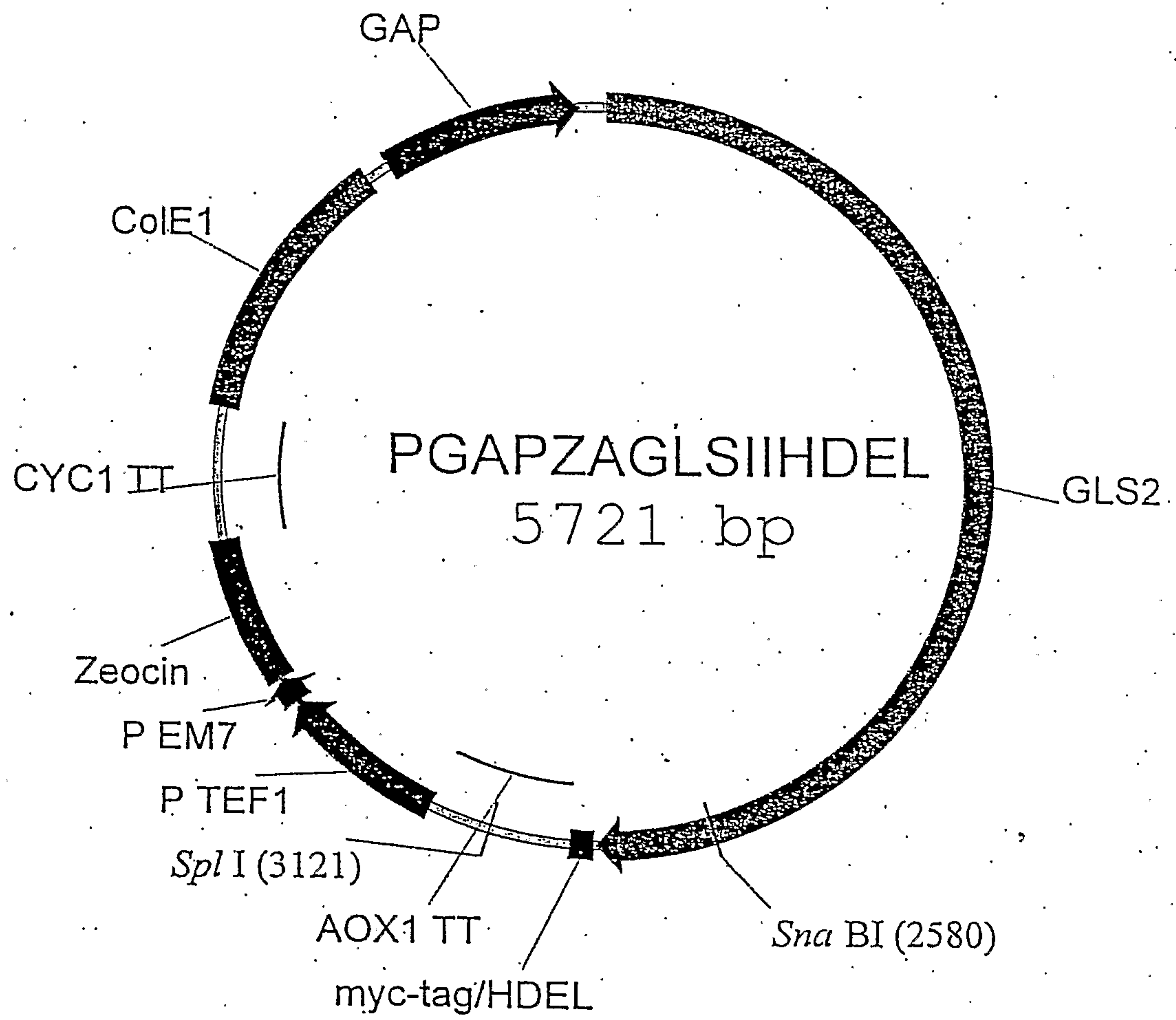


FIG. 21

Glucosidase II assay on commercially available alpha-glucosidase

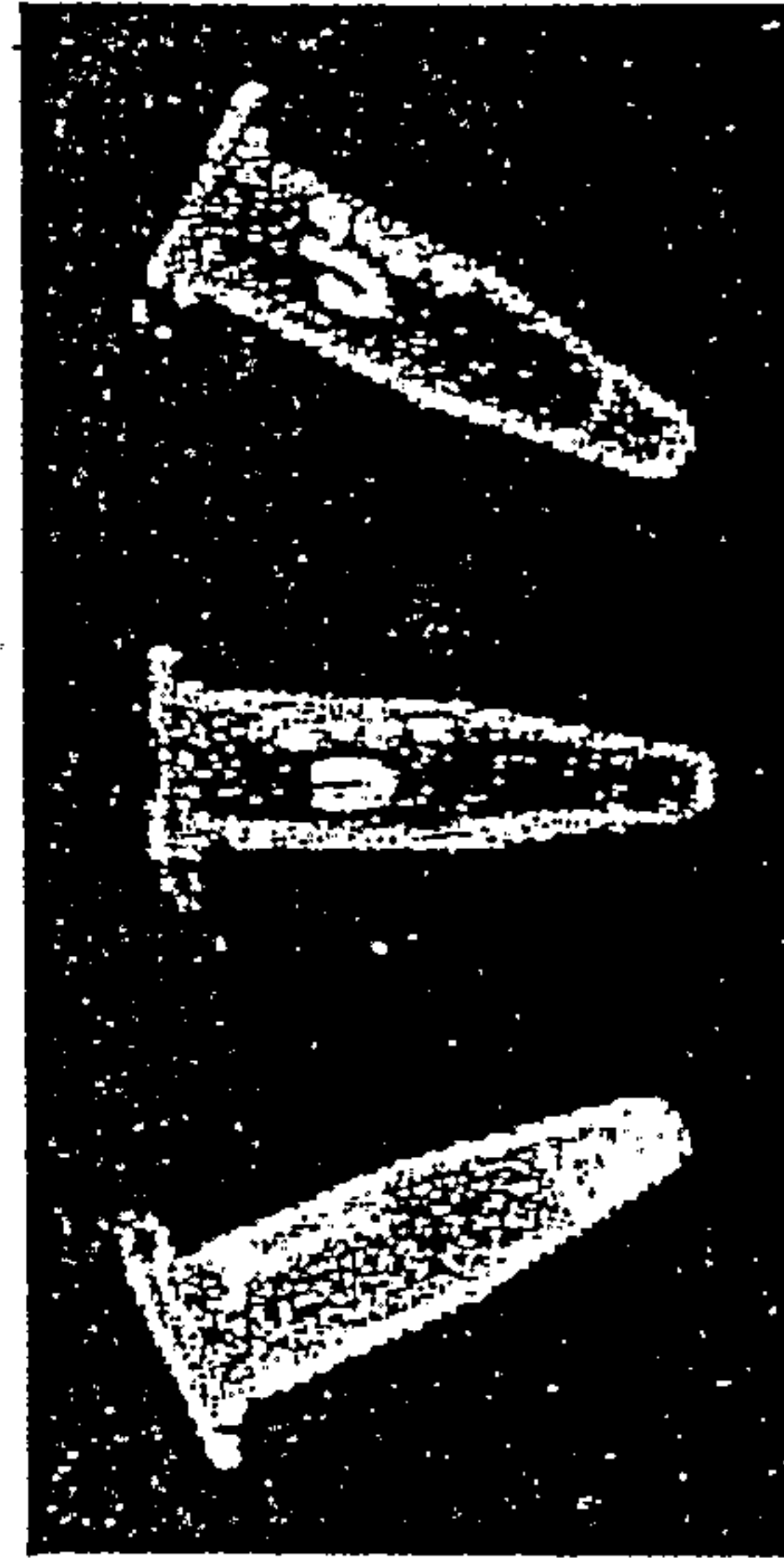


FIG. 22

Glucosidase II assay on heterologously expressed Pichia protein

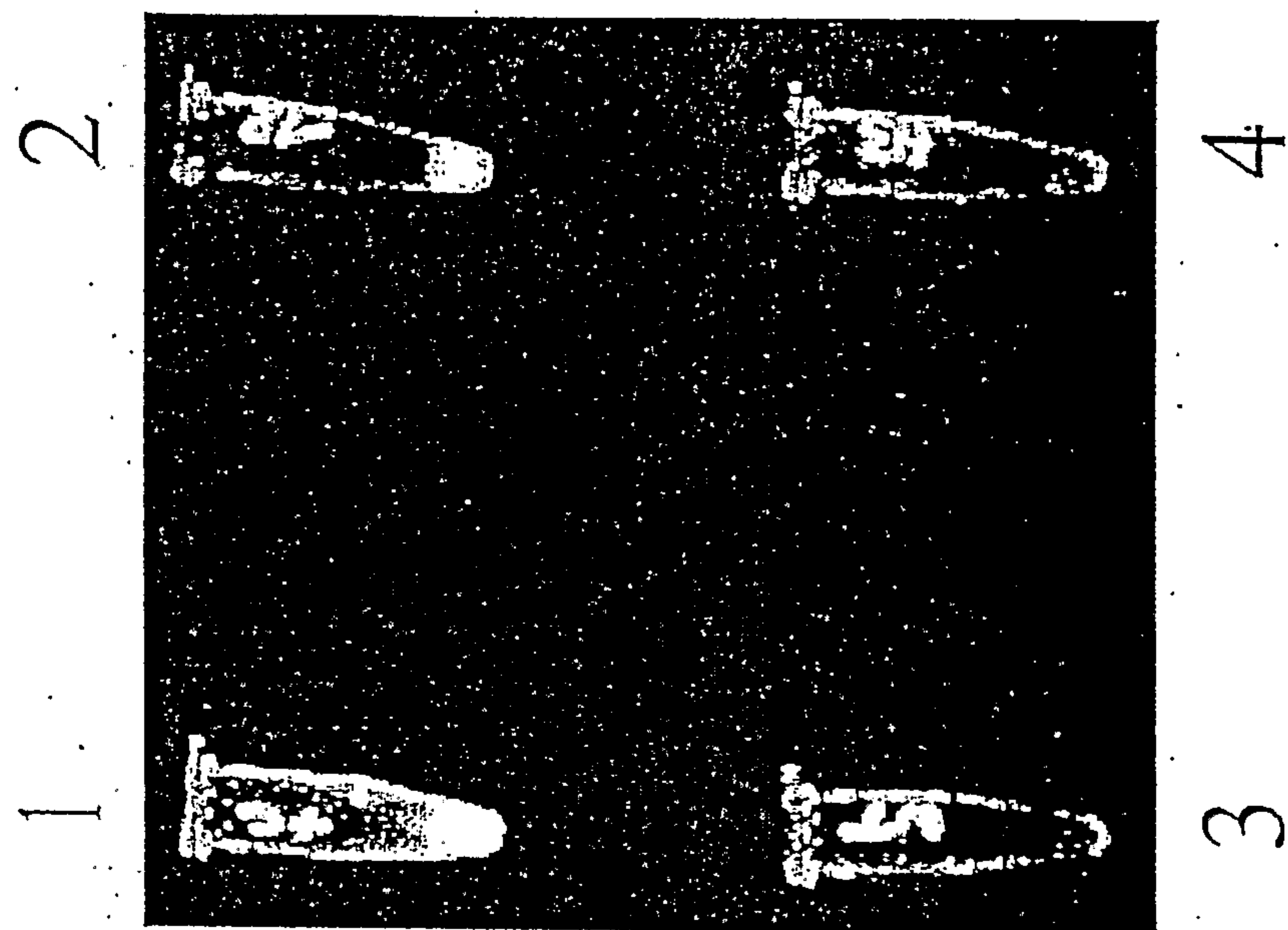


FIG. 23

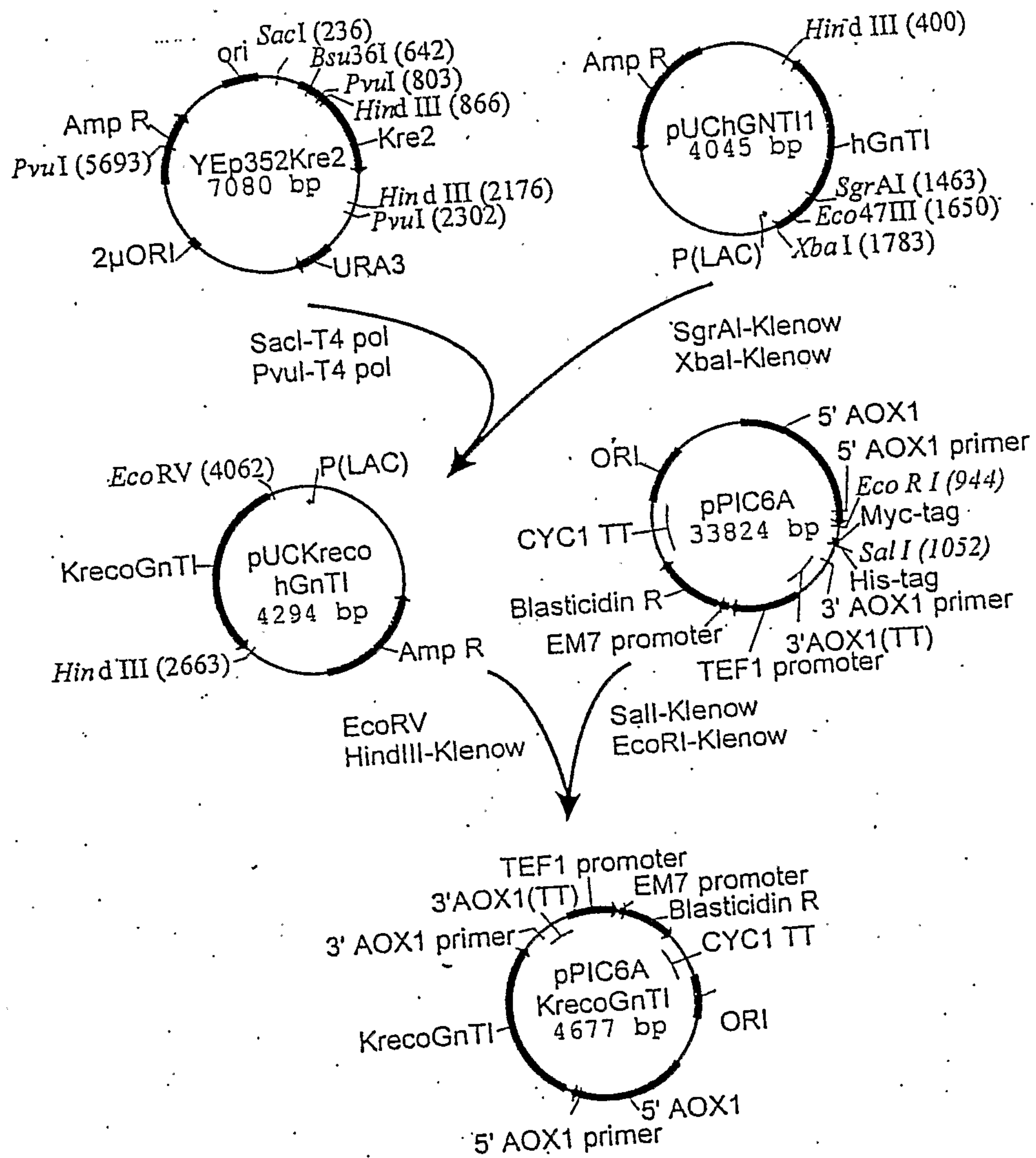


FIG. 24

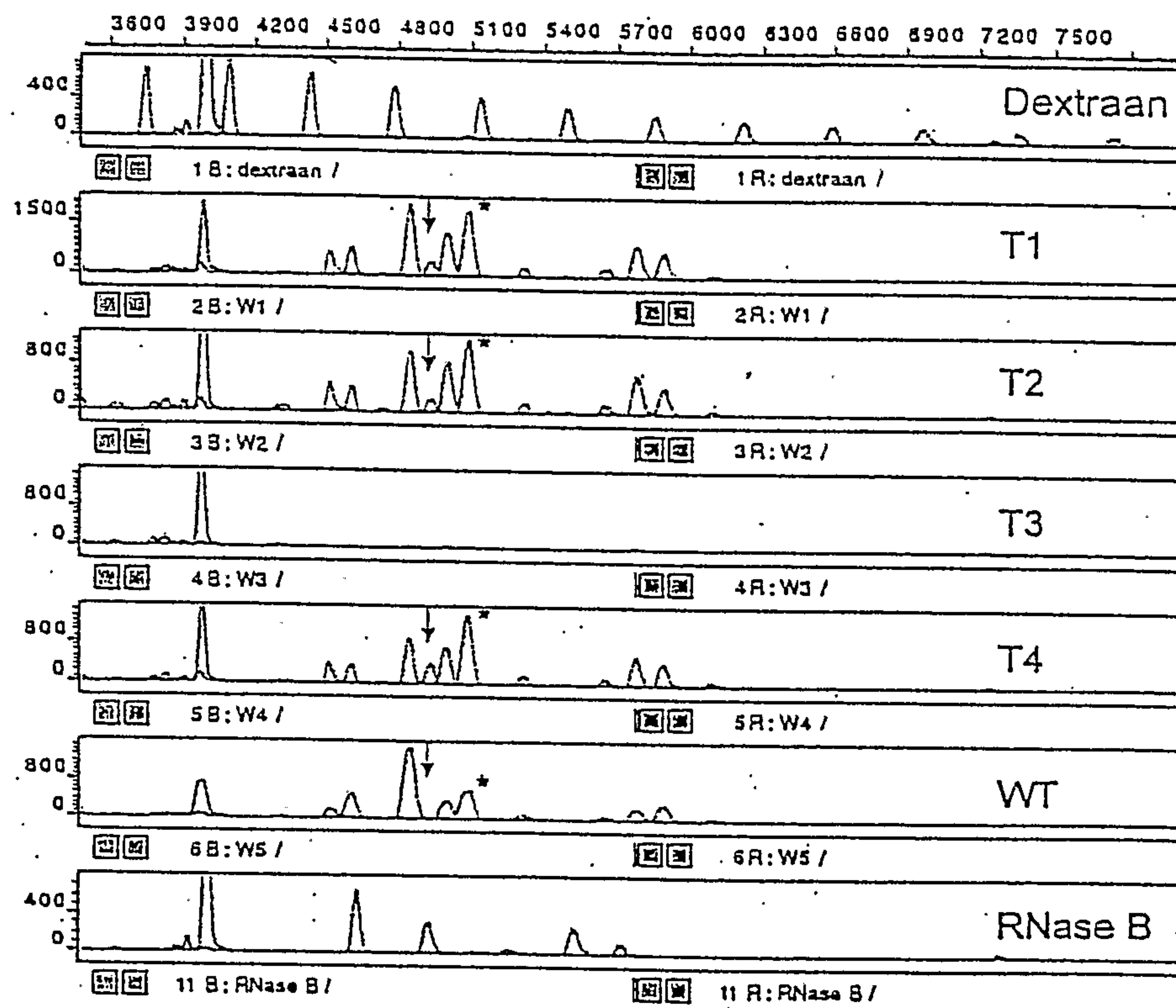


FIG. 25

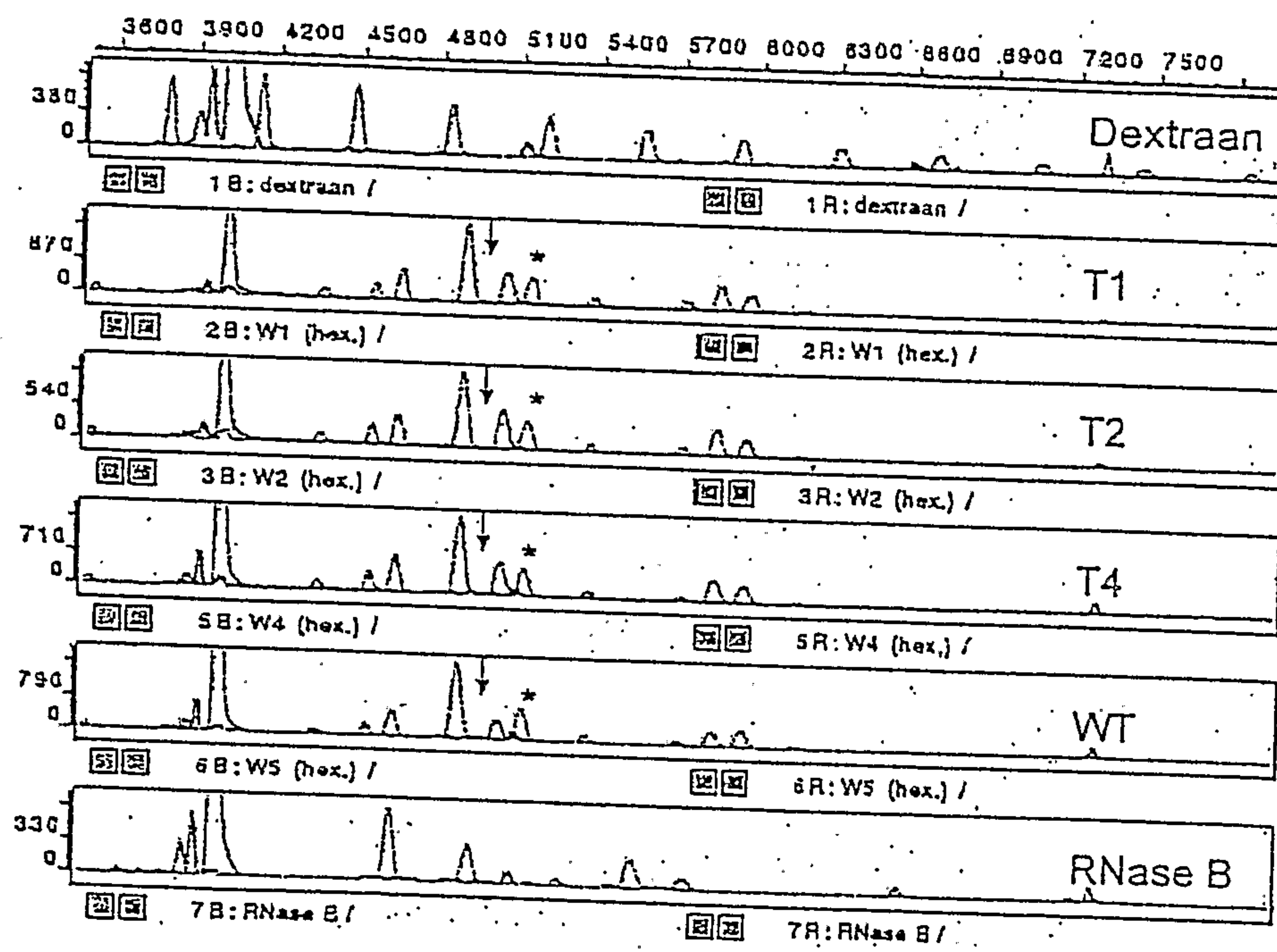


FIG. 26

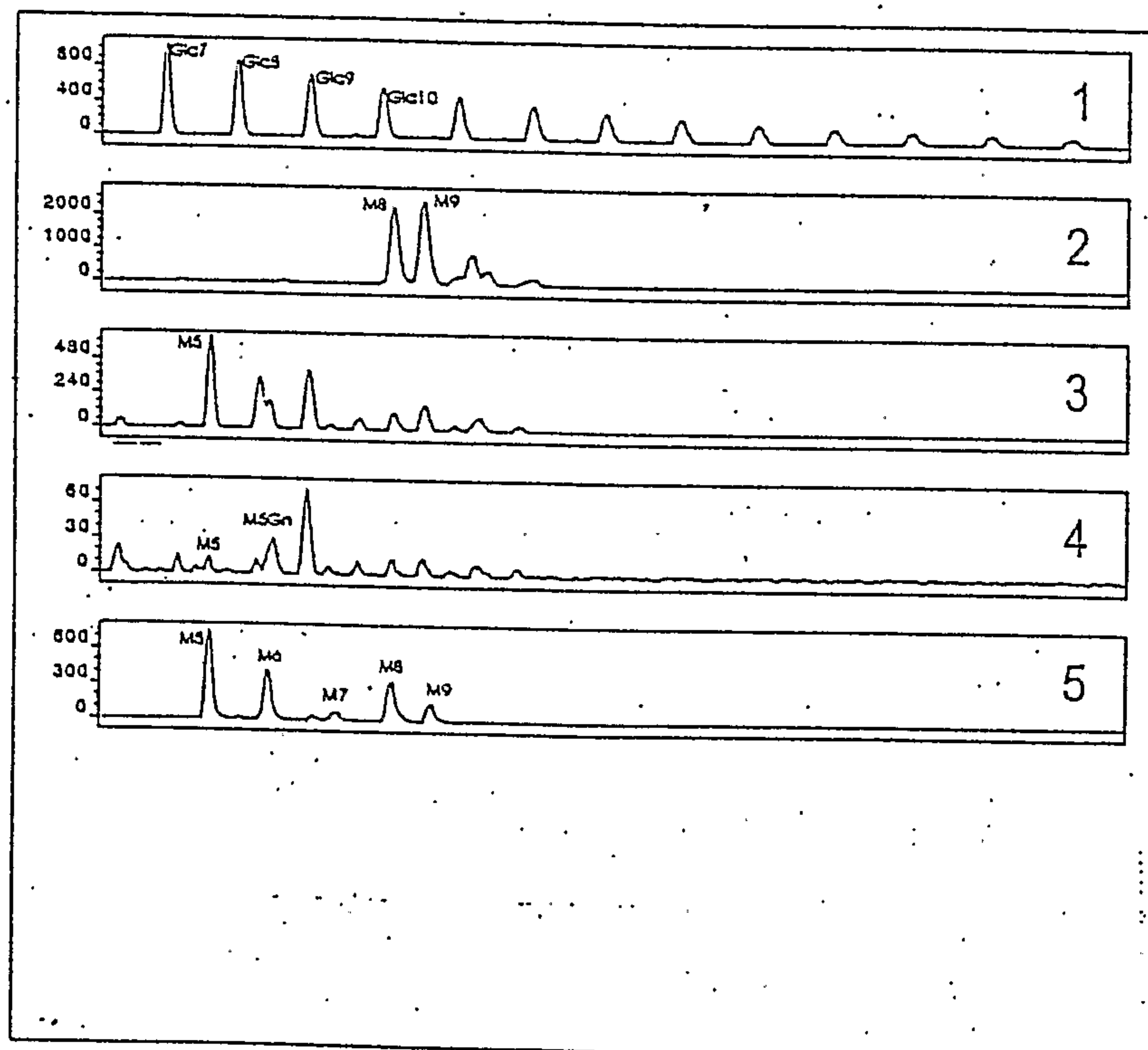


FIG. 27

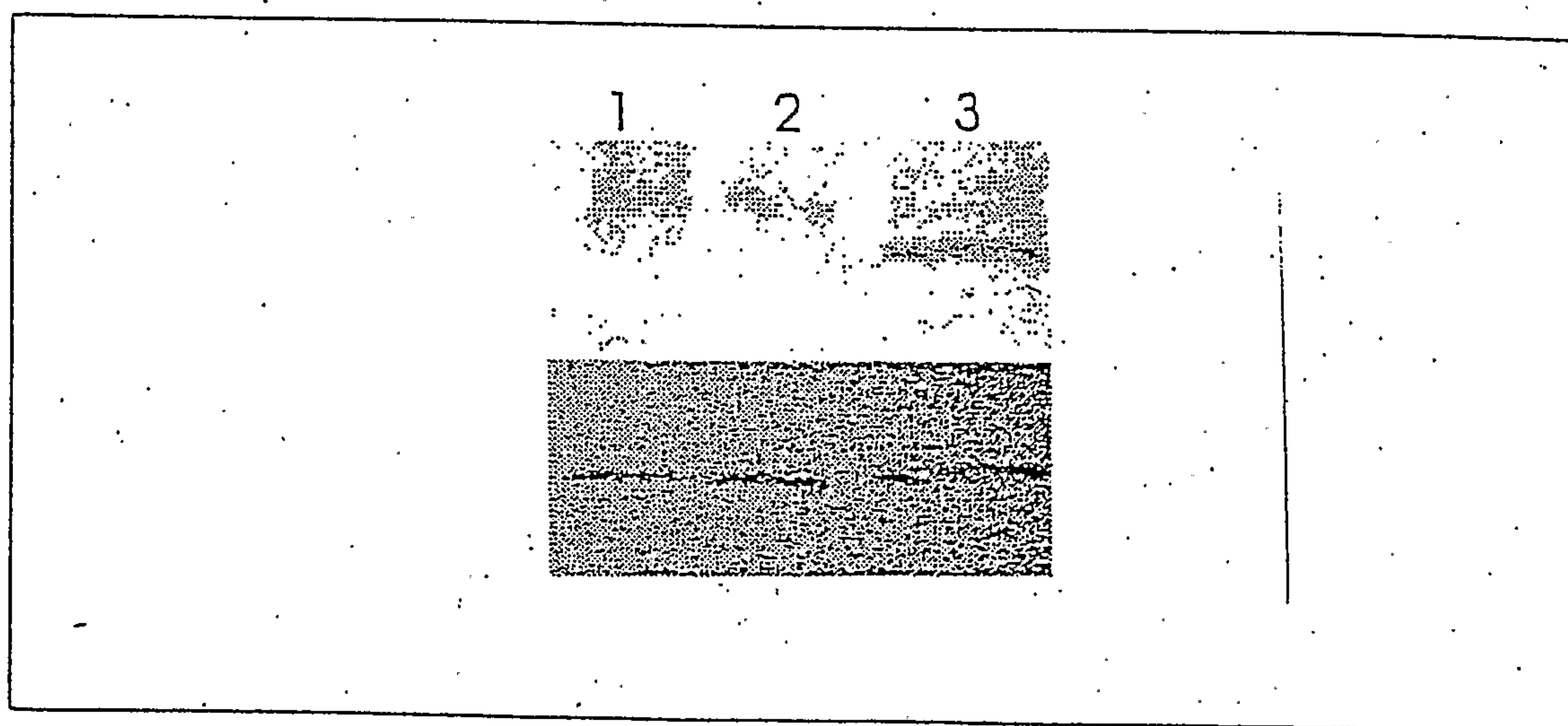


FIG. 28

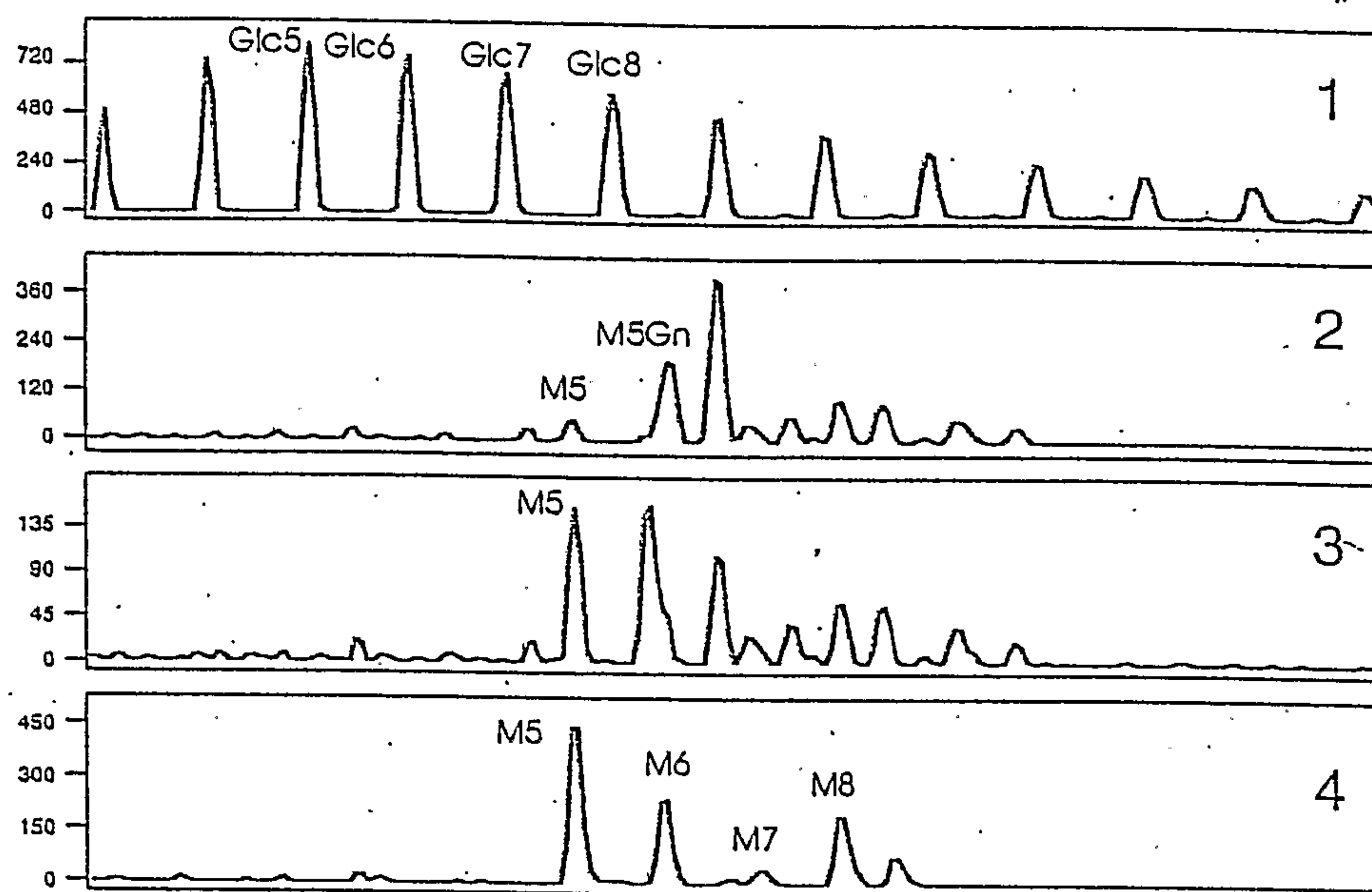


FIG. 29

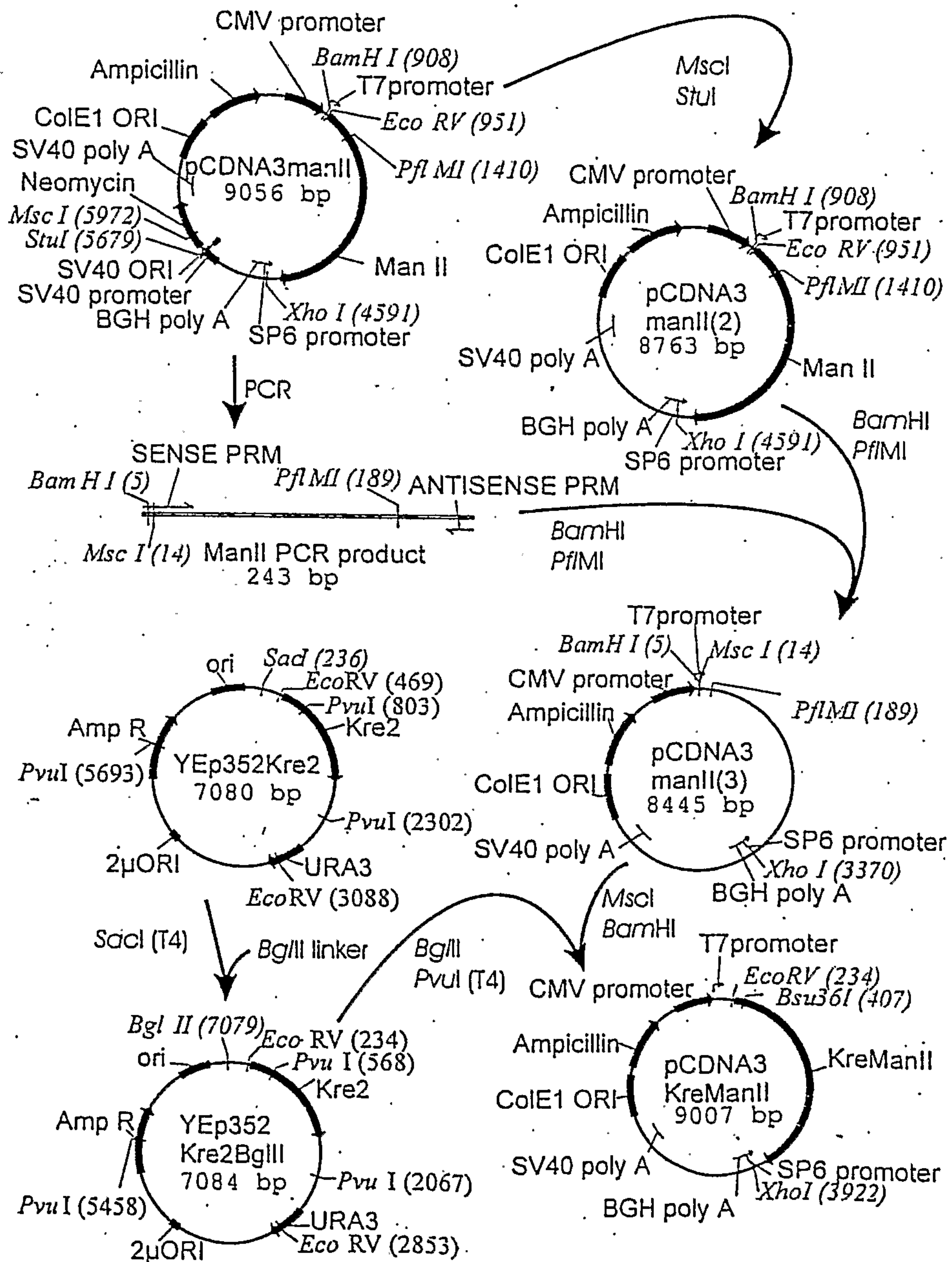


FIG. 30A

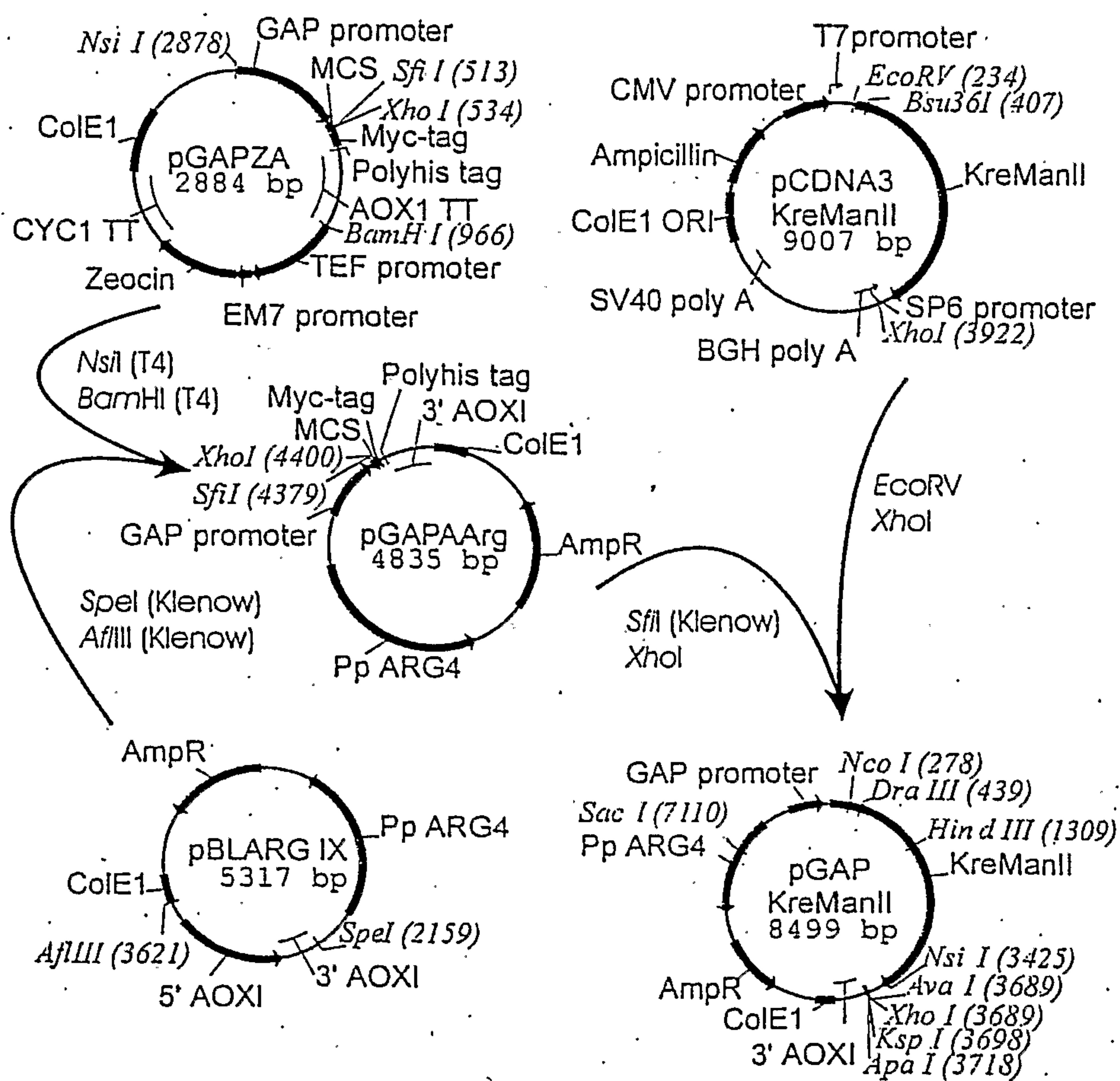


FIG. 30B

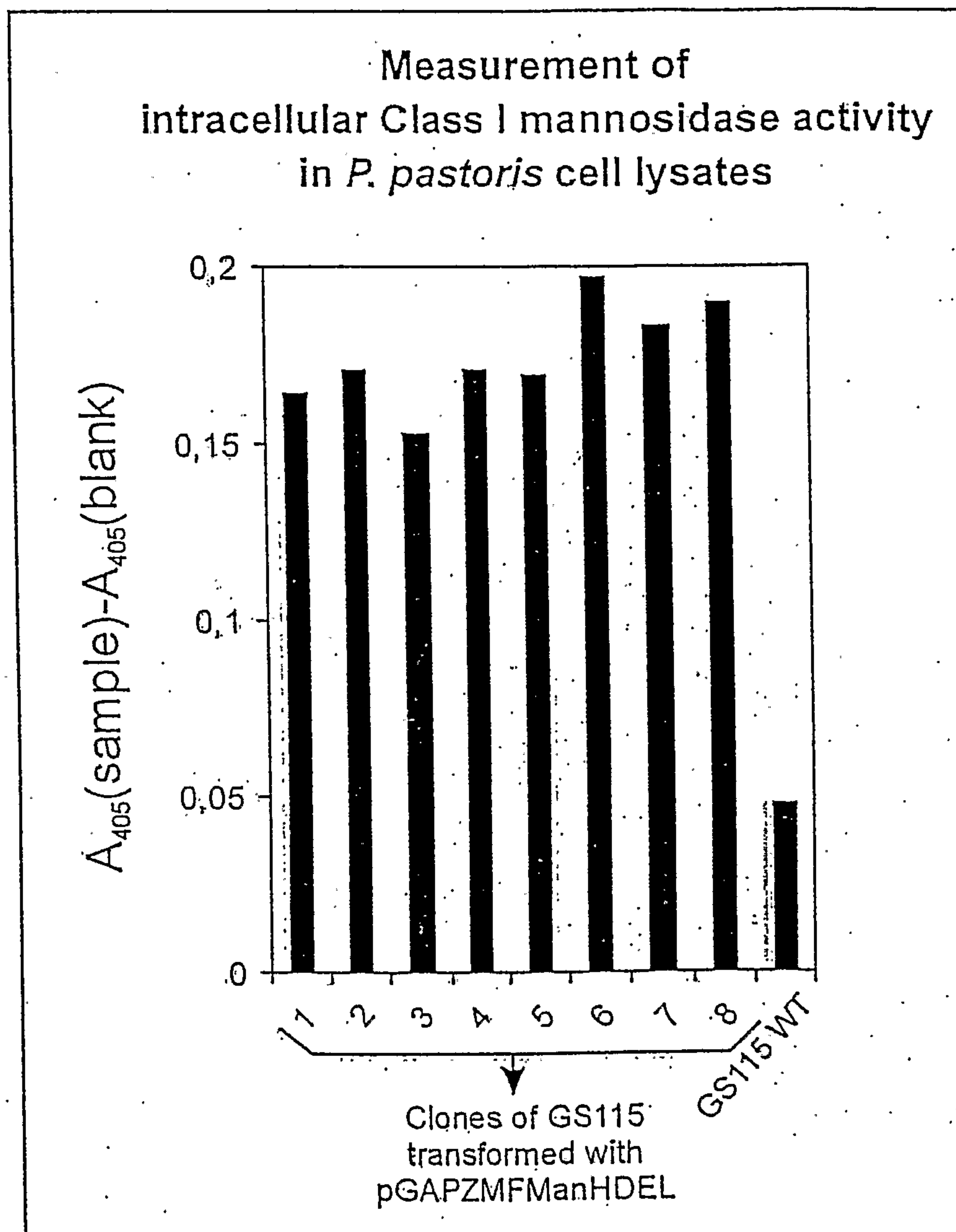
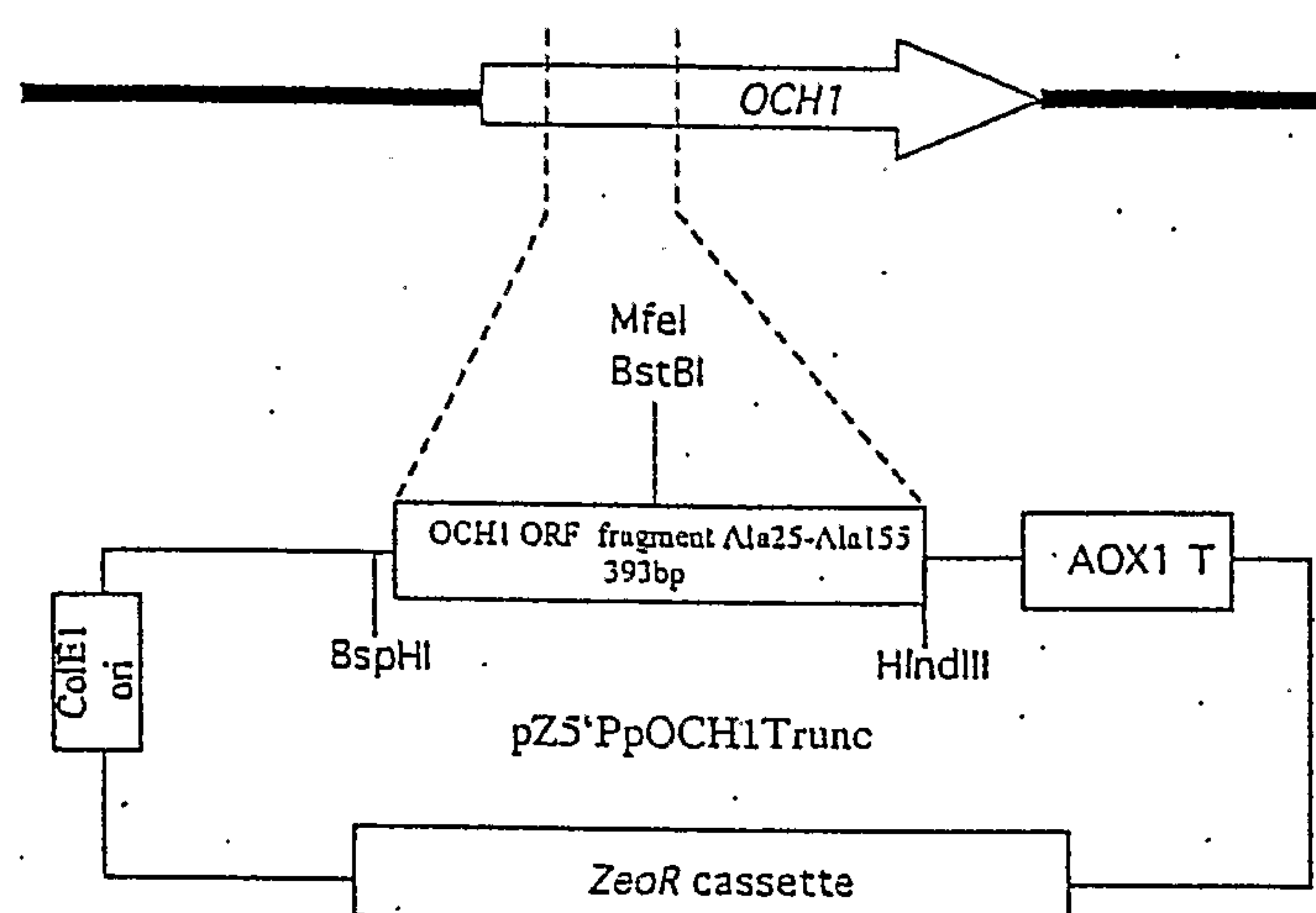
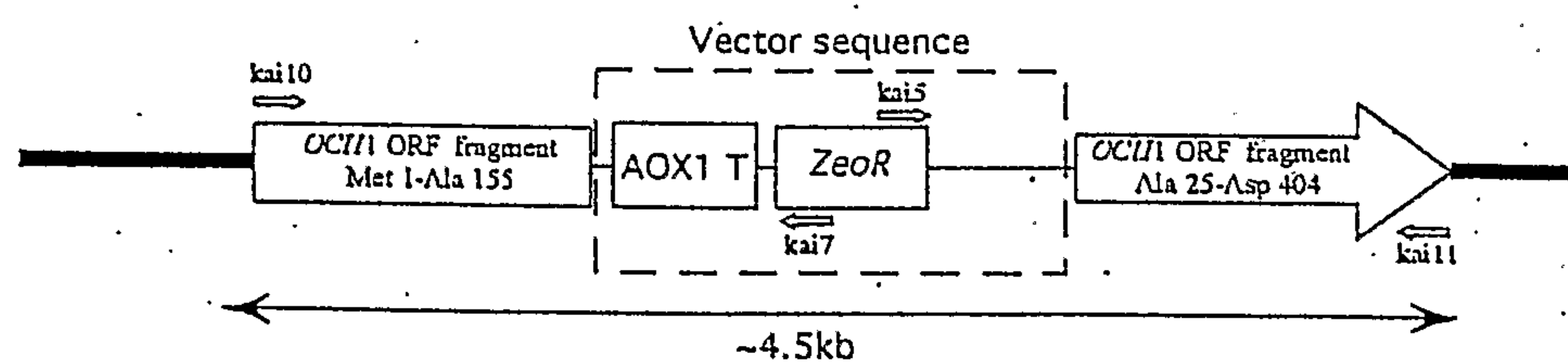


FIG. 31

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

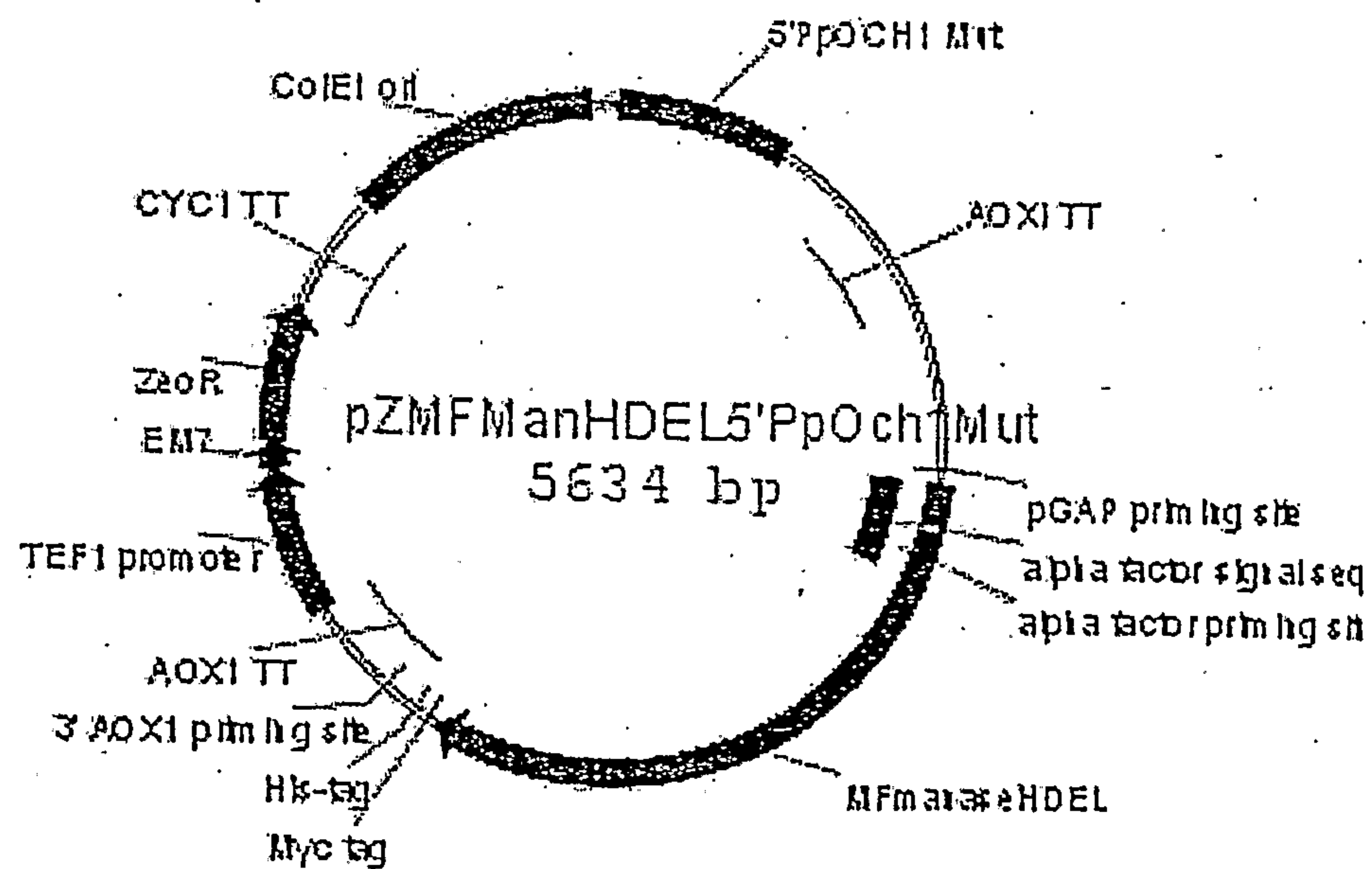


FIG. 33

Clones of GS115 transformed with pZ5'PpOCH1Trunc

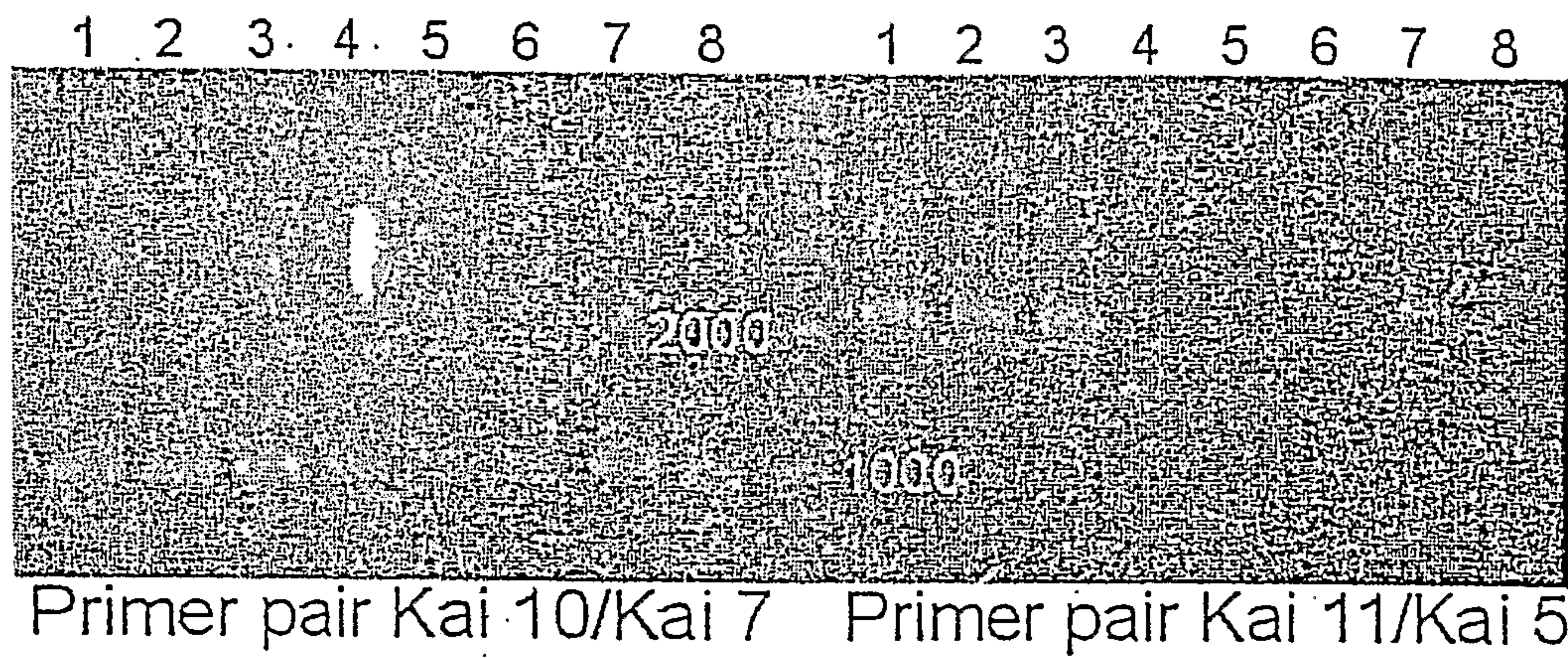
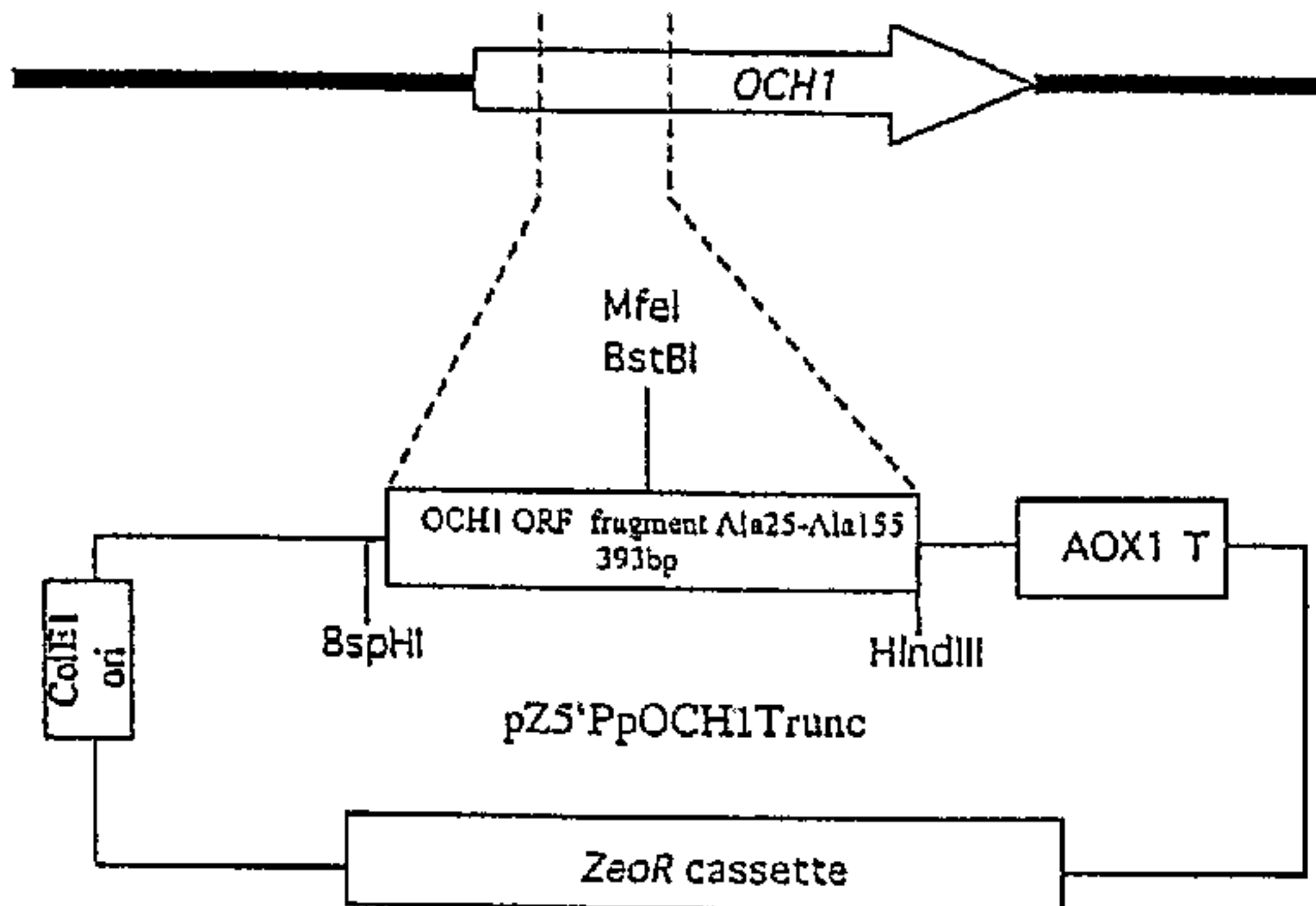
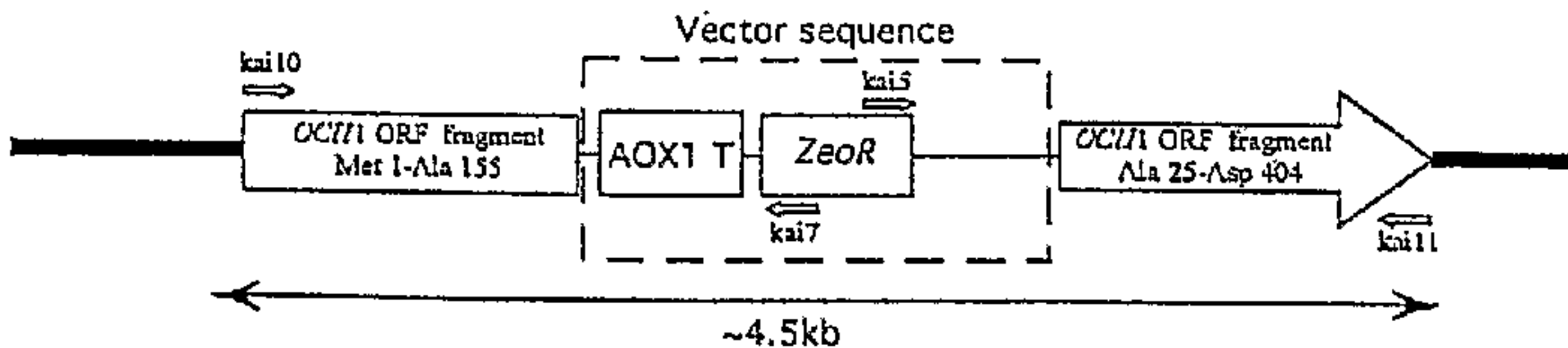


FIG. 34

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