

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
16 March 2006 (16.03.2006)

PCT

(10) International Publication Number
WO 2006/026992 A1

(51) International Patent Classification:

C12P 21/00 (2006.01) C12N 5/10 (2006.01)
C12N 1/14 (2006.01) C12N 15/62 (2006.01)
C12N 9/24 (2006.01) C12N 15/80 (2006.01)
C12N 9/10 (2006.01) C12R 1/645 (2006.01)

(21) International Application Number:

PCT/DK2005/000569

(22) International Filing Date:

7 September 2005 (07.09.2005)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

PA 2004 01346 7 September 2004 (07.09.2004) DK

(71) Applicant (for all designated States except US):
NOVOZYMES A/S [DK/DK]; Krogshøjvej 36, DK-2880
Bagsværd (DK).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **JØRGENSEN,**

Christel, Thea [DK/DK]; Fuglevadsvej 10, DK-2800
Lyngby (DK). **HJORT, Carsten** [DK/DK]; Ravnehus-
parken 6, DK-3500 Værløse (DK).

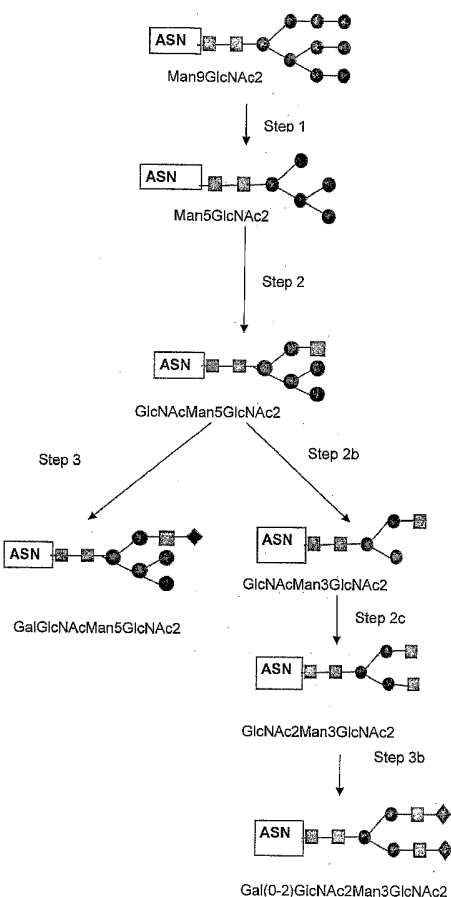
(74) Common Representative: **NOVOZYMES A/S**;
Krogshøjvej 36, DK-2880 Bagsværd (DK).

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

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: ALTERED STRUCTURE OF N-GLYCANS IN A FUNGUS



(57) Abstract: The present invention relates to methods for altering the
structure of N-glycans in fun-gal cells and to fungal cells having an altered
N-glycan structure.

WO 2006/026992 A1



FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

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.

Published:

— *with international search report*

TITLE: ALTERED STRUCTURE OF N-GLYCANS IN A FUNGUS**FIELD OF THE INVENTION**

The present invention relates to a method for altering the structure of N-glycans in a fungus and to a fungal cell having an altered N-glycan structure.

5 BACKGROUND OF THE INVENTION

Glycosylation of proteins is the process of post-translational modification of eukaryotic proteins by attachment of carbohydrates to the protein. Glycosylation is generally divided into; N-glycosylation or O-glycosylation, wherein N-glycosylation is the attachment of carbohydrate(s) (also known as glycans) on the nitrogen atom of an Asparagine (Asn) residue in the protein
10 and O-glycosylation is the attachment of carbohydrates to an oxygen atom of hydroxyls, in particular of a Serine (Ser) or Threonine (Thr) residue in the protein (see e.g. pp. 91-94 of Creighton TE (1993), Proteins; Structure and Molecular Properties, second edition, W.H. Freeman and Company).

Filamentous fungi are widely used to industrially produce proteins, such as enzymes. However, the use of said fungi as production strains for production of pharmaceutical proteins has
15 so far been somewhat limited. One reason for this is that the structure of glycans added to proteins in filamentous fungi is different to that found in mammalian cells and these differences have been shown to be able to elicit an immunological response in mammals, such as humans.

20 Typically the first steps in the N-glycosylation pathway in humans and filamentous fungi, which take place in the endoplasmic reticulum (ER), both lead to the formation of the Man8GlcNAc2 or the Man9GlcNAc2 structure (Man8 or Man9, respectively). In humans Man8 and Man9 is further modified in the Golgi where the following sequential steps take place (Hamilton SR et al. (2003), Science, 301, 1244-1246):

25

- a) Man8 or Man9 is processed by alpha-1,2-mannosidase 1A, 1B and 1C to produce Man5
- b) Man5 is processed by beta-1,2-N-acetylglucosaminyltransferase I to produce GlcNAcMan5
- 30 c) GlcNAcMan5 is processed by Mannosidase II to produce GlcNAcMan3
- d) GlcNAcMan3 is processed by beta-1,2-N-acetylglucosaminyltransferase II (GnTII) to produce GlcNAc2Man3

In yeast more mannose-residues are added to the Man8 or Man9 structures to create so called
35 high mannose structures (see e.g. Hamilton et al. (2003), Science, 301, 1244-1246). The

structure of N-glycans from filamentous fungi is more diverse and only a few have been characterized. Maras M et al. (1999), Glycoconjugate Journal 16, table 2 (page 22) discloses an overview of partially or completely characterised oligosaccharides on glycoproteins from filamentous fungi.

5 US 5,834,251 discloses methods of converting high mannose type glycosylation patterns to hybrid or complex type glycosylation patterns.

WO 01/25406 discloses mannosidase enzymes and use of such enzymes to alter the glycosylation patterns of macromolecules.

WO 02/00879 discloses methods for producing modified glycoproteins.

10 US 5,047,335 discloses a process for controlling the glycosylation of protein in a cell wherein the cell is genetically engineered to produce one or more enzymes which provide internal control of the cell's glycosylation mechanism.

Maras M et al. (1999), FEBS Letters 452, 365-370 discloses in vivo synthesis of complex N-glycans by expression of human N-acetylglucosaminyltransferase I in the filamentous fungus

15 *Trichoderma reesei*.

The present invention provides a method for altering the N-glycan structure of fungi so that it is more similar to the pattern observed in mammals, e.g. humans.

SUMMARY OF THE INVENTION

The present invention relates in a first aspect to a method of altering the structure of N-glycans in a fungus comprising insertion into a fungal cell of a nucleic acid construct or a combination of nucleic acid constructs selected from the group consisting of:

a) a first nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungi

25 b) a second nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase

c) a third nucleic acid construct comprising a nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase

d) a fourth nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase

30 e) a fifth nucleic acid construct comprising a nucleic acid sequence encoding a beta-1,4-galactosyltransferase

In a second aspect the present invention relates to a fungal cell comprising a nucleic acid construct or a combination of nucleic acid constructs selected from the group consisting of:

- 5 a) a first nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungus
- b) a second nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
- 10 c) a third nucleic acid construct comprising a nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase
- d) a fourth nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
- e) a fifth nucleic acid construct comprising a nucleic acid sequence encoding a beta-1,4-galactosyltransferase

15 wherein said cell has an altered N-glycan structure as compared to a parental fungal cell.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows the stepwise altering of the N-glycan structure in a fungal cell of the present invention wherein the glycan-structure before the arrow is the substrate and the glycan-structure after the arrow is the product and wherein "Asn" is an Asparagine residue in a protein, the "square (■)" is a GlcNAc residue, the "circle (●)" is a mannose residue and the "diamond (◆)" is a galactose residue.

Step 1 (the conversion of Man9GlcNAc2 to Man5GlcNAc2) is catalysed by alpha-1,2 mannosidase.

25 Step 2 (the conversion of Man5GlcNAc2 to GlcNAcMan5GlcNAc2) is catalysed by alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase.

Step 3 (the conversion of GlcNAcMan5GlcNAc2 to GalGlcNAcMan5GlcNAc2) is catalysed by beta-1,4-galactosyltransferase.

Step 2b (the conversion of GlcNAcMan5GlcNAc2 to GlcNAcMan3GlcNAc2) is catalysed by mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase.

Step 2c (the conversion of GlcNAcMan3GlcNAc2 to GlcNAc2Man3GlcNAc2) is catalysed by alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase.

Step 3b (the conversion of GlcNAc2Man3GlcNAc2 to Gal(0-2)GlcNAc2Man3GlcNAc2) is catalysed by beta-1,4-galactosyltransferase, wherein the Gal(0-2)GlcNAc2Man3GlcNAc2 structures are described under figure 2.

Figure 2 shows the four different forms of Gal(0-2)GlcNAc2Man3GlcNAc2.

Gal(0)GlcNAc2Man3GlcNAc2 denotes GlcNAc2Man3GlcNAc2 without any Gal residues.

Gal(1)GlcNAc2Man3GlcNAc2 denotes GlcNAc2Man3GlcNAc2 with one Gal residue located on either the 1,3 arm (depicted as the upper arm) or the 1,6 arm (depicted as the lower arm). The two version of the Gal(1)GlcNAc2Man3GlcNAc2 are considered to be equivalents.

Gal(2)GlcNAc2Man3GlcNAc2 denotes GlcNAc2Man3GlcNAc2 with two Gal residues; one on the 1,3 arm and one on the 1,6 arm.

The terms "1,3 arm" and "1,6 arm" are to be understood as referring to that the Gal residue is bound to the GlcNAc residue by a bond between carbon number 1 of the GlcNAc residue and carbon number 3 or 6, respectively of the Gal residue.

DETAILED DESCRIPTION OF THE INVENTION

Definitions and short terms

The term "Man" is in the context of the present invention used as a short term for mannose.

The term "GlcNAc" is in the context of the present invention used as a short term for N-acetylglucosamin.

The term "Gal" is in the context of the present invention used as a short term for a galactose.

The term "Man9" is used as a short term for Man9GlcNAc2, the term "Man8" is used as a short term for Man8GlcNAc2, the term "Man5" is used as a short term for Man5GlcNAc2 and Man3 is used as a short term for Man3GlcNAc2.

The term "ER" is used as short-term for endoplasmic reticulum.

The term "N-glycan" is in the context of the present invention to be understood as Asparagine-linked-glycans, i.e. carbohydrates attached to an Asparagine residue in a polypeptide.

In the context of the present invention, the term "E.C." (Enzyme Class) refers to the internationally recognized enzyme classification system, Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, Academic Press, Inc.

The term "origin" used in the context of amino acid sequences, e.g. proteins, or nucleic acid sequences is to be understood as referring to the organism from which it derives. Said sequence may be expressed by another organism using gene technology methods well known to a person skilled in the art. This also encompasses sequences which have been chemically synthesized. Furthermore, said sequences may comprise minor changes such as codon optimization, i.e. changes in the nucleic acid sequences which do not affect the amino acid sequence.

The term "protein" is in the context of the present invention intended to include polypeptides and combinations of polypeptides.

The term "dominant" when used in reference with a nucleic acid sequence or an allele is to be understood as an allele which determines the phenotype displayed in a heterozygote with another (recessive) allele.

The term "recessive" when used in reference with a nucleic acid sequence or an allele is to be understood as an allele which is obscured in the phenotype of a heterozygote by the dominant allele, often due to inactivity or absence of the product of the recessive allele.

The term "parental cell" is in the context of the present invention to be understood as a cell which is modified according to a method according to the present invention.

The term "open reading frame (ORF)" is in the context of the present invention to be understood with its general meaning, i.e. as a nucleic acid sequence that contains a signal for the start of translation, followed in the correct register by a sufficient length of amino acid-encoding nucleotide triplets to form a protein, followed by a signal for translation.

For purposes of the present invention, alignments of sequences and calculation of homology scores may be done using a full Smith-Waterman alignment, useful for both protein and DNA alignments. The default scoring matrices BLOSUM50 and the identity matrix are used for protein and DNA alignments respectively. The penalty for the first residue in a gap is -12 for proteins and -16 for DNA, while the penalty for additional residues in a gap is -2 for proteins and -4 for DNA. Alignment may be made with the FASTA package version v20u6 (W. R. Pearson and D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85:2444-2448, and W. R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA", Methods in Enzymology, 183:63-98).

Multiple alignments of protein sequences may be made using "ClustalW" (Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-4680). Multiple alignment of DNA sequences may be done using the protein alignment as a template, replacing the amino acids with the corresponding codon from the DNA sequence.

Methods of the present invention

The present invention relates to a method of altering the structure of N-glycans in a fungus comprising insertion into a fungal cell of a nucleic acid construct or a combination of nucleic acid constructs selected from the group consisting of:

- 5 a) a first nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungi
- b) a second nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
- c) a third nucleic acid construct comprising a nucleic acid sequence encoding a man-
10 nosyl-oligosaccharide 1,3-1,6-alpha-mannosidase
- d) a fourth nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosamintransferase
- e) a fifth nucleic acid construct comprising a nucleic acid sequence encoding a beta-1,4-galactosyltransferase

15 The term "altered N-glycan structure" is to be understood as the structure of N-glycan(s) present on a protein produced by a fungal cell of the present invention is different from the N-glycan structure of the same protein produced by a fungal cell which differs from the former cell by having been modified by a method of the present invention and/or by comprising a first, second, third, fourth or fifth or a combination of any of these nucleic acid con-
20 struct of the present invention. In particular the altered N-glycan structure may be a Man5GlcNAc2, or a GlcNAcMan5GlcNAc2, or a GalGlcNAcMan5GlcNAc2, or a GlcNAcMan3GlcNAc2, or a GlcNAc2Man3GlcNAc2 or a Gal(0-2)GlcNAcMan3GlcNAc2. The altered N-glycan structure may also be further modified by other enzymes naturally present in the fungal cell, which are capable of e.g. adding other sugar resi-
25 dues to the N-glycan. Thus in another embodiment the altered N-glycan structure may in particular comprise Man5GlcNAc2, or GlcNAcMan5GlcNAc2, or GalGlcNAcMan5GlcNAc2, or GlcNAcMan3GlcNAc2, or GlcNAc2Man3GlcNAc2 or Gal(0-2)GlcNAcMan3GlcNAc2. Methods for detecting the different glycan-structures may be performed as described in Jackson P and Gallagher JT (1997), A laboratory guide to Glycoconjugate analysis, Biomethods, vol. 9, Birk-
30 häuser, Basel, Switzerland.

Thus a method of the present invention comprises insertion into a fungal cell of a first, or a second, or a third, or a fourth or a fifth nucleic acid construct of the present invention.

In a particular embodiment of the present invention said method may comprise inser-
35 tion of a combination of two or more of the above mentioned, first, second, third, fourth and fifth nucleic acid constructs. Although in principle any combination of said nucleic acid construct may be inserted into a fungal cell, the inventors of the present invention in particular foresee two main routes for altering the N-glycan structure of a fungal cell by insertion of a

combination of the above mentioned, first, second, third, fourth and fifth nucleic acid constructs. These two main routes for altering the N-glycan structure are outlined in figure 1. Figure 1 shows that insertion of a combination of the first and the second nucleic acid construct is the same for both routes, but after the second nucleic acid construct the routes diverge.

5 Furthermore, both routes include as the final step the insertion of a fifth nucleic acid construct.

In a particular embodiment of the present invention a method according to the present invention may further comprise modification of an endogenous gene encoding an enzyme capable of catalyzing the binding of mannose by alpha-1,6-bonds to the 1,3-arm of Man9GlcNAc2 so that fewer mannose residues are bound by alpha-1,6-bonds to the
10 Man9GlcNAc2 structure of proteins than in a parental cell. In the context of the present invention the term "fewer" is to be understood so that the fraction of a particular protein which comprises a mannose residue bound by an alpha-1,6-bond to Man9GlcNAc2 is lower when said protein is produced by a cell in which such an endogenous gene has been modified compared to when said protein is produced by a parental cell. The presence of a mannose residue bound
15 by an alpha-1,6-bond to Man9GlcNAc2 may be measured as described in Jackson P and Gallagher JT (1997), A laboratory guide to Glycoconjugate analysis, Biomethods, vol. 9, Birkhäuser, Basel, Switzerland

The presence of a mannose residue bound by an alpha-1,6-bond to the 1,3 arm of Man9GlcNAc2 is believed to inhibit or at least partly inhibit the hydrolysis of the terminal 1,2-linked alpha-D-mannose in Man9GlcNAc2 to produce Man5GlcNAc2 catalysed by alpha-1,2-mannosidase.
20

In one embodiment of the present invention a fungal cell with an altered N-glycan structure created by a method of the present invention may in particular be used as a host for producing a protein with an altered N-glycan structure. These fungal cells may in particular be
25 used for producing a mammalian protein as the altered N-glycan structure of a fungal cell of the present invention is more similar to that found in mammalian cells than the N-glycan structure of a fungal cell which has not been modified.

In another embodiment of the present invention a fungal cell with an altered N-glycan structure created by a method of the present invention may in particular be used to create a
30 heterokaryon with an altered N-glycan structure. Said heterokaryon with an altered N-glycan structure may in a particular embodiment be used as a host for producing a protein with an altered N-glycan structure. Particularly, said heterokaryon may be used for producing a mammalian protein as the altered N-glycan structure of a fungal heterokaryon of the present invention is more similar to that found in mammalian cells than the unaltered N-glycan structure of a
35 fungal heterokaryon.

If the fungal cell with an altered N-glycan structure created by a method of the present invention is used as a host for production of a protein, then it is advantageous if the combina-

tion of insertion of nucleic acid constructs into said cell creates a stepwise alteration of the N-glycan structure as outlined in figure 1. Thus in this case if the route depicted to the left in figure 1 is followed (step 3) then the following combinations of nucleic acid constructs may in particular be inserted into a fungal cell:

- 5 i) a combination of the first and second nucleic acid construct; or
- ii) a combination of the first, second and fifth nucleic acid construct.

If the route depicted to the right in figure 1 is to be followed (steps 2b-3b) then the following combinations of nucleic acid constructs may in particular be inserted into a fungal cell:

- i) a combination of the first and second nucleic acid construct;
- 10 ii) a combination of the first, second and third nucleic acid construct;
- iii) a combination of the first, second, third and fourth nucleic acid construct; or
- iv) a combination of the first, second, third, fourth and fifth nucleic acid construct.

If two or more fungal cells of the present invention are fused to create a heterokaryon with an altered N-glycan structure the insertion into each individual cell of a combination of nucleic acid constructs of the present invention is envisioned to be different from that of a single cell. For modification of single cell it is necessary to insert each of the relevant nucleic acid constructs into said cell. However, the inventors of the present invention believe that the alteration of the N-glycan structure in one fungal cell will result in the same alteration of a heterokaryon if said cell is fused with one or more other cells to create a heterokaryon. Thus if a combination of nucleic acid sequences of the present invention are to be inserted into a fungal cell and said cell is fused to another cell to create a heterokaryon the insertion of nucleic acid constructs into that particular fungal cell may not necessarily be stepwise as indicated in figure 1. However, the overall modification of the N-glycan structure of the heterokaryon may in particular be stepwise as outlined in figure 1.

25 For example if two genetically different fungal cells of the present invention are used to create a heterokaryon which comprises two or more nucleic acid constructs of the present invention the following combinations for modification of each of the cells are envisioned (the possibility of inserting all of the nucleic acid constructs into one of the cells and none into the other also exists for each of below situations, but is not mentioned):

30

For a heterokaryon comprising a first and second nucleic acid construct

- A first nucleic acid construct may be inserted into one of the fungal cells, while the second nucleic acid construct may be inserted into the other fungal cell.

35

For a heterokaryon comprising a first, second and fifth nucleic acid construct

- A first and second nucleic acid construct may be inserted into one of the fungal cells, while the fifth nucleic acid construct may be inserted into the other fungal cell.

- A first and fifth nucleic acid construct may be inserted into one of the fungal cells, while the second nucleic acid construct may be inserted into the other fungal cell.
- A second and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first nucleic acid construct may be inserted into the other fungal cell.

5

For a heterokaryon comprising a first, second and third nuclei acid construct

- A first and second nucleic acid construct may be inserted into one of the fungal cells, while the third nucleic acid construct may be inserted into the other fungal cell.
- A first and third nucleic acid construct may be inserted into one of the fungal cells, while the second nucleic acid construct may be inserted into the other fungal cell.
- A second and third nucleic acid construct may be inserted into one of the fungal cells, while the first nucleic acid construct may be inserted into the other fungal cell.

10

For a heterokaryon comprising a first, second, third and fourth nuclei acid construct

- A first, second and third nucleic acid construct may be inserted into one of the fungal cells, while the fourth nucleic acid construct may be inserted into the other fungal cell.
- A first, second and fourth nucleic acid construct may be inserted into one of the fungal cells, while the third nucleic acid construct may be inserted into the other fungal cell.
- A first, third and fourth nucleic acid construct may be inserted into one of the fungal cells, while the second nucleic acid construct may be inserted into the other fungal cell.
- A second, third and fourth nucleic acid construct may be inserted into one of the fungal cells, while the first nucleic acid construct may be inserted into the other fungal cell.
- A first and second nucleic acid construct may be inserted into one of the fungal cells, while the third and fourth nucleic acid construct may be inserted into the other fungal cell.
- A first and third nucleic acid construct may be inserted into one of the fungal cells, while the second and fourth nucleic acid construct may be inserted into the other fungal cell.
- A first and fourth nucleic acid construct may be inserted into one of the fungal cells, while the second and third nucleic acid construct may be inserted into the other fungal cell.

15

20

25

30

For a heterokaryon comprising a first, second, third, fourth and fifth nuclei acid construct

- A first, second, third and fourth nucleic acid construct may be inserted into one of the fungal cells, while the fifth nucleic acid construct may be inserted into the other fungal cell.

35

- A first, second, third and fifth nucleic acid construct may be inserted into one of the fungal cells, while the fourth nucleic acid construct may be inserted into the other fungal cell.
- 5 • A first, second, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the third nucleic acid construct may be inserted into the other fungal cell.
- A first, third, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the second nucleic acid construct may be inserted into the other fungal cell.
- 10 • A second, third, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first nucleic acid construct may be inserted into the other fungal cell.
- A first, second and third nucleic acid construct may be inserted into one of the fungal cells, while the fourth and fifth nucleic acid construct may be inserted into the other fungal cell.
- 15 • A first, second and fourth nucleic acid construct may be inserted into one of the fungal cells, while the third and fifth nucleic acid construct may be inserted into the other fungal cell.
- A first, second and fifth nucleic acid construct may be inserted into one of the fungal cells, while the third and fourth nucleic acid construct may be inserted into the other fungal cell.
- 20 • A first, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the second and third nucleic acid construct may be inserted into the other fungal cell.
- 25 • A first, third and fourth nucleic acid construct may be inserted into one of the fungal cells, while the second and fifth nucleic acid construct may be inserted into the other fungal cell.
- A first, third and fifth nucleic acid construct may be inserted into one of the fungal cells, while the second and fourth nucleic acid construct may be inserted into the other fungal cell.
- 30 • A second, third and fourth nucleic acid construct may be inserted into one of the fungal cells, while the first and fifth nucleic acid construct may be inserted into the other fungal cell.
- A second, third and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first and fourth nucleic acid construct may be inserted into the other fungal cell.
- 35 • A second, third and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first and fourth nucleic acid construct may be inserted into the other fungal cell.

- A second, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first and third nucleic acid construct may be inserted into the other fungal cell.
- A third, fourth and fifth nucleic acid construct may be inserted into one of the fungal cells, while the first and second nucleic acid construct may be inserted into the other fungal cell.

If the methods of the present invention are used to create a heterokaryon comprising only one nucleic acid construct of the present invention said nucleic acid construct may be inserted into a fungal cell which then subsequently may be fused with one or more other fungal cells which have not be modified according to a method of the present invention. However, if the heterokaryon is to comprise one or more of the first, second, third, fourth and fifth nucleic acid constructs of the present invention then the insertion of the different nucleic acid constructs may in particular be divided between the different fungal cells which are to be fused to form a heterokaryon. An advantage of dividing the nucleic acid constructs between the different cells which are fused to form the heterokaryon (in contrast to inserting them all into the same cell) is that the same selection marker(s) may be used in each of the cells. Thus a smaller number of different selection markers are necessary as compared to if all nucleic acid constructs were to be inserted into the same cell. Although, the fungal cell in a method of the present invention in principle may be a heterokaryon, i.e. the nucleic acid constructs of the present invention may in principle be inserted into a heterokaryon, the inventors of the present invention believe that it is advantageous to insert the nucleic acid constructs into the single fungal cells and then subsequently fuse said cells to form a heterokaryon, as the heterokaryon is probably be more unstable than single cells thereby making introduction of nucleic acid constructs more difficult.

If a heterokaryon is formed by fusion of more than two genetically different cells then the distribution of nucleic acid constructs which may be inserted into each of the genetically different cells may be different than that suggested above for two genetically different cells. The possible different combinations for insertion of the nucleic acid into each of the cells may be combined similarly as described above for two cells.

If a method of the present invention further comprises modification of an endogenous gene encoding an enzyme capable of catalyzing the binding of mannose by alpha-1,6-bonds to Man5GlcNAc2 and the purpose is to obtain a heterokaryon with such a modified gene, it may be an advantage to modify said gene in each of genetically different cells which are fused to create a heterokaryon. The nuclei of a heterokaryon fungal cell does not fuse, thus if at least one of the genetically different cells which are fused to create said heterokaryons comprises such an endogenous gene which has not been modified it is envisioned that the heterokaryon

will express a functional enzyme encoded by said gene, whereby binding of mannose by alpha-1,6-bonds to Man5GlcNAc2 should still be possible to take place in the heterokaryon.

Fungal cell

The present invention relates both to methods for altering the structure of N-glycans in a fungus and to a fungal cell comprising a first, second, third, fourth or fifth nucleic acid construct or a combination thereof.

Reference to "fungal cell" is in the following intended to encompass both parental fungal cells and fungal cells of the present invention.

"Fungi" as used herein includes the phyla *Ascomycota*, *Basidiomycota*, *Chytridiomycota*, and *Zygomycota* (as defined by Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the *Oomycota* (as cited in Hawksworth et al., 1995, In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK, page 171) and all mitosporic fungi (Hawksworth et al., 1995, supra). Representative groups of *Ascomycota* include, e.g., *Neurospora*, *Eupenicillium* (= *Penicillium*), *Emericella* (= *Aspergillus*), *Eurotium* (= *Aspergillus*), and the true yeasts listed below. Examples of *Basidiomycota* include mushrooms, rusts, and smuts. Representative groups of *Chytridiomycota* include, e.g., *Allomyces*, *Blastocladiella*, *Coelomomyces*, and aquatic fungi. Representative groups of *Oomycota* include, e.g., Saprolegniomycetous aquatic fungi (water molds) such as *Achlya*. Examples of mitosporic fungi include *Aspergillus*, *Penicillium*, *Candida*, and *Alternaria*. Representative groups of *Zygomycota* include, e.g., *Rhizopus* and *Mucor*.

In one embodiment, the fungal cell is a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). The ascosporogenous yeasts are divided into the families *Spermophthoraceae* and *Saccharomycetaceae*. The latter is comprised of four subfamilies, *Schizosaccharomycoideae* (e.g., genus *Schizosaccharomyces*), *Nadsonioideae*, *Lipomycoideae*, and *Saccharomycoideae* (e.g., genera *Pichia*, *Kluyveromyces* and *Saccharomyces*). The basidiosporogenous yeasts include the genera *Leucosporidium*, *Rhodospordium*, *Sporidiobolus*, *Filobasidium*, and *Filobasidiella*. Yeasts belonging to the Fungi Imperfecti are divided into two families, *Sporobolomycetaceae* (e.g., genera *Sorobolomyces* and *Bullera*) and *Cryptococcaceae* (e.g., genus *Candida*). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980. The biology of yeast and manipulation of yeast genetics are well known in the art (see, e.g., Biochemistry and Genetics of Yeast, Bacil, M., Horecker, B.J., and Stopani, A.O.M., editors, 2nd edition, 1987; The Yeasts, Rose, A.H., and Harrison,

J.S., editors, 2nd edition, 1987; and The Molecular Biology of the Yeast *Saccharomyces*, Strathern et al., editors, 1981).

In a more particular embodiment, the yeast cell is a cell of a species of *Candida*, *Kluyveromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Pichia*, or *Yarrowia*. In a most particular
 5 embodiment, the yeast cell is a *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* or *Saccharomyces oviformis* cell. In another embodiment, the yeast cell is a *Kluyveromyces lactis* cell. In yet another embodiment, the yeast cell is a *Yarrowia lipolytica* cell.

10 In another embodiment, the fungal cell is a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision *Eumycota* and *Oomycota* (as defined by Hawksworth et al., 1995, supra). Filamentous fungi are characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegeta-
 15 tive growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

The filamentous fungal cell may be a cell of a species of, but not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium*, or *Trichoderma*.

20 In a most particular embodiment, the filamentous fungal cell is an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* or *Aspergillus oryzae* cell. In another most particular embodiment, the filamentous fungal cell is a *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, or *Fusarium venenatum* cell. In still another embodiment, the filamentous fun-
 25 gal cell is a *Fusarium venenatum* (Nirenberg sp. nov.) cell. In another particular embodiment, the filamentous host cell is a *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*,
 30 *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

If a fungal cell of the present invention, wherein the structure of N-glycans has been modified according to a method of the present invention, is used to create a heterokaryon the
 35 fungal cell may in particular be a yeast or filamentous fungal cell. However, if said fungal cell is not fused with one or more other cells to create a heterokaryon but used e.g. for expression of a protein with an altered N-glycan structure it may in particular be a filamentous fungal cell.

In the context of the present invention a "heterokaryon" is to be understood as a cell with at least two genetically different nuclei. Heterokaryons derive from fusion of two or more genetically different cells wherein the nuclei of said cells does not fuse resulting in a cell comprising two or more nuclei.

5 In particular the heterokaryon fungus may be a filamentous heterokaryon fungus or it may be a yeast heterokaryon. The heterokaryon fungus may be formed naturally between two or more fungi or it may be made artificially. When two or more genetically different fungi fuse the nuclei of each of the individual cells come to coexist in a common cytoplasm. One method to select for heterokaryons is to fuse two or more genetically different cells which each com-
10 prise a genome with a characteristic which renders the survival of each cell dependent on presence of the nuclei from the other cell. For example if two genetically different cells which each depends on a particular nutrient for survival and at the same time is independent of the nutrient the other cell depends on for survival is cultured in the a medium lacking both of the nutrients this will make only cells which arise as a fusion between each of the genetically dif-
15 ferent cells able to survive in this medium.

The heterokaryon filamentous fungus of the present invention may in particular contain nuclei from cells that are homozygous for all heterokaryon compatibility alleles. At least ten chromosomal loci have been identified for heterokaryon incompatibility: het-c, het-d, het-e, het-i, het-5, het-6, het-7, het-8, het-9 and het-10, and more probably exist (see e.g. Perkins et al.,
20 "Chromosomal Loci of *Neurospora crassa*", *Microbiological Reviews* (1982) 46: 462-570, at 478).

Formation of the heterokaryon filamentous fungus may in particular be performed by hyphal or protoplast fusion.

In particular the heterokaryon filamentous fungus of the present invention may be made
25 by fusion of hyphae from two different strains of filamentous fungi, wherein the first nuclei of one the strains contains a genome that results in a characteristic which renders the fungus dependent on the presence of the second nucleus from the other fungus for survival under the conditions provided for fusion to form the heterokaryon, and vice versa. Thus the nucleus of each strain of filamentous fungus confers a characteristic which would result in the failure of
30 the fungus in which it is contained to survive under the culture conditions unless the nucleus from the other filamentous fungus is also present. Examples of characteristics which may be used to render the strains of filamentous fungi dependent on each other include, but are not limited to, a nutritional requirement, resistance to toxic compounds and resistance to extreme environmental conditions. For example if a first strain which requires the presence of a particu-
35 lar nutrient is cultured on a medium lacking said nutrient along with a second strain which does not require said nutrient for survival, the nucleus of the second strain will confer the ability of a fusion of the two strains to survive even in the absence of the particular nutrient. Furthermore,

if the second strain similarly requires the presence of a particular nutrient different from the nutrient required by the first strain, then only fusions comprising a nucleus from each strain will survive in a medium lacking both of said nutrients.

5 Methods for formation of a heterokaryon filamentous fungus are described in US 6,543,745.

Examples of filamentous fungi which may be fused to form a heterokaryon filamentous fungus include those described below as filamentous fungal host cells. In particular different strains of *Aspergillus*, e.g. *A.oryzae* or *A.niger*, *Fusarium* or *Trichoderma* may be used to form a heterokaryon filamentous fungus. In principle, more than two different strains of filamentous
10 fungi may be used to form a heterokaryon, such as 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 different strains. In particular the heterokaryon filamentous fungus of the present invention is formed by fusion of two different strains of filamentous fungi.

Examples of characteristics which make each of the strains of fungi (that are fused to form a heterokaryon filamentous fungus) dependent on the presence of the nucleus from the
15 other fungus for survival under the conditions provided for the fusion include the selectable markers described above. In particular said characteristic may be a characteristic that makes the fungus autotroph. The culture media used for fusion of the different strains of fungi to form a heterokaryon filamentous fungus may be any media which does not complement the particular characteristic of the fungi. Examples of such media are well known to a person skilled in the
20 art as they are generally used to select for recombinant fungi. In the case of fusion of different fungi, however, at least two different characteristics/markers are used for the selection. Examples of characteristics or markers which may be used include those described above as selectable markers useful for the nucleic acid construct. Examples of genes which may make a fungus autotroph include, but are not limited to: *pyrG*, *hemA*, *niaD*, *tpi*, *facC*, *gala*, *biA*, *lysB*,
25 *sC*, *methG* and *phenA*. Thus if a fungus is negative for at least one of these genes said gene may be used as a selectable marker.

In another embodiment the heterokaryon fungus is a yeast heterokaryon. In yeast formation of heterokaryons has been described as a result of a defect in karyogamy (nuclear fusion) during conjugation of two (or more) haploid cells (Olson BL and Siliciano PG, 2003, Yeast, **20**, 893-903). A mutation in the *kar1* gene in one of the cells to be fused has been described as sufficient to block nuclear fusion during mating (Olson BL and Siliciano PG, Yeast, 2003, **20**, 893-903). Thus in a particular embodiment of the present invention at least one of the cells involved in formation of a yeast heterokaryon of the present invention comprises a
30 mutation in the *kar1* gene, e.g. a loss-of-function mutation, such as a null mutation, a non-sense or a missense mutation, whereby the nuclei of said cell can not fuse with the nuclei from another cell during mating. The cloning of *kar1* in yeast is described in Rose MD and Fink GR, Cell, 1987, **48**, 1047-1060.

Methods of transformation of fungi are well known and may be performed as described below for the fungal host cells. Conditions for culturing a heterokaryon fungus are similar to those for culturing the fungi that it arises from with the exception that the heterokaryon is cultured in a medium selecting for at least two different characteristics. The selection for at least two different characteristics needs at least to be present during formation of the heterokaryon but usually it also an advantage to keep this selection pressure, i.e. the selection for at least two characteristic during subsequent culturing to ensure the stability of the heterokaryon. Methods for culturing fungi are well known to a person skilled in the art.

In a particular embodiment the fungal cell may comprise one of the following combinations of nucleic acid constructs of the present invention:

- i) a combination of the first and second nucleic acid construct; or
- ii) a combination of the first, second and fifth nucleic acid construct

Or if the other route of altered N-glycan structure is to be obtained it may comprise one of the following combinations:

- i) a combination of the first and second nucleic acid construct;
- ii) a combination of the first, second and third nucleic acid construct;
- iii) a combination of the first, second, third and fourth nucleic acid construct; or
- iv) a combination of the first, second, third, fourth and fifth nucleic acid construct.

If as mentioned previously in the section of "Methods of the invention" the fungal cell is used for creation of a heterokaryon then the combination of nucleic acid constructs described in that section may in particular be present in such a fungal cell.

If the fungal cell of the present invention is a heterokaryon and if it comprises more than one nucleic acid construct of the present invention it may either comprise the nucleic acid constructs in the same nuclei or it may in particular comprise the nucleic acid constructs in different nuclei. The combination of such nucleic acid constructs in each of the nucleic may as described in the "Methods" section.

In a particular embodiment of the present invention at least one endogenous gene of the fungal cell has been modified so that fewer mannose residues are bound by alpha-1,6-bonds to the 1,3 arm of the Man9GlcNAc2 structure of proteins than in a parental cell. More particularly, said endogenous gene may be Mnn9.

N-glycosylation enzymes and nucleic acid constructs

It is an object of the present invention to alter the structure of N-glycans in a fungal cell, such as a filamentous fungus or in a heterokaryon fungus, wherein the latter in particular may be a filamentous heterokaryon fungus or a yeast heterokaryon. The structure of N-glycans is different in fungi compared to mammalian cells. However, fungal cells are widely used for

industrial production of proteins and it would therefore be an advantage if the N-glycan structure in the fungal cell could be altered so that it resembles that found in mammalian cell.

Figure 1 shows an outline of the different steps of the glycosylation pathway which it is an object of the present invention to introduce and/or modify one or more of in a fungal cell.

5 Each step shows the processing of a particular N-glycan structure by a N-glycosylation enzyme to another N-glycan structure.

The present invention relates both to methods for altering the N-glycan structure of a fungal cell and to such fungal cell, e.g. to fungal cell with an altered N-glycan structure.

10 The N-glycan structure in a fungal cell is in the present invention altered by insertion of a nucleic acid construct comprising a nucleic acid sequence encoding at least one of enzymes shown in figure 1, i.e. which are selected from the group consisting of alpha-1,2-mannosidase, alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase, mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase, alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase and beta-1,4-galactosyltransferase.

15 Cells, in particular eukaryotic cells are compartmentalized with different processes taking place in different compartments. The formation of the N-glycan structure is believed to take place both in the ER and Golgi, thus for a correct formation of the N-glycan structure it is important that the N-glycosylation enzymes involved in formation of said structure are located in the compartment where the reaction they catalyze takes place.

20 First nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase

The first nucleic acid construct of the present invention comprises a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungi.

25 In the context of the present invention the term "alpha-1,2-mannosidase" is to be understood as enzymes capable of catalyzing the hydrolysis of the terminal 1,2-linked alpha-D-mannose residues in the oligosaccharide Man9GlcNAc2, so as to produce Man5GlcNAc2.

The alpha-1,2-mannosidase used in the present invention derives from a filamentous fungi. In a particular embodiment it may derive from a strain of *Aspergillus*, e.g. *A.nidulans* or *A.oryzae*. More particularly it may comprise the amino acid sequence shown in SEQ ID NO: 2 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 2. In another embodiment it may consist of the amino acid sequence shown in SEQ ID NO: 2 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 2.

35 Thus the nucleic acid sequence encoding the alpha-1,2-mannosidase may in particular comprise the nucleic acid sequence shown in SEQ ID NO:1, or it may be comprise a nucleic

acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 1. In another embodiment it may consist of the nucleic acid sequence shown in SEQ ID NO: 1 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 1.

In another embodiment it may comprise the CDS shown in SEQ ID NO:1, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 1. In another embodiment it may consist of the CDS shown in SEQ ID NO: 1 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 1.

For a correct functioning of the alpha-1,2-mannosidase, i.e. catalysing the hydrolysis of terminal 1,2-linked alpha-D-mannose residues in the oligosaccharide Man9GlcNAc2 so as to produce Man5GlcNAc2, it is an advantage if said enzyme is located in the ER.

Alpha-1,2-mannosidases are present in filamentous fungal cells but there they are usually not localized in the ER and will thus only be present in the ER transiently during secretion of the alpha-1,2-mannosidase. So the alpha-1,2-mannosidase is generally not present in the ER in large enough quantities to enable the conversion of all Man9GlcNAc2 structures to Man5GlcNAc2. In a particular embodiment of the present invention an ER retention signal having the amino acid sequence His-Asp-Glu-Leu or Lys-Asp-Glu-Leu may be fused to the carboxy terminal of the alpha-1,2-mannosidase, which means that the amount of alpha-1,2-mannosidase in the ER or the fraction of glycosylated proteins comprising the Man5GlcNAc2 structure is higher in a fungal cell of the present invention (wherein a first nucleic acid construct has been inserted) than in a parental cell. The amount of alpha-1,2-mannosidase may be measured by e.g. isolating the enzyme from the cell and measuring the amount by e.g. gel-electrophoresis or by measuring the activity of the enzyme. Such methods are known to a person skilled in the art. The amount of Man5GlcNAc2 may be measured as described in Jackson P and Gallagher JT (1997), A laboratory guide to Glycoconjugate analysis, Biomethods, vol. 9, Birkhäuser, Basel, Switzerland

As the alpha-1,2-mannosidase in particular may be located in the ER the first nucleic acid construct may in particular further comprise a nucleic acid sequence encoding the amino acid sequence HDEL (His-Asp-Glu-Leu) or KDEL (Lys-Asp-Glu-Leu). An ER retention signal is a signal present in a protein which makes said protein resides in the ER, so far only short amino acid sequences located at the 3'-end of the protein has been identified as being ER retention signals. Thus said ER retention signal may particularly be located downstream of the 3'-end of the nucleic acid sequence encoding an alpha-1,2-mannosidase, typically the nucleic acid sequence encoding the ER retention signal may be located right next to the nucleic acid sequence encoding a protein as it is translated as part of the protein.

To over-express the alpha-1,2-mannosidase the first nucleic acid construct may in particular comprise a constitutive promoter, such as one of those mentioned below, more particularly the *tpiA* or the *gpd* promoter. The *tpiA* promoter has been described in McKnight GL et al, 1986, Cell, 46 (1), pp. 143-147 and the *gpd* promoter has been described in Punt PJ et al, 1990, Gene, 93, pp. 101-109.

Second nucleic acid construct

The second nucleic acid construct of the present invention comprises a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (GnT1). Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferases have been classified as belonging to E.C. 2.4.1.101 which are enzymes capable of catalyzing the following reaction: UDP-N-acetyl-D-glucosamine + 3-(alpha-D-mannosyl)-beta-D-mannosyl-R -> UDP + 3-(2-[N-acetyl-beta-D-glucosaminyl]-alpha-D-mannosyl)-beta-D-mannosyl-R; wherein R represents the remainder of the N-linked oligosaccharide in the glycoprotein acceptor and UDP is uridine diphosphate. In the context of the present invention the short term GnT1 may be used for alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase.

The function of GnT1 with regard to the modification of N-glycans is to catalyse the addition of N-acetylglucosamine to Man5GlcNAc2 to produce GlcNAcMan5GlcNAc2.

In a particular embodiment of the present invention the GnT1 may be of a mammalian origin as only mammalian GnT1's have been identified so far. However, the present invention is not limited to GnT1's of mammalian origin as it is envisioned that GnT1's from other organisms may be identified. For example, the GnT1 may be of a human or rat origin. More particularly GnT1 may comprise the amino acid sequence shown in SEQ ID NO: 4 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 4. In another embodiment it may consist of the amino acid sequence shown in SEQ ID NO: 4 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 4.

The nucleic acid sequence encoding the GnT1 may in particular comprise the nucleic acid sequence shown in SEQ ID NO: 3, or it may comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 3. In another embodiment it may consist of the nucleic acid sequence shown in SEQ ID NO: 3 or a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 3.

In another embodiment it may comprise the CDS shown in SEQ ID NO:3, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 3. In another embodiment it may consist of the CDS shown in SEQ ID NO: 3 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 3.

As it is not trivial to express mammalian proteins in fungi a synthetic gene encoding a mammalian GnT1 may in particular be used. The sequence of the synthetic gene may be identical to e.g. that of the cDNA of a mammalian such as the mRNA or CDS sequence of the human GnT1 shown in SEQ ID NO: 3. However, in a particular embodiment the sequence may have been modified so that it is more similar to the sequence encoding proteins naturally expressed by fungi. For example said sequence may have been codon optimized which means that codon(s) of the nucleic acid sequence encoding the GnT1 may be changed so that they resemble the codon(s) used by the fungal cell into which the second nucleic acid construct is introduced into. The reason for this is that the frequency with which the different nucleic acid codons are used to encode a particular amino acid differs between different species, e.g. between fungi and mammalian cells. Thus a particular codon may in mammalian cells encode one amino acid while the same codon in a fungus may encode a different amino acid or it may not even be recognized by the fungal cell because e.g. the frequency of nucleic acid sequence comprising this codon may be low and instead a different codon may be recognized by the fungal cell to encode that particular amino acid. Thus to obtain a particular amino acid at a given position it may be advantageous to change the codon of the mammalian sequence to a codon which the fungus use more frequently to. Codon optimization is a well-known phenomenon for a person skilled in the art and may e.g. be performed as described in e.g. Sharp PM et al, 1988, Nucleic Acid Res, 16, pp. 8207-8211 or Ramakrishna L et al, 2004, Journal of Virology, 78 (17), pp. 9174-9189.

Another example of minor changes which may be made to the mammalian sequence includes removal of cryptic introns, e.g. as described in WO 97/49821. Cryptic introns are sequences which are not recognized as an intron by the cell which naturally express the protein, e.g. in the present case a mammalian cell, but which is recognized as an intron by the cell into which the nucleic acid is introduced to produce the protein, e.g. in the present case a fungal cell. Thus by modification of the nucleic acid sequence it is possible to avoid that the fungal cell recognises part(s) of the nucleic acid sequence as an intron if it is not recognised as an intron by e.g. the mammalian cell from which it derives.

For the GnT1 to function optimally, i.e. to add N-acetylglucosamine to Man5GlcNAc2, it is an advantage if said enzyme is located in the Golgi, in particular in the early Golgi, wherein "early Golgi" refers to that part of the Golgi which a protein enters from the ER in contrast to the "late Golgi" which is that part of the Golgi which an protein enters before it is e.g. secreted.

As described above it is an advantage if the GnT1 reside in the Golgi. The human GnT1 is known to comprise a Golgi retention signal within its amino acid sequence; however, it is unknown whether or not this retention signal is recognized in a fungal cell. Thus in a particular embodiment the second nucleic acid construct may further comprise a Golgi retention signal. In particular, the Golgi retention signal may be one normally present in a fungal protein so that said signal is recognised as being a Golgi retention signal by the fungal cell. An example of a suitable Golgi retention signal includes but is not limited to the Kre2 signal, in particular it may be the first 100 amino acids of the *KRE2* gene from *Saccharomyces cerevisiae*, which is shown in SEQ ID NO: 7, as describe e.g. by Vervecken W et al. (2004), Applied and Environmental Microbiology, p.2639-2946. In particular said Golgi retention signal may be located upstream of the 5'-end of the nucleic acid sequence encoding the GnT1, in particular it may located right next to said 5'-end in the open reading frame (ORF) encoding the GnT1, so that the Golgi retention signal is translated together with the GnT1 as one amino acid sequence.

In a particular embodiment the second nucleic acid construct may further comprise a constitutive promoter, such as one of those mentioned above, more particularly the *tpiA* or the *gpd* promoter.

Third nucleic acid construct

The third nucleic construct of the present invention comprises a nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6- α -mannosidase also known as Mannosidase II. Mannosidase II have been classified as belonging to E.C. 3.2.1.114 and they are enzymes capable of catalysing the hydrolysis of the terminal 1,2-linked α -D-mannose residues in the oligosaccharide GlcNAcMan5GlcNAc2 to produce GlcNAcMan3GlcNAc2.

In a particular embodiment of the present invention the Mannosidase II may be of a mammalian origin as only mammalian homologous of this enzyme has been identified so far. However, the present invention is not limited to Mannosidase II of mammalian origin as it is envisioned that Mannosidase II from other organisms may be identified. More particularly, the Mannosidase II may be of human or rat origin. More particularly, the Mannosidase II may comprise the amino acid sequence shown in SEQ ID NO: 9 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 9. In another embodiment it may consist of the amino acid sequence shown in SEQ ID NO: 9 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 9. Furthermore, as described above for GnT1 a synthetic gene encoding a mammalian Mannosidase II may in particular be used, e.g. a synthetic gene which has been modified to codon optimize and/or to remove cryptic intron(s).

The nucleic acid sequence encoding the Mannosidase II may in particular comprise the nucleic acid sequence shown in SEQ ID NO:8, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 8. In another embodiment it may consist of the nucleic acid sequence shown in SEQ ID NO: 8 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 8.

In another embodiment it may comprise the CDS shown in SEQ ID NO:8, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 8. In another embodiment it may consist of the CDS shown in SEQ ID NO: 8 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 8.

Furthermore, as also described for GnT1; the Mannosidase II may in particular be located in Golgi, and those Golgi retention signal described for GnT1 may also be used to localise the Mannosidase II in Golgi. Thus the third nucleic acid construct may further comprise a Golgi retention signal, such as the Kre2 signal described above, in particular it may be the first 100 amino acids of the *KRE2* gene from *Saccharomyces cerevisiae* (shown in SEQ ID NO: 7) as describe e.g. by Vervecken W et al. (2004), Applied and Environmental Microbiology, p.2639-2946. In particular said Golgi retention signal may be located upstream of the 5'-end of the nucleic acid sequence encoding the Mannosidase II, in particular it may be located right next to said 5'-end in the open reading frame (ORF) encoding the Mannosidase II, so that the Golgi retention signal is translated together with the Mannosidase II as one amino acid sequence.

In a particular embodiment the third nucleic acid construct may further comprise a constitutive promoter controlling the expression of mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase, such as one of those mentioned above, more particularly the *tpiA* or the *gpd* promoter.

Fourth nucleic acid construct

The fourth nucleic acid construct of the present invention comprises a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase, which is also known as GnT2 and these terms may be used interchangeably in the context of the present invention. GnT2's have been classified as belonging to E.C. 2.4.1.143, which are enzymes capable of catalyzing the following reaction: UDP-N-acetyl-D-glucosamine + 6-(alpha-D-mannosyl)-beta-D-mannosyl-R -> UDP + 6-(2-[N-acetyl-beta-D-glucosaminyl]-alpha-D-mannosyl)-beta-D-mannosyl-R, wherein R represents the remainder of the N-linked oligosaccharide in the glycoprotein acceptor and UDP is uridine diphosphate. The function of GnT2

with regard to the modification of N-glycans is to catalyze the binding of GlcNAc to GlcNAcMan3GlcNAc2 to produce GlcNAc2Man3GlcNAc2.

In a particular embodiment of the present invention the GnT2 may be of a mammalian origin as only mammalian homologous of this enzyme has been identified so far. However, the present invention is not limited to GnT2 of mammalian origin as it is envisioned that GnT2 from other organisms may be identified. More particularly, the GnT2 may be of human or rat origin. More particularly, the GnT2 may comprise the amino acid sequence shown in SEQ ID NO: 11 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 11. In another embodiment it may consist of the amino acid sequence shown in SEQ ID NO: 11 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 11.

The nucleic acid sequence encoding the GnT2 may in particular comprise the nucleic acid sequence shown in SEQ ID NO: 10, or it may comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 10. In another embodiment it may consist of the nucleic acid sequence shown in SEQ ID NO: 10 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 10.

In another embodiment it may comprise the CDS shown in SEQ ID NO: 10, or it may comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 10. In another embodiment it may consist of the CDS shown in SEQ ID NO: 10 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 10.

Furthermore, as described above for GnT1 a synthetic gene encoding a mammalian GnT2 may in particular be used, such as a synthetic gene which has been modified to codon optimize and/or to remove cryptic intron.

Furthermore, as also described for GnT1; the GnT2 may in particular be located in Golgi, and those Golgi retention signal described for GnT1 may also be used to localise the GnT2 in Golgi.

Thus in a particular embodiment the fourth nucleic acid construct may further comprise a Golgi retention signal, such as the Kre2 signal described above, in particular it may be the first 100 amino acids of the *KRE2* gene from *Saccharomyces cerevisiae* (shown in SEQ ID NO: 7) as describe e.g. by Vervecken W et al. (2004), Applied and Environmental Microbiology, p.2639-2946. In particular said Golgi retention signal may be located upstream of the 5'-end of the nucleic acid sequence encoding the GnT2, in particular it may located right next to

said 5'-end in the open reading frame (ORF) encoding the GnT2, so that the Golgi retention signal is translated together with the GnT2 as one amino acid sequence.

In a particular embodiment the fourth nucleic acid construct may further comprise a constitutive promoter controlling the expression of the alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase, such as one of those mentioned above, more particularly the tpiA or the gpd promoter.

Fifth nucleic acid construct

The fifth nucleic acid construct of the present invention comprises a nucleic acid sequence encoding a beta-1,4-galactosyltransferase (GalT), which are enzymes that have been classified as belonging to E.C. 2.4.1.38 and are capable of catalyzing the following reaction:

UDP-galactose + N-acetyl-beta-D-glucosaminylglycopeptide → UDP + beta-D-galactosyl-1,4-N-acetyl-beta-D-glucosaminylglycopeptide.

The function of GalT with regard to the modification of N-glycans is to catalyse the binding of galactose (Gal) to GlcNAcMan5GlcNAc2 to produce GalGlcNAcMan5GlcNAc2 or to bind Gal to GlcNAc2Man3GlcNAc2 to produce Gal(0-2)GlcNAc2Man3GlcNAc2. As shown in figure 2 the binding of Gal to GlcNAc2Man3GlcNAc2 typically results in the formation of four different N-glycan structures; G0 which denotes GlcNAc2Man3GlcNAc2 (no Gal residue was bound), G1 which denotes GalGlcNAc2Man3GlcNAc2 (one Gal residue was bound) where the Gal residue is bound either to the 1,3 arm or the 1,6 arm of GlcNAc2Man3GlcNAc2 and G2 which denotes Gal2GlcNAc2Man3GlcNAc2 (two Gal residues were bound, one Gal residue bound to the 1,3 arm and one to the 1,6 arm of GlcNAc2Man3GlcNAc2). In a particular embodiment of the present invention at least 10%, such as between 10-50%, 10-40%, 10-30%, 15-50%, 15-40%, 15-35%, 20-40%, 25-35% of the Gal(0-2)GlcNAc2Man3GlcNAc2 structures are Gal(1)GlcNAc2Man3GlcNAc2, wherein Gal(1)GlcNAc2Man3GlcNAc2 includes both of the above mentioned forms with the Gal bound to the 1,3 arm or the 1,6 arm, respectively. Furthermore, the fraction of Gal(2) GlcNAc2Man3GlcNAc2 structures may in particular be less than 15%, such as between 0-15%, 0-10%, 0-5%, 0-3% or 1-5%. This may in particular be relevant if e.g. a mammalian antibody is expressed in a fungal cell of the present invention as approximately 30% of the molecules of a particular antibody expressed in humans has the Gal(1)GlcNAc2Man3GlcNAc2 structure while the other 70% has the Gal(0)GlcNAc2Man3GlcNAc2.

In contrast to the above situation GalT only catalyze binding of Gal to the 1,3 arm of GlcNAcMan5GlcNAc2. Thus for this reaction only two glycans are formed, i.e. G0 and G1, wherein the formed denotes GlcNAcMan5GlcNAc2 (no Gal residue was bound) and

GalGlcNAcMan5GlcNAc2 (one Gal residue was bound) with the Gal residue bound to the 1,3 arm.

In a particular embodiment of the present invention the GalT may be of a mammalian origin as only mammalian homologous of this enzyme has been identified so far. However, the present invention is not limited to GalT of mammalian origin as it is envisioned that GalT from other organisms may be identified. More particularly, the GalT may be of human or rat origin. More particularly, the GalT may comprise the amino acid sequence shown in SEQ ID NO: 6 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 6. In another embodiment it may consist of the amino acid sequence shown in SEQ ID NO: 6 or an amino acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the amino acid sequence shown in SEQ ID NO: 6.

The nucleic acid sequence encoding the GalT may in particular comprise the nucleic acid sequence shown in SEQ ID NO:5, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 5. In another embodiment it may consist of the nucleic acid sequence shown in SEQ ID NO: 5 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence shown in SEQ ID NO: 5.

In another embodiment it may comprise the CDS shown in SEQ ID NO:5, or it may comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 5. In another embodiment it may consist of the CDS shown in SEQ ID NO: 5 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the CDS shown in SEQ ID NO: 5.

In another embodiment it may comprise the nucleic acid sequence encoding the mature peptide shown in SEQ ID NO:5, or it may be comprise a nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence encoding the mature peptide shown in SEQ ID NO: 5. In another embodiment it may consist of the nucleic acid sequence encoding the mature peptide shown in SEQ ID NO: 5 or an nucleic acid sequence having at least 70% identity, such as 75% or 80% or 85% or 90% or 95% identity with the nucleic acid sequence encoding the mature peptide shown in SEQ ID NO: 5.

Furthermore, as described above for GnT1 a synthetic gene encoding a mammalian GalT may in particular be used including a synthetic gene, such as a synthetic gene which has been modified for the purpose of codon optimization and/or to remove cryptic intron(s).

Furthermore, as also described for GnT1; the GalT may in particular be located in Golgi, and those Golgi retention signal described for GnT1 may also be used to localise the

GalT in Golgi, such as the Kre2 signal described above, in particular it may be the first 100 amino acids of the *KRE2* gene from *Saccharomyces cerevisiae* (shown in SEQ ID NO: 7) as described e.g. by Vervecken W et al. (2004), Applied and Environmental Microbiology, p.2639-2946. In particular said Golgi retention signal may be located upstream of the 5'-end of the nucleic acid sequence encoding the GalT, in particular it may be located right next to said 5'-end
5 end in the open reading frame (ORF) encoding the GalT, so that the Golgi retention signal is translated together with the GalT as one amino acid sequence.

In a particular embodiment the fifth nucleic acid construct may further comprise a constitutive promoter, such as one of those mentioned above, more particularly the *tpiA* or the *gpd*
10 promoter.

Alpha-1,6-mannosyltransferase

Some yeast cells have been found to comprise an endogenous gene encoding alpha-1,6-mannosyltransferase, which are enzymes capable of binding a mannose residue by an alpha-1,6-bond to the 1,3 arm of Man9GlcNAc2. If this reaction takes place in a fungal cell of the present invention it may be an advantage to inactivate such endogenous enzymes as the presence of a mannose residue on the 1,3 arm of Man9GlcNAc2 bound by an alpha-1,6-bond is
15 believed to inhibit the action of the alpha-1,2-mannosidase and thereby inhibit the processing of Man9GlcNAc2 to Man5GlcNAc2.

Thus in a particular embodiment of the present invention at least one endogenous gene
20 encoding an alpha-1,6-mannosyltransferase capable of catalyzing the binding of mannose by alpha-1,6-bonds to Man9GlcNAc2, in a fungal cell of the present invention has been modified so that said cell forms fewer alpha-1,6 mannose bonds than a parental fungal cell.

In *Saccharomyces cerevisiae* an alpha-1,6-mannosyltransferase called MNN9 has been identified, the amino acid sequence of which is shown in SEQ ID NO: 12. Thus in a particular embodiment an endogenous gene encoding an alpha-1,6-mannosyltransferase which
25 has at least 40%, e.g. 50% or 60% or 70% or 80% or 90% identity with the sequence of SEQ ID NO: 12 has been modified in a fungal cell of the present invention so that said cell forms fewer alpha-1,6 mannose bonds than a parental fungal cell.

Methods for inactivation of endogenous genes are well known to a person skilled in the art and may be performed as described by e.g. Miller et al., 1985, Molecular and Cell Biology,
30 5, pp. 1714-1721 or Fincham JR, 1989, Microbiological reviews, 53 (1), pp. 148-170. In particular modification of such genes may result in that at least 50%, such as 60% or 70% or 80% or 90% or 95% of the N-glycans in the fungal cell does not contain any mannose residues bound to Man9GlcNAc by an alpha-1,6 bond.

The presence of a mannose residue bound by an alpha-1,6-bond to Man9GlcNAc2 may be measured as described by Jackson P and Gallagher JT, 1997, A laboratory guide to Glycoconjugate analysis, Biomethods, vol. 9, Birkhäuser, Basel, Switzerland.

General

5 The nucleic acid constructs of the present invention may in particular be expression vectors capable of facilitating expression of one of the N-glycosylation enzymes of the present invention (the alpha-1,2-mannosidase, alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase, mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase, alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosamintransferase and beta-1,4-
10 galactosyltransferase) which is encoded by the nucleic acid sequence comprised in nucleic acid construct, when said nucleic acid construct is inserted into a fungal cell.

Expression includes transcription and translation of the nucleic acid sequence into a polypeptide sequence. For secreted polypeptides this usually also includes secretion of said polypeptide.

Expression vectors

15 The choice of expression vector will often depend on the host cell into which it is to be introduced. Examples of a suitable vector include a linear or closed circular plasmid or a virus. The vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromo-
20 some. The vector may contain any means for assuring self-replication. Examples of origin of replications for use in a yeast host cell are the 2 micron origin of replication, the combination of CEN6 and ARS4, and the combination of CEN3 and ARS1. The origin of replication may be one having a mutation which makes it function as temperature-sensitive in the host cell (see,
25 e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75:1433).

Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Vectors which are integrated into the genome of the host cell may contain any nucleic acid sequence enabling integration into the genome; in particular it may contain nucleic
30 acid sequences facilitating integration into the genome by homologous or non-homologous recombination. The vector system may be a single vector, e.g. plasmid or virus, or two or more vectors, e.g. plasmids or virus', which together contain the total nucleotide sequence to be introduced into the genome of the host cell, or a transposon.

The vector may in particular be an expression vector in which the nucleic acid sequence encoding the polypeptide is operably linked to additional segments or control sequences required for transcription of the DNA. The term, "operably linked" indicates that the
35

segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in a promoter and proceeds through the nucleotide sequence encoding the modified polypeptide or the parent polypeptide. Additional segments or control sequences include a promoter, a leader, a polyadenylation sequence, a propeptide sequence, a signal sequence and a transcription terminator. At a minimum the control sequences include a promoter and transcriptional and translational stop signals.

The promoter may be any nucleotide sequence that shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell.

Examples of suitable promoters for use in a filamentous fungal host cell are promoters obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, *Fusarium oxysporum* trypsin-like protease (as described in U.S. Patent No. 4,288,627, which is incorporated herein by reference), and hybrids thereof. Particularly, promoters for use in filamentous fungal host cells are the TAKA amylase, NA2-tpi (a hybrid of the promoters from the genes encoding *Aspergillus niger* neutral (alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and glaA promoters. Further suitable promoters for use in filamentous fungus host cells are the ADH3 promoter (McKnight et al., The EMBO J. 4 (1985), 2093 - 2099), the tpiA promoter or the gpd promoter. The tpiA promoter has been described in McKnight GL et al, 1986, Cell, 46 (1), pp. 143-147 and the gpd promoter has been described in Punt PJ et al, 1990, Gene, 93, pp. 101-109.

Examples of suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., J. Biol. Chem. 255 (1980), 12073 - 12080; Alber and Kawasaki, J. Mol. Appl. Gen. 1 (1982), 419 - 434) or alcohol dehydrogenase genes (Young et al., in Genetic Engineering of Microorganisms for Chemicals (Hollaender et al, eds.), Plenum Press, New York, 1982), or the TPI1 (US 4,599,311) or ADH2-4c (Russell et al., Nature 304 (1983), 652 - 654) promoters.

Further useful promoters are obtained from the *Saccharomyces cerevisiae* enolase (ENO-1) gene, the *Saccharomyces cerevisiae* galactokinase gene (GAL1), the *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase genes (ADH2/GAP), and the *Saccharomyces cerevisiae* 3-phosphoglycerate kinase gene. Other useful promoters for yeast host cells are described by Romanos et al., 1992, Yeast 8:423-488. In a mammalian host cell, useful promoters include viral promoters such as those from Simian Virus 40 (SV40), Rous sarcoma virus (RSV), adenovirus, and bovine papilloma virus (BPV).

The nucleic acid sequence encoding a polypeptide may also, if necessary, be operably connected to a suitable terminator.

The recombinant vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question.

5 The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, or a gene encoding resistance to e.g. antibiotics like ampicillin, kanamycin, chloramphenicol, erythromycin, tetracycline, spectinomycin, neomycin, hygromycin, methotrexate, or resistance to heavy metals, virus or herbicides, or which provides for prototrophy or auxotrophs. A selectable marker for use in a filamentous fungal host
10 cell may be selected from the group including, but not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hygB (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC (sulfate adenylyltransferase), trpC (anthranilate synthase), and glufosinate resistance markers, as well as equivalents from other species. Particularly, for use in an *Aspergillus* cell are
15 the amdS and pyrG markers of *Aspergillus nidulans* or *Aspergillus oryzae* and the bar marker of *Streptomyces hygroscopicus*. Furthermore, selection may be accomplished by co-transformation, e.g., as described in WO 91/17243, where the selectable marker is on a separate vector.

20 The procedures used to ligate the nucleotide sequences coding for the modified polypeptide or the parent polypeptide of the present invention, the promoter and optionally the terminator and/or secretory signal sequence, respectively, or to assemble these sequences by suitable PCR amplification schemes, and to insert them into suitable vectors containing the information necessary for replication or integration, are well known to persons skilled in the art (cf., for instance, Sambrook et al.).

25 An expression vector typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and optionally a repressor gene or various activator genes.

Transformation of a fungal cell

30 Filamentous fungal host cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023, EP 184,438, and Yelton *et al.*, 1984, Proceedings of the National Academy of Sciences USA 81:1470-1474. A suitable method of transforming *Fusarium* species is described by Malardier *et al.*, 1989, Gene 78:147-156 or in co-pending US Serial No. 08/269,449.

Proteins

The present invention also relates to a method of producing a protein with an altered N-glycan structure comprising culturing a fungal cell of the present invention comprising a nucleic acid sequence encoding the protein and recovering the protein. An altered N-glycan structure is to be understood as an N-glycan structure which is different from the N-glycan structure of said protein when expressed naturally. Said protein may be a protein which is naturally expressed by the cell or it may in particular be a protein which is not naturally expressed by the cell, such as a protein of mammalian origin.

In a particular embodiment the protein may be of mammalian origin. In the context of the present invention the term "mammalian origin" is intended to encompass amino acid sequences, e.g. proteins, or nucleic acid sequences which have been isolated from a mammalian cell. However, it is also intended to encompass such sequences which are identical to a sequence isolated from a mammalian cell but which has been chemically synthesised or expressed by a non-mammalian cell using e.g. well-known methods of gene technology. Furthermore, it is also intended to encompass sequences which comprise minor differences as compared to the mammalian counterpart such as codon optimization, i.e. changes in the nucleic acid sequences which do not affect the amino acid sequence and/or the removal of cryptic introns.

In a particular embodiment the protein may be a monoclonal antibody (mAb), e.g. a human mAb or antibody fragments, such as Fab, single chain antibodies, diabodies or triabodies. The heavy chain of a human IgG antibody contains only one N-linked glycan in the constant domain, known to be important for the effect and/or function of antibodies. Thus in a particular embodiment a human mAb, wherein one or both of heavy chains of said mAb comprises a Man5GlcNAc2, or a GlcNAcMan5GlcNAc2, or GalGlcNAcMan5GlcNAc2, or a GlcNAcMan3GlcNAc2, or a GlcNAc2Man3GlcNAc2a or a Gal(0-2)GlcNAc2Man3GlcNAc2 N-glycan structure may be produced in a fungal cell of the present invention. If it comprises a Gal(0-2)GlcNAc2Man3GlcNAc2 N-glycan structure approximately 30% of the heavy chains may comprise a Gal1GlcNAc2Man3GlcNAc2 N-glycan structure while approximately 70% of the heavy chains may comprise a Gal0GlcNAc2Man3GlcNAc2 N-glycan structure.

MATERIALS AND METHODS

EXAMPLES

Example 1

Trimming of high mannose glycans to Man5GlcNac2.

An *Aspergillus oryzae* Mnn9 homologue is identified. The mnn9 gene product is involved in alfa-1,6 mannose bond formation. The mnn9 gene is disrupted from HowB425 (WO 98/12300) by standard gene disruption methods applying the pyrG gene as selective marker. Disruptants are identified by Southern blot analysis. A pyrG mutant is selected by first selecting for growth on minimal plates containing uridine and 5-flouorotic acid (5-FOA), and secondly demonstrate that the selected strains can be complemented by transformation with the *Aspergillus oryzae* pyrG gene. The selected strain is termed *Aspergillus oryzae* mnn9-A.

The *A. nidulans* mannosidase 1C gene (SEQ ID NO: 1) is cloned into an *Aspergillus oryzae* expression vector down-stream of the gapdh (glycer aldehyde phosphate dehydrogenase) promoter. The mannosidase is C-terminally fused to a sequence encoding the amino-acid sequence His-Asp-Glu-Leu. These four amino acids constitute an ER retention signal that will localize the mannosidase to the ER. The expression vector further more harbors the *A. oryzae* pyrG gene as a selective marker flanked by an upstream sequence for the *A. oryzae* alkaline protease *alp* (SEQ ID NO: 3 of WO 98/12300), the 5' end of the gapdh promoter upstream the pyrG gene and flanked by the mannosidase expression cassette and the downstream sequence for the alkaline protease downstream the pyrG gene.

The DNA constructs are linarized so that the alkaline protease flanking sequences are at the ends of the plasmid are transformed into *Aspergillus oryzae* mnn9-A. Transformants are analyzed by sourthern blot analysis, and a transformant of each mannosidase construct where the pyrG gene and the mannosidase expression cassette replaced the alkaline protease is selected.

The effect of the transformation on glycosylations is analysed using the *Thermomyces lanuginosus* lipase (the nucleic-and amino acid sequence of which is shown in US 5869438) as a reporter protein as this contains one N-glycosylation site. This lipase gene is transformed into the selected strains using amdS based expression plasmids and the TAKA promoter as described in patent WO 91/17243.

It is confirmed that the N-linked structure of the lipase consist of Man5GlcNac2.

Having confirmed the desired glycosylation pattern, pyrG mutants are selected by first selecting for growth on minimal plates containing uridine and 5-FOA, and secondly demonstrate that the selected strain can be complemented by transformation with the *Aspergillus oryzae* pyrG gene. The selected strain is termed *Aspergillus oryzae* AnMan (for the *Aspergillus nidulans* mannosidase strain).

Example 2Adding N-acetyl glucosamine to Man5GlcNac2.

The human GnT 1 (alpha-1,3-mannosyl-glycoprotein beta-1, 2-N-acetylglucosaminyltransferase) gene is expressed in *Aspergillus oryzae* in the following way:

5 A synthetic gene codon optimized for *Aspergillus oryzae* encoding the human GnT 1 gene (the amino acid sequence shown in SEQ ID NO: 2) is designed and cloned into an *Aspergillus oryzae* expression vector downstream of the tpi (triose phosphate isomerase) promoter. The mannosidase is N-terminal fused to an alpha-1,2-mannosyl-transferase (Kre2p) homologue Golgi retention signal. The expression vector further more harbors the *A. oryzae* pyrG gene as a selective marker flanked by an upstream sequence for the *A. oryzae* Neutral metalloprotease I (SEQ ID NO: 2 of WO 98/12300), the 5' end of the tpi promoter upstream the pyrG gene and flanked by the GnT 1 expression cassette and the downstream sequence for the neutral metallo protease downstream the pyrG gene.

10 The DNA constructs are linearized so that the neutral metallo protease flanking sequences are at the ends of the plasmid before they are transformed into *Aspergillus oryzae* AnMan. Transformants are analyzed by Southern blot analysis, and a transformant where the pyrG gene and the GnT 1 expression cassette has replaced the neutral metallo protease is selected.

20 The effect of the transformation on glycosylations is analysed the same way as in example 1 using the *Thermomyces lanuginosus* lipase. It is confirmed that the N linked structure of the lipase consist of GlcNacMan5GlcNac2.

Having confirmed the desired glycosylation pattern, pyrG mutants are selected as in example 1. The selected strain is termed *Aspergillus oryzae* GnT1.

Example 325 Trimming GlcNacMan5GlcNac2 to GlcNacMan3glcNac2

The human mannosidase II (Mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase), gene is expressed in *Aspergillus oryzae* in the following way:

30 A synthetic gene codon optimized for *Aspergillus oryzae* encoding the human mannosidase II gene (the amino acid sequence is shown in SEQ ID NO: 9) is designed and cloned into an *Aspergillus oryzae* expression vector downstream of the tpi (triose phosphate isomerase) promoter. The mannosidase is N-terminal fused to an alpha-1,2 mannosyl-transferase (Kre2p) homologue Golgi retention signal. The expression vector furthermore harbours the *A. oryzae* pyrG gene as a selective marker flanked by an upstream sequence for the *A. oryzae* pepE gene (SEQ ID NO: 3 in WO 97/22705), the 5' end of the tpi promoter upstream the pyrG

gene and flanked by the mannosidase II expression cassette and the downstream sequence for the pepE gene downstream the pyrG gene.

The DNA constructs are linearized so that pepE flanking sequences are at the ends of the plasmid and are transformed into *Aspergillus oryzae* GnT1. Transformants are analyzed by Southern blot analysis, and a transformant where the pyrG gene and the mannosidase II expression cassette replaced the pepE gene is selected.

The effect of the transformation on glycosylations is analysed the same way as in example 1 using the *Thermomyces lanuginosus* lipase. It is confirmed that the N-linked structure of the lipase consist of GlcNAcMan3GlcNAc2.

Having confirmed the desired glycosylation pattern, pyrG mutants are selected as in example 1. The selected strain is termed *Aspergillus oryzae* Man2.

Example 4

Adding N acetyl glucosamine to GlcNAcMan3GlcNAc2.

The human GnT 2 (Alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase), gene is expressed in *Aspergillus oryzae* in the following way:

A synthetic gene codon optimized for *Aspergillus oryzae* encoding the human GnT 2 gene (the amino acid sequence is shown in SEQ ID NO: 11) is designed and cloned into an *Aspergillus oryzae* expression vector downstream of the tpi (triose phosphate isomerase promoter). The mannosidase is N-terminal fused to an alpha-1,2-mannosyltransferase (Kre2p) homologue Golgi retention signal. The expression vector further more harbors the *A. oryzae* pyrG gene as a selective marker flanked by an upstream sequence for the *A. oryzae* pepC gene ((SEQ ID NO: 1 in WO 97/22705)), the 5' end of the tpi promoter upstream the pyrG gene and flanked by the GnT 2 expression cassette and the downstream sequence for the pepC gene downstream the pyrG gene.

The DNA constructs are linearized so that pepC flanking sequences are at the ends of the plasmid are transformed into *Aspergillus oryzae* GnT1. Transformants are analyzed by Southern blot analysis, and a transformant where the pyrG gene and the GnT2 expression cassette replaced the pepC gene is selected.

The effect of the transformation on glycosylations is analysed the same way as in example 1 using the *Thermomyces lanuginosus* lipase. It is confirmed that the N-linked structure of the lipase consist of primarily GlcNAc2Man3GlcNAc2.

Having confirmed the desired glycosylation pattern, pyrG mutants are selected as in example 1. The selected strain is termed *Aspergillus oryzae* GnT2.

Example 5Adding Galactose to GlcNac2Man3GlcNac2.

The human galT (beta-1,4-galactosyltransferase) gene is expressed in *Aspergillus oryzae* in the following way:

5 A synthetic gene codon optimized for *Aspergillus oryzae* encoding the human galT gene (the amino acid sequence is shown in SEQ ID NO: 6) is designed and cloned into an *Aspergillus oryzae* expression vector downstream of the tpi (triose phosphate isomerase) promoter. The mannosidase is N terminal fused to the kexB Golgi retention signal. The expression vector further more harbors the *A. oryzae* pyrG gene as a selective marker flanked by an up-
10 stream sequence for the *A. oryzae* AMG gene, the 5' end of the tpi promoter upstream the pyrG gene and flanked by the galT expression cassette and the downstream sequence for the AMG gene downstream the pyrG gene.

The DNA constructs are linearized so that AMG flanking sequences are at the ends of the plasmid are transformed into *Aspergillus oryzae* GtT1. Transformants are analyzed by
15 Southern blot analysis, and a transformant where the pyrG gene and the galT expression cassette replaced the AMG gene is selected.

The effect of the transformation on glycosylations is analysed the same way as in example 1 using the *Thermomyces lanuginosus* lipase. It is confirmed that the N-linked structure of the lipase consist of a significant amount of GalGlcNAc2Man3GlcNAc2 and some
20 Gal2GlcNAc2Man3GlcNAc2

Having confirmed the desired glycosylation pattern, pyrG mutants are selected as in example 1. The selected strain is termed *Aspergillus oryzae* GalT.

CLAIMS

1. Method of altering the structure of N-glycans in a fungus comprising insertion into a fungal cell of a nucleic acid construct or a combination of nucleic acid constructs selected from the group consisting of:
- 5 a) a first nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungi
- b) a second nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
- 10 c) a third nucleic acid construct comprising a nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase
- d) a fourth nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
- e) a fifth nucleic acid construct comprising a nucleic acid sequence encoding a beta-1,4-galactosyltransferase.
- 15 2. A method according to claim 1, wherein the first nucleic acid construct further comprises a nucleic acid sequence encoding an ER retention signal.
3. A method according to claim 2, wherein the nucleic acid sequence encoding the ER retention signal is located at the 3'-end of the nucleic acid sequence encoding the alpha-1,2-mannosidase.
- 20 4. A method according to any of claims 2-3, wherein the ER retention signal has the amino acid sequence HDEL.
5. A method according to any of the preceding claims, wherein the first nucleic acid construct further comprises a constitutive promoter, e.g. tpi or gpd promoter.
- 25 6. A method according to any of the preceding claims, wherein the nucleic acid sequence encoding the alpha-1,2-mannosidase comprises the nucleic acid sequence shown in

SEQ ID NO: 1 or a sequence which has at least 70% identity with the sequence of SEQ ID NO: 1.

7. A method according to any of the preceding claims, wherein the nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyl-transferase is of mammalian origin.

8. A method according to any of the preceding claims, wherein the second nucleic acid construct further comprises a constitutive promoter, e.g. the tpi or gpd promoter.

9. A method according to any of the preceding claims, wherein the second nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase.

10. A method according to any of the preceding claims, wherein the nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase is of mammalian origin.

11. A method according to any of the preceding claims, wherein the third nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase.

12. A method according to any of the preceding claims, wherein the nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyl-transferase is of mammalian origin.

13. A method according to any of the preceding claims, wherein the fourth nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyl-transferase.

14. A method according to any of the preceding claims, wherein the third, fourth or fifth nucleic acid construct further comprises a constitutive promoter, e.g. the tpi or gpd promoter.

15. A method according to any of the preceding claims, wherein the nucleic acid sequence encoding a beta-1,4-galactosyltransferase is of mammalian origin.

16. A method according to any of the preceding claims, wherein the fifth nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the beta-1,4-galactosyltransferase.

17. A method according to any of the preceding claims, wherein the fungal cell is a filamentous fungal cell, such as a strain of *Aspergillus*, e.g. *A.oryzae*.

18. A method according to any of the preceding claims, wherein at least one endogenous gene encoding an alpha-1,6-mannosyltransferase capable of catalyzing the binding of mannose by alpha-1,6-bonds to the 1,3 arm of Man9GlcNAc2 has been modified so that fewer mannose residues are bound by alpha-1,6-bonds to the 1,3 arm of Man9GlcNAc2 structure of proteins than in a parental cell.

19. A method according to claim 18, wherein the endogenous gene encodes an enzyme having at least 40% identity with the sequence of SEQ ID NO: 12.

20. A method of preparing a heterokaryon fungus comprising an altered structure of N-glycans comprising fusion of at least two genetically different fungal cells under conditions suitable for formation of a heterokaryon, wherein at least one of the genetically different cells have been altered according to any of claims 1-19.

21. A fungal cell comprising a nucleic acid construct or a combination of nucleic acid constructs selected from the group consisting of:

a) a first nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,2-mannosidase obtained from a filamentous fungus

b) a second nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetyl-glucosaminyltransferase

c) a third nucleic acid construct comprising a nucleic acid sequence encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase

- d) a fourth nucleic acid construct comprising a nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosamintransferase
- e) a fifth nucleic acid construct comprising a nucleic acid sequence encoding a beta-1,4-galactosyltransferase,

5 wherein said cell has an altered N-glycan structure as compared to a parental fungal cell.

22. A fungal cell according to claim 21, wherein the first nucleic acid construct further comprises a nucleic acid sequence encoding an ER retention signal.

23. A fungal cell according to any of claims 21-22, wherein the nucleic acid sequence encoding the ER retention signal is located at the 3'-end of the nucleic acid sequence
10 encoding the alpha-1,2-mannosidase.

24. A fungal cell according to any of claims 22-23, wherein the ER retention signal has the amino acid sequence HDEL.

25. A fungal cell according to any of claims 21-24, wherein the first nucleic acid construct further comprises a constitutive promoter, e.g. tpi or gpd promoter.

15 26. A fungal cell according to any of claims 21-25, wherein the alpha-1,2-mannosidase is obtained from a strain of *Aspergillus*, e.g. *A.oryzae* or *A.nidulans*.

27. A fungal cell according to any of claims 21-26, wherein the nucleic acid sequence encoding the alpha-1,2-mannosidase comprises the nucleic acid sequence shown in SEQ ID NO: 1 or a sequence which has at least 70% identity with the sequence of SEQ ID NO: 1.

20 28. A fungal cell according to claim 21-27, wherein the second nucleic acid construct further comprises a constitutive promoter, e.g. the tpi or gpd promoter.

29. A fungal cell according to any of claims 21-28, wherein the nucleic acid sequence encoding a alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase is of mammalian origin.

25 30. A fungal cell according to any of claims 21-29, wherein the second nucleic acid construct further comprises a nucleic acid sequence encoding an Golgi retention signal, such

as such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase.

31. A fungal cell according to any of claims 21-30, wherein the nucleic acid sequence
5 encoding a mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase is of mammalian origin.

32. A fungal cell according to any of claims 21-31, wherein the third nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase.

10 33. A fungal cell according to any of claims 21-32, wherein the nucleic acid sequence encoding an alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosamintransferase is of mammalian origin.

34. A fungal cell according to any of claims 21-33, wherein the fourth nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as
15 such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence encoding the alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosamintransferase.

35. A fungal cell according to any of claims 21-34, wherein the fifth nucleic acid construct further comprises a constitutive promoter, e.g. the tpi or gpd promoter.

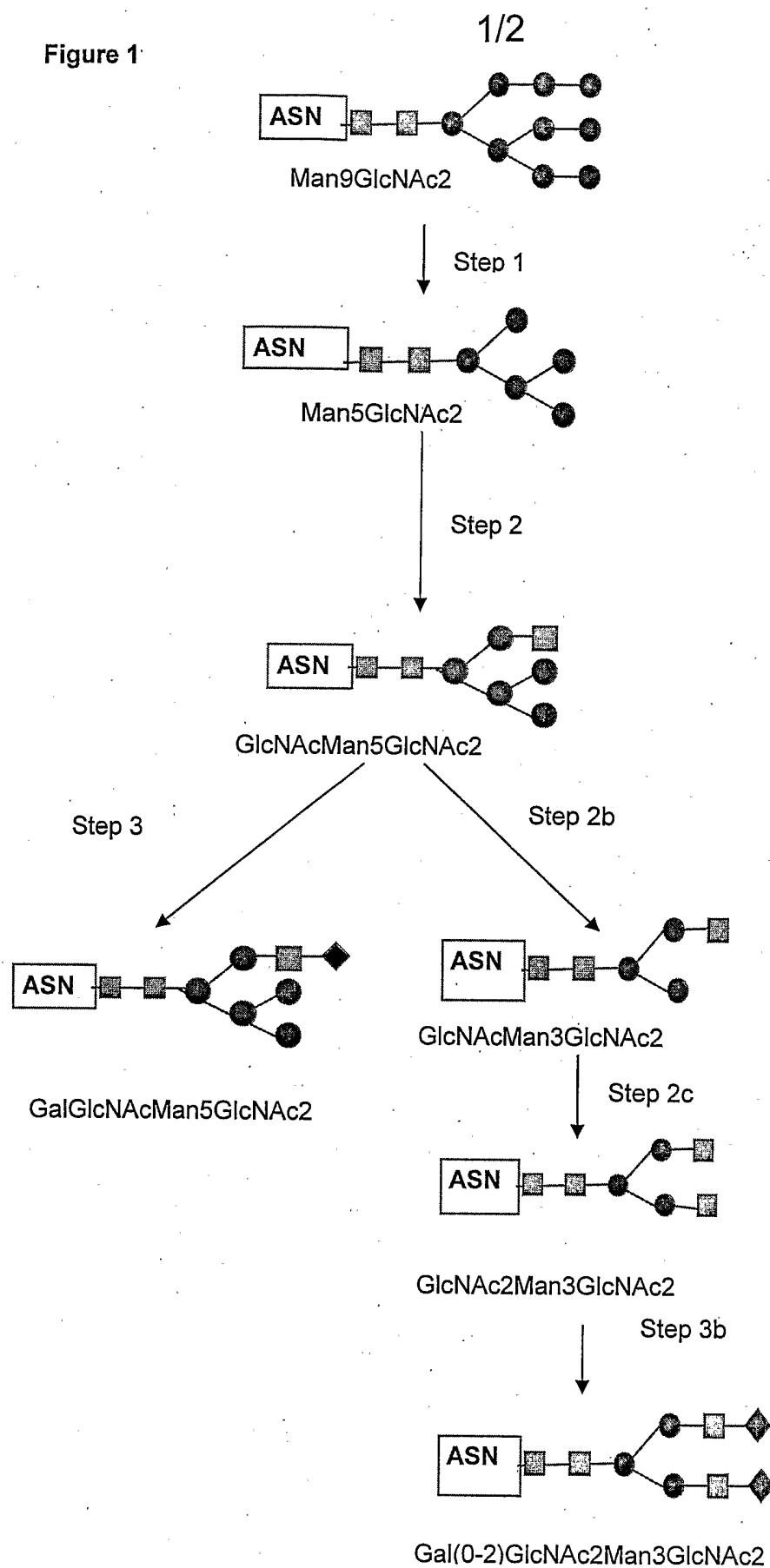
20 36. A fungal cell according to any of claims 21-35, wherein the nucleic acid sequence encoding a beta-1,4-galactosyltransferase is of mammalian origin.

37. A fungal cell according to any of claims 21-36, wherein the fifth nucleic acid construct further comprises a nucleic acid sequence encoding a Golgi retention signal, such as such as the amino acids 1-100 of SEQ ID NO: 7 located at the 5'-end of the nucleic acid sequence
25 encoding the beta-1,4-galactosyltransferase.

38. A fungal cell according to any of claims 21-37, wherein at least one endogenous gene of the fungal cell has been modified so that fewer mannose residues are bound by alpha-1,6-bonds to the 1,3 arm of the Man9GlcNAc2 structure of proteins than in a parental cell.

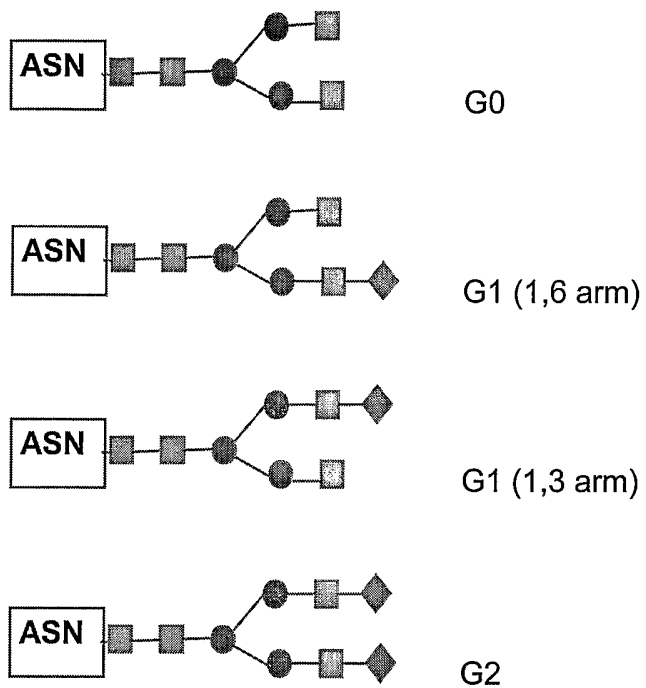
39. A fungal cell according to claim 38, wherein the endogenous gene encodes an enzyme having at least 40% identity with the sequence of SEQ ID NO: 12.
40. A fungal cell according to any of claims 21-39, wherein the fungal cell is a filamentous fungal cell, such as a strain of *Aspergillus*, e.g. *A.oryzae*.
- 5 41. A fungal cell according to any of claims 21-40, wherein the fungal cell is a heterokaryon.
42. A fungal cell according to claim 41, wherein the heterokaryon is a filamentous heterokaryon fungus.
43. A fungal cell according to claim 41, wherein the heterokaryon is a yeast heterokaryon.
- 10 44. A method of producing a protein with an altered glycan structure comprising culturing a fungal cell of any of claims 21-43, comprising a nucleic acid sequence encoding the protein and recovering the protein.
45. A method according to claim 44, wherein the protein is a monoclonal antibody.

Figure 1



2/2

Figure 2



SEQUENCE LISTING

<110> Novozymes A/S

<120> Modification of N-glycans

<130> 10628

<160> 12

<170> PatentIn version 3.2

<210> 1

<211> 2032

<212> DNA

<213> aspergillus nidulans

<220>

<221> CDS

<222> (36)..(1796)

<400> 1

atcgatgtct gctgcataaa ggcagacgga ggaag atg ccg aga cgg tgg tcc 53
 Met Pro Arg Arg Trp Ser
 1 5

tcc ctc atc agc atc aca gcc atc ttc ttg gtc ctc ttc ttc ctc ctt 101
 Ser Leu Ile Ser Ile Thr Ala Ile Phe Leu Val Leu Phe Phe Leu Leu
 10 15 20

cat agg aat aca gac aca cca cgc gcc gcc aat agg gct aca aac ggc 149
 His Arg Asn Thr Asp Thr Pro Arg Ala Ala Asn Arg Ala Thr Asn Gly
 25 30 35

cct gcc aac ggc ttt gct agg cag caa agc ata tgt cca tca aca ccc 197
 Pro Ala Asn Gly Phe Ala Arg Gln Gln Ser Ile Cys Pro Ser Thr Pro
 40 45 50

cct cag cct cca tat aac cga acc agc acg gga ggg ttc aac tgg ggt 245
 Pro Gln Pro Pro Tyr Asn Arg Thr Ser Thr Gly Gly Phe Asn Trp Gly
 55 60 65 70

gaa atc cca gtc aga tac cct gta tcc gac ttc atc ccg ctg tca acc 293
 Glu Ile Pro Val Arg Tyr Pro Val Ser Asp Phe Ile Pro Leu Ser Thr
 75 80 85

aac tct cct gca aca ctt ccg cgc atc caa cgc tct tcc ttc cca ctt 341
 Asn Ser Pro Ala Thr Leu Pro Arg Ile Gln Arg Ser Ser Phe Pro Leu
 90 95 100

caa tcc tca atc act aaa tcc cgc cag gca gca gtc aaa ggt gcc ttt 389
 Gln Ser Ser Ile Thr Lys Ser Arg Gln Ala Ala Val Lys Gly Ala Phe
 105 110 115

cag cgc gca tgg acc tcc tac aca acc cac gcc tgg aag gcg gac gag 437
 Gln Arg Ala Trp Thr Ser Tyr Thr Thr His Ala Trp Lys Ala Asp Glu
 120 125 130

gta cgg ccc atc acg gcc gga tct cga aac aac ttt ggc gga tgg gga Val Arg Pro Ile Thr Ala Gly Ser Arg Asn Asn Phe Gly Gly Trp Gly 135 140 145 150	485
gcg acc cta gtc gac aat ctc gac aca ctg cta atc atg ggg ctg gac Ala Thr Leu Val Asp Asn Leu Asp Thr Leu Leu Ile Met Gly Leu Asp 155 160 165	533
gag gag ttc gca gcg gca gtc gac gcg ctc gca gat ata gaa ttc agc Glu Glu Phe Ala Ala Ala Val Asp Ala Leu Ala Asp Ile Glu Phe Ser 170 175 180	581
ccg cac tcg tcc cca tcc tcc tcc cag agc aca atc aac ata ttc gaa Pro His Ser Ser Pro Ser Ser Gln Ser Thr Ile Asn Ile Phe Glu 185 190 195	629
acg aca atc cgg tat ctg ggc ggc ttg ctc gcg gcg tat gat ctc act Thr Thr Ile Arg Tyr Leu Gly Gly Leu Leu Ala Ala Tyr Asp Leu Thr 200 205 210	677
ggc tgt cga gag act cgg ctg ctg gac aaa gca atc cag ctt ggg gag Gly Cys Arg Glu Thr Arg Leu Leu Asp Lys Ala Ile Gln Leu Gly Glu 215 220 225 230	725
atg atc tac acc tcc ttc gac aca gag aac cgc atg ccc gta cca cgg Met Ile Tyr Thr Ser Phe Asp Thr Glu Asn Arg Met Pro Val Pro Arg 235 240 245	773
tgg aat ctg cac aaa gca ggc aac gga gag cct cag cgc gcg gca gtg Trp Asn Leu His Lys Ala Gly Asn Gly Glu Pro Gln Arg Ala Ala Val 250 255 260	821
cag ggc gtg ctc gct gaa ctc gcc agc agc agt ctc gag ttc acg cgg Gln Gly Val Leu Ala Glu Leu Ala Ser Ser Ser Leu Glu Phe Thr Arg 265 270 275	869
ctg tcg cag ctg acg ggg gat atg cgg tat ttc gat gcg gca tcc cgc Leu Ser Gln Leu Thr Gly Asp Met Arg Tyr Phe Asp Ala Ala Ser Arg 280 285 290	917
att acc gat ctg ctt gac tcc caa gcc ggc cat acc cgg atc ccg ggg Ile Thr Asp Leu Leu Asp Ser Gln Ala Gly His Thr Arg Ile Pro Gly 295 300 305 310	965
ttg tgg cca gtc agc gtg aac ctg cag aaa ggc gat ctg acc cgt ggg Leu Trp Pro Val Ser Val Asn Leu Gln Lys Gly Asp Leu Thr Arg Gly 315 320 325	1013
tcg aca ttc agt ttt ggc ggg atg gcc gat agc gcc tac gag tat ctc Ser Thr Phe Ser Phe Gly Gly Met Ala Asp Ser Ala Tyr Glu Tyr Leu 330 335 340	1061
ggc aag acg tat cgg ctc ctc ggt ggt gtg ggg aaa ggg cca cag tac Gly Lys Thr Tyr Arg Leu Leu Gly Gly Val Gly Lys Gly Pro Gln Tyr 345 350 355	1109
gag cgt ctg gcg cga aac gca cta gat gcc ggg att cga cat ctc ctc Glu Arg Leu Ala Arg Asn Ala Leu Asp Ala Gly Ile Arg His Leu Leu 360 365 370	1157

ttc cga ccg atg acg cct gat cat gca gat atc ctc cta ccc ggg gtc Phe Arg Pro Met Thr Pro Asp His Ala Asp Ile Leu Leu Pro Gly Val 375 380 385 390	1205
gcg cac gca acc agc tct tcc gtg gga ctc gag ccc cgg aca gag cat Ala His Ala Thr Ser Ser Val Gly Leu Glu Pro Arg Thr Glu His 395 400 405	1253
ctc gcc tgt ttt gtg ggt ggg atg tac gcg ctc gcc ggg aag ctt ttc Leu Ala Cys Phe Val Gly Gly Met Tyr Ala Leu Ala Gly Lys Leu Phe 410 415 420	1301
tca aac cag acg tac ctc gac acc ggc cgg aag ctg aca gac ggt tgt Ser Asn Gln Thr Tyr Leu Asp Thr Gly Arg Lys Leu Thr Asp Gly Cys 425 430 435	1349
atc tgg tac tac gat aat tca ccg cta ggt atc atg ccg gag atg ttc Ile Trp Tyr Tyr Asp Asn Ser Pro Leu Gly Ile Met Pro Glu Met Phe 440 445 450	1397
acc gtg ccg gct tgt ccg tca gtg gct gaa tgt cct tgg gac gaa aca Thr Val Pro Ala Cys Pro Ser Val Ala Glu Cys Pro Trp Asp Glu Thr 455 460 465 470	1445
agg ggt ggt atc tac acc tac gtg cgt gat ggg cac tac ttt ctg cgt Arg Gly Gly Ile Tyr Thr Tyr Val Arg Asp Gly His Tyr Phe Leu Arg 475 480 485	1493
cct gag gca atg gag agt atc ttc tat atg tgg cgc att aca ggg gac Pro Glu Ala Met Glu Ser Ile Phe Tyr Met Trp Arg Ile Thr Gly Asp 490 495 500	1541
gaa aag tac cgc gag gct gca tgg aga atg ttc acg gct atc gaa gcg Glu Lys Tyr Arg Glu Ala Ala Trp Arg Met Phe Thr Ala Ile Glu Ala 505 510 515	1589
gtt aca aag acg gag ttt ggg aat gcg gcg gtg cgg gat gtt atg gtt Val Thr Lys Thr Glu Phe Gly Asn Ala Ala Val Arg Asp Val Met Val 520 525 530	1637
gag gaa gga aat gta aag aga gaa gat agc atg gag agt ttc tgg atg Glu Glu Gly Asn Val Lys Arg Glu Asp Ser Met Glu Ser Phe Trp Met 535 540 545 550	1685
gca gag acg ttg aag tat ctg tat ctg ata ttt ggg gag acc gat ttg Ala Glu Thr Leu Lys Tyr Leu Tyr Leu Ile Phe Gly Glu Thr Asp Leu 555 560 565	1733
gtc agc ttg gac gac tgg gtg ttc aat acg gag gcg cac cct ttg agg Val Ser Leu Asp Trp Val Phe Asn Thr Glu Ala His Pro Leu Arg 570 575 580	1781
ggt gca ggg agt tga cattgtattc acacatcggt atagacaaat tatagagtag Gly Ala Gly Ser 585	1836
acgttcaaaa cggccaaaac tgaatggata gactccatat gcattgaata tacaatgtat	1896

tcgctgcaaa gcatggataa aataaagatg tacaaagtgt ctttggtgtc gctttgaaag 1956
 tggatatatca tcccatcata aggtggcagt gtaaccaacc ctctatatca cctacataga 2016
 cagctgatag accggc 2032

<210> 2
 <211> 586
 <212> PRT
 <213> aspergillus nidulans

<400> 2

Met Pro Arg Arg Trp Ser Ser Leu Ile Ser Ile Thr Ala Ile Phe Leu
 1 5 10 15

Val Leu Phe Phe Leu Leu His Arg Asn Thr Asp Thr Pro Arg Ala Ala
 20 25 30

Asn Arg Ala Thr Asn Gly Pro Ala Asn Gly Phe Ala Arg Gln Gln Ser
 35 40 45

Ile Cys Pro Ser Thr Pro Pro Gln Pro Pro Tyr Asn Arg Thr Ser Thr
 50 55 60

Gly Gly Phe Asn Trp Gly Glu Ile Pro Val Arg Tyr Pro Val Ser Asp
 65 70 75 80

Phe Ile Pro Leu Ser Thr Asn Ser Pro Ala Thr Leu Pro Arg Ile Gln
 85 90 95

Arg Ser Ser Phe Pro Leu Gln Ser Ser Ile Thr Lys Ser Arg Gln Ala
 100 105 110

Ala Val Lys Gly Ala Phe Gln Arg Ala Trp Thr Ser Tyr Thr Thr His
 115 120 125

Ala Trp Lys Ala Asp Glu Val Arg Pro Ile Thr Ala Gly Ser Arg Asn
 130 135 140

Asn Phe Gly Gly Trp Gly Ala Thr Leu Val Asp Asn Leu Asp Thr Leu
 145 150 155 160

Leu Ile Met Gly Leu Asp Glu Glu Phe Ala Ala Ala Val Asp Ala Leu
 165 170 175

Ala Asp Ile Glu Phe Ser Pro His Ser Ser Pro Ser Ser Ser Gln Ser

180	185	190
Thr Ile Asn Ile Phe Glu Thr Thr Ile Arg Tyr Leu Gly Gly Leu Leu		
195	200	205
Ala Ala Tyr Asp Leu Thr Gly Cys Arg Glu Thr Arg Leu Leu Asp Lys		
210	215	220
Ala Ile Gln Leu Gly Glu Met Ile Tyr Thr Ser Phe Asp Thr Glu Asn		
225	230	235
Arg Met Pro Val Pro Arg Trp Asn Leu His Lys Ala Gly Asn Gly Glu		
	245	250
Pro Gln Arg Ala Ala Val Gln Gly Val Leu Ala Glu Leu Ala Ser Ser		
	260	265
Ser Leu Glu Phe Thr Arg Leu Ser Gln Leu Thr Gly Asp Met Arg Tyr		
	275	280
Phe Asp Ala Ala Ser Arg Ile Thr Asp Leu Leu Asp Ser Gln Ala Gly		
	290	295
His Thr Arg Ile Pro Gly Leu Trp Pro Val Ser Val Asn Leu Gln Lys		
305	310	315
Gly Asp Leu Thr Arg Gly Ser Thr Phe Ser Phe Gly Gly Met Ala Asp		
	325	330
Ser Ala Tyr Glu Tyr Leu Gly Lys Thr Tyr Arg Leu Leu Gly Gly Val		
	340	345
Gly Lys Gly Pro Gln Tyr Glu Arg Leu Ala Arg Asn Ala Leu Asp Ala		
	355	360
Gly Ile Arg His Leu Leu Phe Arg Pro Met Thr Pro Asp His Ala Asp		
	370	375
Ile Leu Leu Pro Gly Val Ala His Ala Thr Ser Ser Ser Val Gly Leu		
385	390	395
Glu Pro Arg Thr Glu His Leu Ala Cys Phe Val Gly Gly Met Tyr Ala		
	405	410
		415

Leu Ala Gly Lys Leu Phe Ser Asn Gln Thr Tyr Leu Asp Thr Gly Arg
 420 425 430

Lys Leu Thr Asp Gly Cys Ile Trp Tyr Tyr Asp Asn Ser Pro Leu Gly
 435 440 445

Ile Met Pro Glu Met Phe Thr Val Pro Ala Cys Pro Ser Val Ala Glu
 450 455 460

Cys Pro Trp Asp Glu Thr Arg Gly Gly Ile Tyr Thr Tyr Val Arg Asp
 465 470 475 480

Gly His Tyr Phe Leu Arg Pro Glu Ala Met Glu Ser Ile Phe Tyr Met
 485 490 495

Trp Arg Ile Thr Gly Asp Glu Lys Tyr Arg Glu Ala Ala Trp Arg Met
 500 505 510

Phe Thr Ala Ile Glu Ala Val Thr Lys Thr Glu Phe Gly Asn Ala Ala
 515 520 525

Val Arg Asp Val Met Val Glu Glu Gly Asn Val Lys Arg Glu Asp Ser
 530 535 540

Met Glu Ser Phe Trp Met Ala Glu Thr Leu Lys Tyr Leu Tyr Leu Ile
 545 550 555 560

Phe Gly Glu Thr Asp Leu Val Ser Leu Asp Asp Trp Val Phe Asn Thr
 565 570 575

Glu Ala His Pro Leu Arg Gly Ala Gly Ser
 580 585

<210> 3
 <211> 4705
 <212> DNA
 <213> homo sapiens

<220>
 <221> mRNA
 <222> (2056)..(4592)

<220>
 <221> CDS
 <222> (2182)..(3519)

<400> 3

tgatctctgt aatatagatt agttctctatt ttcggcagtt ttatgtcaat ggaatcacag	60
tgtatgtgct cttttttgtt tttcccaaaa agaccatat tgtgtaatta tttggagatt	120
gtgttggtga tgtcagtggt ttattcagtt ttattgctga atattattcc attctatgag	180
taaaccacaa tttgtttatc tgtatttccc tgttgatggg cctttgggtt atttccagtt	240
ttgagtgatt atgaataaag ttaccatcaa cactcctgtg taggtttttg tagggacatg	300
gctgaaagta gaaagaaatg gggcatgcag gagagactgg gcttggttat aacagcctgt	360
tatccctggg gaactagctc actcaccaga ggacgttagg aaggatctgc ccctgcccc	420
cttgtgatcc aaacacctac cactagacct cacctccggc agcaccacca cacaggggat	480
caaatcccaa catgtgtttc agagaacaaa ccatattcaa gctatagcat agctacgtaa	540
ttattataac actattaata agaccatctt tgccttggga attttaagta gcacctttat	600
tacatattag atttttttgt gtaatgagtc tgtttctgga gcttaaattc tatttctactg	660
gatggctctgt ttatttatgt accagtacca cgcagttgta attatggagg ttttgtagta	720
tgccatatat atatatatat aataaaaaaa aatttttttg agacagagtc tctctctctt	780
gcccaggctg gagtgcaatg gcacaatcca atcacagctc actacagcct cagcctccca	840
gactcaggtg atcccatctc agcctcctga gtagctggga ccacaggtgc gtgccacat	900
gcccagcttt tttttttttt ttttttaatt gttacttgta gagacagggg ctccgtgtgc	960
tgcccaggct attcttgaac tcctgggctc aagtgatcct cctgtattgg cctcccaaag	1020
tgctgggatt gcaggtagga gctactgcat ctggccagta tgcagtattt tcatatctgg	1080
tagggctgag tccctcctca cagcttattt ttattgtttt cctagttttt catgtgaact	1140
ttactgtcag ttcgtctagc tccataaaaa ggtattttta ttgggatttt tgatgatgct	1200
gaaatgtcct gaccaagaac aagggatatt cagatgttct tttgtgtctt tcaggaatga	1260
tagttttgtt cacataagtt ttgaatgttt aagtttattt aagtttattt ctaaataattt	1320
tctcatttct ctggcttttg taagtagggg tttctcatcc atgttttctt ctcatgagtt	1380
atgtgtggat atgaaggcta tccattagta tatgttgatt tttatattac acttccttgc	1440
tcagttcatt attgattctt tttgagtttt ccaggcatat tctcacaagt aaagataata	1500
gaaatagttt gcttcctttc cacttctgct ttgaattttt ttttcttggt tcatttgcac	1560
tggtctgctc ctccagcaaa atgttaaata accctggaga tgatgggcaa ctctgttttg	1620
ctcctgacat tcgtgggggtg cctctggtgc ttccctgttg gtaaggggtt aactgtagcc	1680
ctgaggtggg acatttgatt ttaaaaaatca gtcactcttg ggcgcttagg ttagaggaat	1740
ggtaggcaga tgctgtcact ccttgccctt cccctcctcc tcccacctg gaggggaaat	1800

gaaatctgac	aggtagaaaag	agggggagttg	gggttctttt	tctctctccc	tccaccagca	1860										
tcactctctg	cctctccctc	aaaaatacgt	tcctgggtca	ggatatatgt	tgactcccta	1920										
gagagctctg	gagtcaacct	cctggccttc	ctccaccctc	actcttggcc	ttttctgccc	1980										
cccatttcct	ctacctgtgg	ggcatggagc	cacgagcctt	tgtgtgacgg	tttgctttct	2040										
ctctcctgtc	tttaggtgca	tggctgcctc	ctaateccat	agtcagagg	aggcatccct	2100										
aggactgcgg	gcaagggagc	cgggcaagcc	cagggcagcc	ttgaaccgtc	ccctggcctg	2160										
ccctccccgg	tggggggccag	g	atg	ctg	aag	aag	cag	tct	gca	ggg	ctt	gtg			2211	
		Met	Leu	Lys	Lys	Gln	Ser	Ala	Gly	Leu	Val					
		1				5						10				
ctg	tgg	ggc	gct	atc	ctc	ttt	gtg	gcc	tgg	aat	gcc	ctg	ctg	ctc	ctc	2259
Leu	Trp	Gly	Ala	Ile	Leu	Phe	Val	Ala	Trp	Asn	Ala	Leu	Leu	Leu	Leu	
				15					20					25		
ttc	ttc	tgg	acg	cgc	cca	gca	cct	ggc	agg	cca	ccc	tca	gtc	agc	gct	2307
Phe	Phe	Trp	Thr	Arg	Pro	Ala	Pro	Gly	Arg	Pro	Pro	Ser	Val	Ser	Ala	
				30					35					40		
ctc	gat	ggc	gac	ccc	gcc	agc	ctc	acc	cgg	gaa	gtg	att	cgc	ctg	gcc	2355
Leu	Asp	Gly	Asp	Pro	Ala	Ser	Leu	Thr	Arg	Glu	Val	Ile	Arg	Leu	Ala	
				45				50					55			
caa	gac	gcc	gag	gtg	gag	ctg	gag	cgg	cag	cgt	ggg	ctg	ctg	cag	cag	2403
Gln	Asp	Ala	Glu	Val	Glu	Leu	Glu	Arg	Gln	Arg	Gly	Leu	Leu	Gln	Gln	
						65					70					
atc	ggg	gat	gcc	ctg	tcg	agc	cag	cgg	ggg	agg	gtg	ccc	acc	gcg	gcc	2451
Ile	Gly	Asp	Ala	Leu	Ser	Ser	Gln	Arg	Gly	Arg	Val	Pro	Thr	Ala	Ala	
						80				85					90	
cct	ccc	gcc	cag	ccg	cgt	gtg	cct	gtg	acc	ccc	gcg	ccg	gcg	gtg	att	2499
Pro	Pro	Ala	Gln	Pro	Arg	Val	Pro	Val	Thr	Pro	Ala	Pro	Ala	Val	Ile	
						95				100					105	
ccc	atc	ctg	gtc	atc	gcc	tgt	gac	cgc	agc	act	gtt	cgg	cgc	tgc	ctg	2547
Pro	Ile	Leu	Val	Ile	Ala	Cys	Asp	Arg	Ser	Thr	Val	Arg	Arg	Cys	Leu	
						110				115				120		
gac	aag	ctg	ctg	cat	tat	cgg	ccc	tcg	gct	gag	ctc	ttc	ccc	atc	atc	2595
Asp	Lys	Leu	Leu	His	Tyr	Arg	Pro	Ser	Ala	Glu	Leu	Phe	Pro	Ile	Ile	
						125			130				135			
gtt	agc	cag	gac	tgc	ggg	cac	gag	gag	acg	gcc	cag	gcc	atc	gcc	tcc	2643
Val	Ser	Gln	Asp	Cys	Gly	His	Glu	Glu	Thr	Ala	Gln	Ala	Ile	Ala	Ser	
						145					150					
tac	ggc	agc	gcg	gtc	acg	cac	atc	cgg	cag	ccc	gac	ctg	agc	agc	att	2691
Tyr	Gly	Ser	Ala	Val	Thr	His	Ile	Arg	Gln	Pro	Asp	Leu	Ser	Ser	Ile	
						160					165				170	
gcg	gtg	ccg	ccg	gac	cac	cgc	aag	ttc	cag	ggc	tac	tac	aag	atc	gcg	2739
Ala	Val	Pro	Pro	Asp	His	Arg	Lys	Phe	Gln	Gly	Tyr	Tyr	Lys	Ile	Ala	

175							180							185							
cgc	cac	tac	cgc	tgg	gcg	ctg	ggc	cag	gtc	ttc	cgg	cag	ttt	cgc	ttc	2787					
Arg	His	Tyr	Arg	Trp	Ala	Leu	Gly	Gln	Val	Phe	Arg	Gln	Phe	Arg	Phe						
			190					195					200								
ccc	gcg	gcc	gtg	gtg	gtg	gag	gat	gac	ctg	gag	gtg	gcc	ccg	gac	ttc	2835					
Pro	Ala	Ala	Val	Val	Val	Glu	Asp	Asp	Leu	Glu	Val	Ala	Pro	Asp	Phe						
		205					210					215									
ttc	gag	tac	ttt	cgg	gcc	acc	tat	ccg	ctg	ctg	aag	gcc	gac	ccc	tcc	2883					
Phe	Glu	Tyr	Phe	Arg	Ala	Thr	Tyr	Pro	Leu	Leu	Lys	Ala	Asp	Pro	Ser						
	220					225					230										
ctg	tgg	tgc	gtc	tcg	gcc	tgg	aat	gac	aac	ggc	aag	gag	cag	atg	gtg	2931					
Leu	Trp	Cys	Val	Ser	Ala	Trp	Asn	Asp	Asn	Gly	Lys	Glu	Gln	Met	Val						
	235				240					245					250						
gac	gcc	agc	agg	cct	gag	ctg	ctc	tac	cgc	acc	gac	ttt	ttc	cct	ggc	2979					
Asp	Ala	Ser	Arg	Pro	Glu	Leu	Leu	Tyr	Arg	Thr	Asp	Phe	Phe	Pro	Gly						
				255					260					265							
ctg	ggc	tgg	ctg	ctg	ttg	gcc	gag	ctc	tgg	gct	gag	ctg	gag	ccc	aag	3027					
Leu	Gly	Trp	Leu	Leu	Leu	Ala	Glu	Leu	Trp	Ala	Glu	Leu	Glu	Pro	Lys						
			270					275					280								
tgg	cca	aag	gcc	ttc	tgg	gac	gac	tgg	atg	cgg	cgg	ccg	gag	cag	cgg	3075					
Trp	Pro	Lys	Ala	Phe	Trp	Asp	Asp	Trp	Met	Arg	Arg	Pro	Glu	Gln	Arg						
		285					290					295									
cag	ggg	cgg	gcc	tgc	ata	cgc	cct	gag	atc	tca	aga	acg	atg	acc	ttt	3123					
Gln	Gly	Arg	Ala	Cys	Ile	Arg	Pro	Glu	Ile	Ser	Arg	Thr	Met	Thr	Phe						
	300					305					310										
ggc	cgc	aag	ggt	gtg	agc	cac	ggg	cag	ttc	ttt	gac	cag	cac	ctc	aag	3171					
Gly	Arg	Lys	Gly	Val	Ser	His	Gly	Gln	Phe	Phe	Asp	Gln	His	Leu	Lys						
	315				320				325					330							
ttt	atc	aag	ctg	aac	cag	cag	ttt	gtg	cac	ttc	acc	cag	ctg	gac	ctg	3219					
Phe	Ile	Lys	Leu	Asn	Gln	Gln	Phe	Val	His	Phe	Thr	Gln	Leu	Asp	Leu						
				335				340						345							
tct	tac	ctg	cag	cgg	gag	gcc	tat	gac	cga	gat	ttc	ctc	gcc	cgc	gtc	3267					
Ser	Tyr	Leu	Gln	Arg	Glu	Ala	Tyr	Asp	Arg	Asp	Phe	Leu	Ala	Arg	Val						
			350					355					360								
tac	ggt	gct	ccc	cag	ctg	cag	gtg	gag	aaa	gtg	agg	acc	aat	gac	cgg	3315					
Tyr	Gly	Ala	Pro	Gln	Leu	Gln	Val	Glu	Lys	Val	Arg	Thr	Asn	Asp	Arg						
		365					370					375									
aag	gag	ctg	ggg	gag	gtg	cgg	gtg	cag	tat	acg	ggc	agg	gac	agc	ttc	3363					
Lys	Glu	Leu	Gly	Glu	Val	Arg	Val	Gln	Tyr	Thr	Gly	Arg	Asp	Ser	Phe						
		380				385					390										
aag	gct	ttc	gcc	aag	gct	ctg	ggt	gtc	atg	gat	gac	ctt	aag	tcg	ggg	3411					
Lys	Ala	Phe	Ala	Lys	Ala	Leu	Gly	Val	Met	Asp	Asp	Leu	Lys	Ser	Gly						
	395				400				405						410						
gtt	ccg	aga	gct	ggc	tac	cgg	ggt	att	gtc	acc	ttc	cag	ttc	cgg	ggc	3459					

Val Pro Arg Ala Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly
 415 420 425
 cgc cgt gtc cac ctg gcg ccc cca ccg acg tgg gag ggc tat gat cct 3507
 Arg Arg Val His Leu Ala Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro
 430 435 440
 agc tgg aat tag cacctgcctg tccttcctgg gccctcctt gccacatcat 3559
 Ser Trp Asn
 445
 gagctgaggt gggaccacag tccccagget gcatcggcct gcctgtgttt cctccttagg 3619
 tgcatttata tttttgattt ttccgagtgg catttaagtg cacaaatgat aacaagagga 3679
 ttattctccc gttctcaagg gagtcagatc aggggaacta ttctagggta tgttgccggg 3739
 tattaagcag gaaaccactg tgtgggtggg ggcactgggc ttgttggggc cagaaatgtc 3799
 cacgtcctga gctttctcct ggagcatgtg cagagagttt ggcaacgttc gctctcttga 3859
 ccagacccct tctccctgac ctggctcttc cagccagggc acgagccctc cttctatacc 3919
 tgctccccct cccccagtgg ggactgagtt atgggagaag gggacatatt tgtggccaaa 3979
 atgataactaa ccaaaggggc ttccttgtca gggcctgggt gagttgggtg gtcacgggg 4039
 ctcaactgct cctgcccttc tctcctgtct gacccccact tagcccttct ctcccttgag 4099
 cctagcagtt tatagttctg agatggaaag ttgaaggggg caagcaagac ctctcctcag 4159
 cccatgccc gctgtcagga gagaggtgca gggaggaagg ccttggtgct ggacaacctc 4219
 tctcttgctt tacctcagag agggactatg ccctgacccc tcctttctga aaatcagtgc 4279
 cctccctggt gctctaggag gctcctgctg gcttggtaga agacagaatt cgatctgcct 4339
 gtcccttttt cccctggggg ttgacacaca ggctcctctc agcatgaggt ggagcagtga 4399
 ccaggtggag cagtgaccag gacgcctctg gcccagtgtt gcccagcctc cccgcccgt 4459
 cccaggcgcc ccatgtcctc acaggccagg acgcatggc aggatggaga ggacttggtg 4519
 gatTTTTgtt tcttgctga cctcagtttc atgaaagaaa gtggaagcta cagaattatt 4579
 ttctaaaata aaggctgaat tgtctgaaaa atatttatgt gtgtgtgtcc tggaaaagaa 4639
 ggtggcaggc agggaaagaa aggaaaaggg agaataaaga gttaagaaga ggtctagacg 4699
 ggtggg 4705
 <210> 4
 <211> 445
 <212> PRT
 <213> homo sapiens
 <400> 4

Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asp Gly Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp Ala Leu Ser
 65 70 75 80
 Ser Gln Arg Gly Arg Val Pro Thr Ala Ala Pro Pro Ala Gln Pro Arg
 85 90 95
 Val Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
 100 105 110
 Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr
 115 120 125
 Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
 130 135 140
 His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr
 145 150 155 160
 His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
 165 170 175
 Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
 180 185 190
 Leu Gly Gln Val Phe Arg Gln Phe Arg Phe Pro Ala Ala Val Val Val
 195 200 205
 Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Arg Ala
 210 215 220
 Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
 225 230 235 240

Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Arg Pro Glu
 245 250 255
 Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
 260 265 270
 Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp
 275 280 285
 Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Ile
 290 295 300
 Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
 305 310 315 320
 His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
 325 330 335
 Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu
 340 345 350
 Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln Leu
 355 360 365
 Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val
 370 375 380
 Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala
 385 390 395 400
 Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr
 405 410 415
 Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His Leu Ala
 420 425 430
 Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 435 440 445
 <210> 5
 <211> 1464
 <212> DNA
 <213> homo sapiens

<220>

<221> sig_peptide

<222> (64)..(294)

<220>

<221> CDS

<222> (64)..(1263)

<220>

<221> mat_peptide

<222> (295)..(1263)

<400> 5

cggcccgcg ggcgggtccgt cccctcctgt agccacacacc cttcttaaag cggcggcggg 60

aag atg agg ctt cgg gag ccg ctc ctg agc ggc gcc gcg atg cca ggc 108
 Met Arg Leu Arg Glu Pro Leu Leu Ser Gly Ala Ala Met Pro Gly
 -75 -70 -65

gcg tcc cta cag cgg gcc tgc cgc ctg ctc gtg gcc gtc tgc gct ctg 156
 Ala Ser Leu Gln Arg Ala Cys Arg Leu Leu Val Ala Val Cys Ala Leu
 -60 -55 -50

cac ctt ggc gtc acc ctc gtt tac tac ctg gct ggc cgc gac ctg agc 204
 His Leu Gly Val Thr Leu Val Tyr Tyr Leu Ala Gly Arg Asp Leu Ser
 -45 -40 -35

cgc ctg ccc caa ctg gtc gga gtc tcc aca ccg ctg cag ggc ggc tcg 252
 Arg Leu Pro Gln Leu Val Gly Val Ser Thr Pro Leu Gln Gly Gly Ser
 -30 -25 -20 -15

aac agt gcc gcc gcc atc ggg cag tcc tcc ggg gag ctc cgg acc gga 300
 Asn Ser Ala Ala Ala Ile Gly Gln Ser Ser Gly Glu Leu Arg Thr Gly
 -10 -5 -1 1

ggg gcc cgg ccg ccg cct cct cta ggc gcc tcc tcc cag ccg cgc ccg 348
 Gly Ala Arg Pro Pro Pro Pro Leu Gly Ala Ser Ser Gln Pro Arg Pro
 5 10 15

ggg gcc gac tcc agc cca gtc gtg gat tct ggc cct ggc ccc gct agc 396
 Gly Gly Asp Ser Ser Pro Val Val Asp Ser Gly Pro Gly Pro Ala Ser
 20 25 30

aac ttg acc tcg gtc cca gtg ccc cac acc acc gca ctg tcg ctg ccc 444
 Asn Leu Thr Ser Val Pro Val Pro His Thr Thr Ala Leu Ser Leu Pro
 35 40 45 50

gcc tgc cct gag gag tcc ccg ctg ctt gtg ggc ccc atg ctg att gag 492
 Ala Cys Pro Glu Glu Ser Pro Leu Leu Val Gly Pro Met Leu Ile Glu
 55 60 65

ttt aac atg cct gtg gac ctg gag ctc gtg gca aag cag aac cca aat 540
 Phe Asn Met Pro Val Asp Leu Glu Leu Val Ala Lys Gln Asn Pro Asn
 70 75 80

gtg aag atg ggc ggc cgc tat gcc ccc agg gac tgc gtc tct cct cac 588
 Val Lys Met Gly Gly Arg Tyr Ala Pro Arg Asp Cys Val Ser Pro His
 85 90 95

aaa gtg gcc atc atc att cca ttc cgc aac cgg cag gag cac ctc aag	636
Lys Val Ala Ile Ile Ile Pro Phe Arg Asn Arg Gln Glu His Leu Lys	
100 105 110	
tac tgg cta tat tat ttg cac cca gtc ctg cag cgc cag cag ctg gac	684
Tyr Trp Leu Tyr Tyr Leu His Pro Val Leu Gln Arg Gln Gln Leu Asp	
115 120 125 130	
tat ggc atc tat ggc atc tat gtt atc aac cag gcg gga gac act ata	732
Tyr Gly Ile Tyr Gly Ile Tyr Val Ile Asn Gln Ala Gly Asp Thr Ile	
135 140 145	
ttc aat cgt gct aag ctc ctc aat gtt ggc ttt caa gaa gcc ttg aag	780
Phe Asn Arg Ala Lys Leu Leu Asn Val Gly Phe Gln Glu Ala Leu Lys	
150 155 160	
gac tat gac tac acc tgc ttt gtg ttt agt gac gtg gac ctc atc cca	828
Asp Tyr Asp Tyr Thr Cys Phe Val Phe Ser Asp Val Asp Leu Ile Pro	
165 170 175	
atg aat gac cat aat gcg tac agg tgt ttt tca cag cca cgg cac att	876
Met Asn Asp His Asn Ala Tyr Arg Cys Phe Ser Gln Pro Arg His Ile	
180 185 190	
tcc gtt gca atg gat aag ttt gga ttc agc cta cct tat gtt cag tat	924
Ser Val Ala Met Asp Lys Phe Gly Phe Ser Leu Pro Tyr Val Gln Tyr	
195 200 205 210	
ttt gga ggt gtc tct gct cta agt aaa caa cag ttt cta acc atc aat	972
Phe Gly Gly Val Ser Ala Leu Ser Lys Gln Gln Phe Leu Thr Ile Asn	
215 220 225	
gga ttt cct aat aat tat tgg ggc tgg gga gga gaa gat gat gac att	1020
Gly Phe Pro Asn Asn Tyr Trp Gly Trp Gly Gly Glu Asp Asp Asp Ile	
230 235 240	
ttt aac aga tta gtt ttt aga ggc atg tct ata tct cgc cca aat gct	1068
Phe Asn Arg Leu Val Phe Arg Gly Met Ser Ile Ser Arg Pro Asn Ala	
245 250 255	
gtg gtc ggg agg tgt cgc atg atc cgc cac tca aga gac aaa aaa aat	1116
Val Val Gly Arg Cys Arg Met Ile Arg His Ser Arg Asp Lys Lys Asn	
260 265 270	
gaa ccc aat cct cag agg ttt gac cga att gca cac aca aag gag aca	1164
Glu Pro Asn Pro Gln Arg Phe Asp Arg Ile Ala His Thr Lys Glu Thr	
275 280 285 290	
atg ctc tct gat ggt ttg aac tca ctc acc tac cag gtg ctg gat gta	1212
Met Leu Ser Asp Gly Leu Asn Ser Leu Thr Tyr Gln Val Leu Asp Val	
295 300 305	
cag aga tac cca ttg tat acc caa atc aca gtg gac atc ggg aca ccg	1260
Gln Arg Tyr Pro Leu Tyr Thr Gln Ile Thr Val Asp Ile Gly Thr Pro	
310 315 320	
agc tagcggttttg gtacacggat aagagactga aattagccag ggacctctgc	1313
Ser	

tgtgtgtctc tgccaatctg ctgggctggg ccctctcatt tttaccagtc tgagtgcacag 1373
 gtcccccttct tcgctcatca ttcagatggc tttccagatg accaggacga gtgggatatt 1433
 ttgcccccaa cttggctcgg catgtgaatt c 1464

<210> 6
 <211> 400
 <212> PRT
 <213> homo sapiens

<400> 6

Met Arg Leu Arg Glu Pro Leu Leu Ser Gly Ala Ala Met Pro Gly Ala
 -75 -70 -65

Ser Leu Gln Arg Ala Cys Arg Leu Leu Val Ala Val Cys Ala Leu His
 -60 -55 -50

Leu Gly Val Thr Leu Val Tyr Tyr Leu Ala Gly Arg Asp Leu Ser Arg
 -45 -40 -35 -30

Leu Pro Gln Leu Val Gly Val Ser Thr Pro Leu Gln Gly Gly Ser Asn
 -25 -20 -15

Ser Ala Ala Ala Ile Gly Gln Ser Ser Gly Glu Leu Arg Thr Gly Gly
 -10 -5 -1 1

Ala Arg Pro Pro Pro Pro Leu Gly Ala Ser Ser Gln Pro Arg Pro Gly
 5 10 15

Gly Asp Ser Ser Pro Val Val Asp Ser Gly Pro Gly Pro Ala Ser Asn
 20 25 30 35

Leu Thr Ser Val Pro Val Pro His Thr Thr Ala Leu Ser Leu Pro Ala
 40 45 50

Cys Pro Glu Glu Ser Pro Leu Leu Val Gly Pro Met Leu Ile Glu Phe
 55 60 65

Asn Met Pro Val Asp Leu Glu Leu Val Ala Lys Gln Asn Pro Asn Val
 70 75 80

Lys Met Gly Gly Arg Tyr Ala Pro Arg Asp Cys Val Ser Pro His Lys
 85 90 95

Val Ala Ile Ile Ile Pro Phe Arg Asn Arg Gln Glu His Leu Lys Tyr
 100 105 110 115
 Trp Leu Tyr Tyr Leu His Pro Val Leu Gln Arg Gln Gln Leu Asp Tyr
 120 125 130
 Gly Ile Tyr Gly Ile Tyr Val Ile Asn Gln Ala Gly Asp Thr Ile Phe
 135 140 145
 Asn Arg Ala Lys Leu Leu Asn Val Gly Phe Gln Glu Ala Leu Lys Asp
 150 155 160
 Tyr Asp Tyr Thr Cys Phe Val Phe Ser Asp Val Asp Leu Ile Pro Met
 165 170 175
 Asn Asp His Asn Ala Tyr Arg Cys Phe Ser Gln Pro Arg His Ile Ser
 180 185 190 195
 Val Ala Met Asp Lys Phe Gly Phe Ser Leu Pro Tyr Val Gln Tyr Phe
 200 205 210
 Gly Gly Val Ser Ala Leu Ser Lys Gln Gln Phe Leu Thr Ile Asn Gly
 215 220 225
 Phe Pro Asn Asn Tyr Trp Gly Trp Gly Gly Glu Asp Asp Asp Ile Phe
 230 235 240
 Asn Arg Leu Val Phe Arg Gly Met Ser Ile Ser Arg Pro Asn Ala Val
 245 250 255
 Val Gly Arg Cys Arg Met Ile Arg His Ser Arg Asp Lys Lys Asn Glu
 260 265 270 275
 Pro Asn Pro Gln Arg Phe Asp Arg Ile Ala His Thr Lys Glu Thr Met
 280 285 290
 Leu Ser Asp Gly Leu Asn Ser Leu Thr Tyr Gln Val Leu Asp Val Gln
 295 300 305
 Arg Tyr Pro Leu Tyr Thr Gln Ile Thr Val Asp Ile Gly Thr Pro Ser
 310 315 320

<210> 7.
 <211> 442
 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 7

Met Ala Leu Phe Leu Ser Lys Arg Leu Leu Arg Phe Thr Val Ile Ala
 1 5 10 15

Gly Ala Val Ile Val Leu Leu Leu Thr Leu Asn Ser Asn Ser Arg Thr
 20 25 30

Gln Gln Tyr Ile Pro Ser Ser Ile Ser Ala Ala Phe Asp Phe Thr Ser
 35 40 45

Gly Ser Ile Ser Pro Glu Gln Gln Val Ile Ser Glu Glu Asn Asp Ala
 50 55 60

Lys Lys Leu Glu Gln Ser Ala Leu Asn Ser Glu Ala Ser Glu Asp Ser
 65 70 75 80

Glu Ala Met Asp Glu Glu Ser Lys Ala Leu Lys Ala Ala Ala Glu Lys
 85 90 95

Ala Asp Ala Pro Ile Asp Thr Lys Thr Thr Met Asp Tyr Ile Thr Pro
 100 105 110

Ser Phe Ala Asn Lys Ala Gly Lys Pro Lys Ala Cys Tyr Val Thr Leu
 115 120 125

Val Arg Asn Lys Glu Leu Lys Gly Leu Leu Ser Ser Ile Lys Tyr Val
 130 135 140

Glu Asn Lys Ile Asn Lys Lys Phe Pro Tyr Pro Trp Val Phe Leu Asn
 145 150 155 160

Asp Glu Pro Phe Thr Glu Glu Phe Lys Glu Ala Val Thr Lys Ala Val
 165 170 175

Ser Ser Glu Val Lys Phe Gly Ile Leu Pro Lys Glu His Trp Ser Tyr
 180 185 190

Pro Glu Trp Ile Asn Gln Thr Lys Ala Ala Glu Ile Arg Ala Asp Ala
 195 200 205

Ala Thr Lys Tyr Ile Tyr Gly Gly Ser Glu Ser Tyr Arg His Met Cys
 210 215 220

Arg Tyr Gln Ser Gly Phe Phe Trp Arg His Glu Leu Leu Glu Glu Tyr
 225 230 235 240

Asp Trp Tyr Trp Arg Val Glu Pro Asp Ile Lys Leu Tyr Cys Asp Ile
 245 250 255

Asn Tyr Asp Val Phe Lys Trp Met Gln Glu Asn Glu Lys Val Tyr Gly
 260 265 270

Phe Thr Val Ser Ile His Glu Tyr Glu Val Thr Ile Pro Thr Leu Trp
 275 280 285

Gln Thr Ser Met Asp Phe Ile Lys Lys Asn Pro Glu Tyr Leu Asp Glu
 290 295 300

Asn Asn Leu Met Ser Phe Leu Ser Asn Asp Asn Gly Lys Thr Tyr Asn
 305 310 315 320

Leu Cys His Phe Trp Ser Asn Phe Glu Ile Ala Asn Leu Asn Leu Trp
 325 330 335

Arg Ser Pro Ala Tyr Arg Glu Tyr Phe Asp Thr Leu Asp His Gln Gly
 340 345 350

Gly Phe Phe Tyr Glu Arg Trp Gly Asp Ala Pro Val His Ser Ile Ala
 355 360 365

Ala Ala Leu Phe Leu Pro Lys Asp Lys Ile His Tyr Phe Ser Asp Ile
 370 375 380

Gly Tyr His His Pro Pro Tyr Asp Asn Cys Pro Leu Asp Lys Glu Val
 385 390 395 400

Tyr Asn Ser Asn Asn Cys Glu Cys Asp Gln Gly Asn Asp Phe Thr Phe
 405 410 415

Gln Gly Tyr Ser Cys Gly Lys Glu Tyr Tyr Asp Ala Gln Gly Leu Val
 420 425 430

Lys Pro Lys Asn Trp Lys Lys Phe Arg Glu
 435 440

<210> 8
 <211> 3599

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (4)..(3438)

<400> 8

aaa atg aag tta agc cgc cag ttc acc gtg ttc ggc agt gcg atc ttc	48
Met Lys Leu Ser Arg Gln Phe Thr Val Phe Gly Ser Ala Ile Phe	
1 5 10 15	

tgt gtg gtg att ttc tcg ctc tac ctg atg ctg gac cgg ggt cac tta	96
Cys Val Val Ile Phe Ser Leu Tyr Leu Met Leu Asp Arg Gly His Leu	
20 25 30	

gac tac ccc agg aac ccg cgc cgc gag ggc tcc ttc cct cag ggc cag	144
Asp Tyr Pro Arg Asn Pro Arg Arg Glu Gly Ser Phe Pro Gln Gly Gln	
35 40 45	

ctc tca atg ttg caa gaa aaa ata gac cat ttg gag cgt ttg cta gct	192
Leu Ser Met Leu Gln Glu Lys Ile Asp His Leu Glu Arg Leu Leu Ala	
50 55 60	

gag aat aat gag atc atc tca aat att aga gac tca gtc atc aat ttg	240
Glu Asn Asn Glu Ile Ile Ser Asn Ile Arg Asp Ser Val Ile Asn Leu	
65 70 75	

agt gag tct gtg gag gat ggt ccg aaa agt tca caa agc aat ttc agc	288
Ser Glu Ser Val Glu Asp Gly Pro Lys Ser Ser Gln Ser Asn Phe Ser	
80 85 90 95	

caa ggt gct ggc tca cat ctt ctg ccc tca caa tta tcc ctc tca gtt	336
Gln Gly Ala Gly Ser His Leu Leu Pro Ser Gln Leu Ser Leu Ser Val	
100 105 110	

gac act gca gac tgt ctg ttt gct tca caa agt gga agt cac aat tca	384
Asp Thr Ala Asp Cys Leu Phe Ala Ser Gln Ser Gly Ser His Asn Ser	
115 120 125	

gat gtg cag atg ttg gat gtt tac agt cta att tct ttt gac aat cca	432
Asp Val Gln Met Leu Asp Val Tyr Ser Leu Ile Ser Phe Asp Asn Pro	
130 135 140	

gat ggt gga gtt tgg aag caa gga ttt gac att act tat gaa tct aat	480
Asp Gly Gly Val Trp Lys Gln Gly Phe Asp Ile Thr Tyr Glu Ser Asn	
145 150 155	

gaa tgg gac act gaa ccc ctt caa gtc ttt gtg gtg cct cat tcc cat	528
Glu Trp Asp Thr Glu Pro Leu Gln Val Phe Val Val Pro His Ser His	
160 165 170 175	

aac gac cca ggt tgg ttg aag act ttc aat gac tac ttt aga gac aag	576
Asn Asp Pro Gly Trp Leu Lys Thr Phe Asn Asp Tyr Phe Arg Asp Lys	
180 185 190	

act cag tat att ttt aat aac atg gtc cta aag ctg aaa gaa gac tca	624
Thr Gln Tyr Ile Phe Asn Asn Met Val Leu Lys Leu Lys Glu Asp Ser	

195						200						205						
cgg	agg	aag	ttt	att	tgg	tct	gag	atc	tct	tac	ctt	tca	aag	tgg	tgg	672		
Arg	Arg	Lys	Phe	Ile	Trp	Ser	Glu	Ile	Ser	Tyr	Leu	Ser	Lys	Trp	Trp			
		210				215						220						
gat	att	ata	gat	att	cag	aag	aag	gat	gct	gtt	aaa	agt	tta	ata	gaa	720		
Asp	Ile	Ile	Asp	Ile	Gln	Lys	Lys	Asp	Ala	Val	Lys	Ser	Leu	Ile	Glu			
		225				230						235						
aat	ggt	cag	ctt	gaa	att	gtg	aca	ggt	ggc	tgg	gtt	atg	cct	gat	gaa	768		
Asn	Gly	Gln	Leu	Glu	Ile	Val	Thr	Gly	Gly	Trp	Val	Met	Pro	Asp	Glu			
240				245						250				255				
gct	act	cca	cat	tat	ttt	gcc	tta	att	gat	caa	cta	att	gaa	gga	cat	816		
Ala	Thr	Pro	His	Tyr	Phe	Ala	Leu	Ile	Asp	Gln	Leu	Ile	Glu	Gly	His			
				260										270				
cag	tgg	ctg	gaa	aat	aat	ata	gga	gtg	aaa	cct	cgg	tcc	ggc	tgg	gct	864		
Gln	Trp	Leu	Glu	Asn	Asn	Ile	Gly	Val	Lys	Pro	Arg	Ser	Gly	Trp	Ala			
		275						280				285						
att	gat	ccc	ttt	gga	cac	tca	cca	aca	atg	gct	tat	ctt	cta	aac	cgt	912		
Ile	Asp	Pro	Phe	Gly	His	Ser	Pro	Thr	Met	Ala	Tyr	Leu	Leu	Asn	Arg			
		290				295						300						
gct	gga	ctt	tct	cac	atg	ctt	atc	cag	aga	gtt	cat	tat	gca	gtt	aaa	960		
Ala	Gly	Leu	Ser	His	Met	Leu	Ile	Gln	Arg	Val	His	Tyr	Ala	Val	Lys			
305						310						315						
aaa	cac	ttt	gca	ctg	cat	aaa	aca	ttg	gag	ttt	ttt	tgg	aga	cag	aat	1008		
Lys	His	Phe	Ala	Leu	His	Lys	Thr	Leu	Glu	Phe	Phe	Trp	Arg	Gln	Asn			
320				325						330				335				
tgg	gat	ctg	gga	tct	gtc	aca	gat	att	tta	tgc	cac	atg	atg	ccc	ttc	1056		
Trp	Asp	Leu	Gly	Ser	Val	Thr	Asp	Ile	Leu	Cys	His	Met	Met	Pro	Phe			
				340				345				350						
tac	agc	tat	gac	atc	cct	cac	act	tgt	gga	cct	gat	cct	aaa	ata	tgc	1104		
Tyr	Ser	Tyr	Asp	Ile	Pro	His	Thr	Cys	Gly	Pro	Asp	Pro	Lys	Ile	Cys			
		355						360				365						
tgc	cag	ttt	gat	ttt	aaa	cgt	ctt	cct	gga	ggc	aga	ttt	ggt	tgt	ccc	1152		
Cys	Gln	Phe	Asp	Phe	Lys	Arg	Leu	Pro	Gly	Gly	Arg	Phe	Gly	Cys	Pro			
		370				375						380						
tgg	gga	gtc	ccc	cca	gaa	aca	ata	cat	cct	gga	aat	gtc	caa	agc	agg	1200		
Trp	Gly	Val	Pro	Pro	Glu	Thr	Ile	His	Pro	Gly	Asn	Val	Gln	Ser	Arg			
385						390				395								
gct	cgg	atg	cta	cta	gat	cag	tac	cga	aag	aag	tca	aag	ctt	ttt	cga	1248		
Ala	Arg	Met	Leu	Leu	Asp	Gln	Tyr	Arg	Lys	Lys	Ser	Lys	Leu	Phe	Arg			
400				405						410				415				
acc	aaa	gtt	ctc	ctg	gct	cca	cta	gga	gat	gat	ttc	cgc	tac	tgt	gaa	1296		
Thr	Lys	Val	Leu	Leu	Ala	Pro	Leu	Gly	Asp	Asp	Phe	Arg	Tyr	Cys	Glu			
				420				425						430				
tac	acg	gaa	tgg	gat	tta	cag	ttt	aag	aat	tat	cag	cag	ctt	ttt	gat	1344		

Tyr Thr Glu Trp Asp Leu Gln Phe Lys Asn Tyr Gln Gln Leu Phe Asp	
435 440 445	
tat atg aat tct cag tcc aag ttt aaa gtt aag ata cag ttt gga act	1392
Tyr Met Asn Ser Gln Ser Lys Phe Lys Val Lys Ile Gln Phe Gly Thr	
450 455 460	
tta tca gat ttt ttt gat gcg ctg gat aaa gca gat gaa act cag aga	1440
Leu Ser Asp Phe Phe Asp Ala Leu Asp Lys Ala Asp Glu Thr Gln Arg	
465 470 475	
gac aag ggc caa tcg atg ttc cct gtt tta agt gga gat ttt ttc act	1488
Asp Lys Gly Gln Ser Met Phe Pro Val Leu Ser Gly Asp Phe Phe Thr	
480 485 490 495	
tat gcc gat cga gat gat cat tac tgg agt ggc tat ttt aca tcc aga	1536
Tyr Ala Asp Arg Asp Asp His Tyr Trp Ser Gly Tyr Phe Thr Ser Arg	
500 505 510	
ccc ttt tac aaa cga atg gac aga atc atg gaa tct cat tta agg gct	1584
Pro Phe Tyr Lys Arg Met Asp Arg Ile Met Glu Ser His Leu Arg Ala	
515 520 525	
gct gaa att ctt tac tat ttc gcc ctg aga caa gct cac aaa tac aag	1632
Ala Glu Ile Leu Tyr Tyr Phe Ala Leu Arg Gln Ala His Lys Tyr Lys	
530 535 540	
ata aat aaa ttt ctc tca tca tca ctt tac acg gca ctg aca gaa gcc	1680
Ile Asn Lys Phe Leu Ser Ser Ser Leu Tyr Thr Ala Leu Thr Glu Ala	
545 550 555	
aga agg aat ttg gga ctg ttt caa cat cat gat gct atc aca gga act	1728
Arg Arg Asn Leu Gly Leu Phe Gln His His Asp Ala Ile Thr Gly Thr	
560 565 570 575	
gca aaa gac tgg gtg gtt gtg gat tat ggt acc aga ctt ttt cat tcg	1776
Ala Lys Asp Trp Val Val Val Asp Tyr Gly Thr Arg Leu Phe His Ser	
580 585 590	
tta atg gtt ttg gag aag ata att gga aat tct gca ttt ctt ctt att	1824
Leu Met Val Leu Glu Lys Ile Ile Gly Asn Ser Ala Phe Leu Leu Ile	
595 600 605	
ggg aag gac aaa ctc aca tac gac tct tac tct cct gat acc ttc ctg	1872
Gly Lys Asp Lys Leu Thr Tyr Asp Ser Tyr Ser Pro Asp Thr Phe Leu	
610 615 620	
gag atg gat ttg aaa caa aaa tca caa gat tct ctg cca caa aaa aat	1920
Glu Met Asp Leu Lys Gln Lys Ser Gln Asp Ser Leu Pro Gln Lys Asn	
625 630 635	
ata ata agg ctg agt gcg gag cca agg tac ctt gtg gtc tat aat cct	1968
Ile Ile Arg Leu Ser Ala Glu Pro Arg Tyr Leu Val Val Tyr Asn Pro	
640 645 650 655	
tta gaa caa gac cga atc tcg ttg gtc tca gtc tat gtg agt tcc ccg	2016
Leu Glu Gln Asp Arg Ile Ser Leu Val Ser Val Tyr Val Ser Ser Pro	
660 665 670	

aca gtg caa gtg ttc tct gct tca gga aaa cct gtg gaa gtt caa gtc Thr Val Gln Val Phe Ser Ala Ser Gly Lys Pro Val Glu Val Gln Val 675 680 685	2064
agc gca gtt tgg gat aca gca aat act att tca gaa aca gcc tat gag Ser Ala Val Trp Asp Thr Ala Asn Thr Ile Ser Glu Thr Ala Tyr Glu 690 695 700	2112
atc tct ttt cga gca cat ata ccg cca ttg gga ctg aaa gtg tat aag Ile Ser Phe Arg Ala His Ile Pro Pro Leu Gly Leu Lys Val Tyr Lys 705 710 715	2160
att ttg gaa tca gca agt tca aat tca cat tta gct gat tat gtc ttg Ile Leu Glu Ser Ala Ser Ser Asn Ser His Leu Ala Asp Tyr Val Leu 720 725 730 735	2208
tat aag aat aaa gta gaa gat agc gga att ttc acc ata aag aat atg Tyr Lys Asn Lys Val Glu Asp Ser Gly Ile Phe Thr Ile Lys Asn Met 740 745 750	2256
ata aat act gaa gaa ggt ata aca cta gag aac tcc ttt gtt tta ctt Ile Asn Thr Glu Glu Gly Ile Thr Leu Glu Asn Ser Phe Val Leu Leu 755 760 765	2304
cgg ttt gat caa act gga ctt atg aag caa atg atg act aaa gaa gat Arg Phe Asp Gln Thr Gly Leu Met Lys Gln Met Met Thr Lys Glu Asp 770 775 780	2352
ggt aaa cac cat gaa gta aat gtg caa ttt tca tgg tat gga acc aca Gly Lys His His Glu Val Asn Val Gln Phe Ser Trp Tyr Gly Thr Thr 785 790 795	2400
att aaa aga gac aaa agt ggt gcc tac ctc ttc tta cct gat ggt aat Ile Lys Arg Asp Lys Ser Gly Ala Tyr Leu Phe Leu Pro Asp Gly Asn 800 805 810 815	2448
gcc aag cct tat gtt tac aca aca ccg ccc ttt gtc aga gtg aca cat Ala Lys Pro Tyr Val Tyr Thr Thr Pro Pro Phe Val Arg Val Thr His 820 825 830	2496
gga agg att tat tcg gaa gtg act tgc ttt ttt gac cat gtt act cat Gly Arg Ile Tyr Ser Glu Val Thr Cys Phe Phe Asp His Val Thr His 835 840 845	2544
aga gtc cga cta tac cac ata cag gga ata gaa gga cag tct gtg gaa Arg Val Arg Leu Tyr His Ile Gln Gly Ile Glu Gly Gln Ser Val Glu 850 855 860	2592
gtt tcc aat att gtg gac atc cga aaa gta tat aac cgt gag att gca Val Ser Asn Ile Val Asp Ile Arg Lys Val Tyr Asn Arg Glu Ile Ala 865 870 875	2640
atg aaa att tct tct gat ata aaa agc caa aat aga ttt tat act gac Met Lys Ile Ser Ser Asp Ile Lys Ser Gln Asn Arg Phe Tyr Thr Asp 880 885 890 895	2688
cta aat ggg tac cag att caa cct aga atg aca ctg agc aaa ttg cct Leu Asn Gly Tyr Gln Ile Gln Pro Arg Met Thr Leu Ser Lys Leu Pro 900 905 910	2736

ctt caa gca aat gtc tat ccc atg acc aca atg gcc tat atc cag gat Leu Gln Ala Asn Val Tyr Pro Met Thr Thr Met Ala Tyr Ile Gln Asp 915 920 925	2784
gcc aaa cat cgt ttg aca ctg ctc tct gct cag tca tta ggg gtt tcg Ala Lys His Arg Leu Thr Leu Ser Ala Gln Ser Leu Gly Val Ser 930 935 940	2832
agt ttg aat agt ggt cag att gaa gtt atc atg gat cga aga ctc atg Ser Leu Asn Ser Gly Gln Ile Glu Val Ile Met Asp Arg Arg Leu Met 945 950 955	2880
caa gat gat aat cgt ggc ctt gag caa ggt atc cag gat aac aag att Gln Asp Asp Asn Arg Gly Leu Glu Gln Gly Ile Gln Asp Asn Lys Ile 960 965 970 975	2928
aca gct aat cta ttt cga ata cta cta gaa aaa aga agt gct gtt aat Thr Ala Asn Leu Phe Arg Ile Leu Leu Glu Lys Arg Ser Ala Val Asn 980 985 990	2976
acg gaa gaa gaa aag aag tcg gtc agt tat cct tct ctc ctt agc cac Thr Glu Glu Glu Lys Lys Ser Val Ser Tyr Pro Ser Leu Leu Ser His 995 1000 1005	3024
ata act tct tct ctc atg aat cat cca gtc att cca atg gca aat Ile Thr Ser Ser Leu Met Asn His Pro Val Ile Pro Met Ala Asn 1010 1015 1020	3069
aag ttc tcc tca cct acc ctt gag ctg caa ggt gaa ttc tct cca Lys Phe Ser Ser Pro Thr Leu Glu Leu Gln Gly Glu Phe Ser Pro 1025 1030 1035	3114
tta cag tca tct ttg cct tgt gac att cat ctg gtt aat ttg aga Leu Gln Ser Ser Leu Pro Cys Asp Ile His Leu Val Asn Leu Arg 1040 1045 1050	3159
aca ata cag tca aag gtg ggc aat ggg cac tcc aat gag gca gcc Thr Ile Gln Ser Lys Val Gly Asn Gly His Ser Asn Glu Ala Ala 1055 1060 1065	3204
ttg atc ctc cac aga aaa ggg ttt gat tgt cgg ttc tct agc aaa Leu Ile Leu His Arg Lys Gly Phe Asp Cys Arg Phe Ser Ser Lys 1070 1075 1080	3249
ggc aca ggg ctg ttt tgt tct act act cag gga aag ata ttg gta Gly Thr Gly Leu Phe Cys Ser Thr Thr Gln Gly Lys Ile Leu Val 1085 1090 1095	3294
cag aaa ctt tta aac aag ttt att gtc gaa agt ctc aca cct tca Gln Lys Leu Leu Asn Lys Phe Ile Val Glu Ser Leu Thr Pro Ser 1100 1105 1110	3339
tca cta tcc ttg atg cat tca cct ccc ggc act cag aat ata agt Ser Leu Ser Leu Met His Ser Pro Pro Gly Thr Gln Asn Ile Ser 1115 1120 1125	3384
gag atc aac ttg agt cca atg gaa atc agc aca ttc cga atc cag Glu Ile Asn Leu Ser Pro Met Glu Ile Ser Thr Phe Arg Ile Gln	3429

1130 1135 1140

ttg agg tga acctgacttt cacatttgga ttgagaatca ttggctttta 3478
 Leu Arg

tacctttctt ggtttgacgt gcaataaaga agcacattat tttagcttct ggctactgtg 3538

agaacatgaa ttctgtgatt ctgtggggtt tttctttttt tcttttacca gtacagtaag 3598

a 3599

<210> 9
 <211> 1144
 <212> PRT
 <213> Homo sapiens

<400> 9

Met Lys Leu Ser Arg Gln Phe Thr Val Phe Gly Ser Ala Ile Phe Cys
 1 5 10 15

Val Val Ile Phe Ser Leu Tyr Leu Met Leu Asp Arg Gly His Leu Asp
 20 25 30

Tyr Pro Arg Asn Pro Arg Arg Glu Gly Ser Phe Pro Gln Gly Gln Leu
 35 40 45

Ser Met Leu Gln Glu Lys Ile Asp His Leu Glu Arg Leu Leu Ala Glu
 50 55 60

Asn Asn Glu Ile Ile Ser Asn Ile Arg Asp Ser Val Ile Asn Leu Ser
 65 70 75 80

Glu Ser Val Glu Asp Gly Pro Lys Ser Ser Gln Ser Asn Phe Ser Gln
 85 90 95

Gly Ala Gly Ser His Leu Leu Pro Ser Gln Leu Ser Leu Ser Val Asp
 100 105 110

Thr Ala Asp Cys Leu Phe Ala Ser Gln Ser Gly Ser His Asn Ser Asp
 115 120 125

Val Gln Met Leu Asp Val Tyr Ser Leu Ile Ser Phe Asp Asn Pro Asp
 130 135 140

Gly Gly Val Trp Lys Gln Gly Phe Asp Ile Thr Tyr Glu Ser Asn Glu
 145 150 155 160

Trp Asp Thr Glu Pro Leu Gln Val Phe Val Val Pro His Ser His Asn
 165 170 175

Asp Pro Gly Trp Leu Lys Thr Phe Asn Asp Tyr Phe Arg Asp Lys Thr
 180 185 190

Gln Tyr Ile Phe Asn Asn Met Val Leu Lys Leu Lys Glu Asp Ser Arg
 195 200 205

Arg Lys Phe Ile Trp Ser Glu Ile Ser Tyr Leu Ser Lys Trp Trp Asp
 210 215 220

Ile Ile Asp Ile Gln Lys Lys Asp Ala Val Lys Ser Leu Ile Glu Asn
 225 230 235 240

Gly Gln Leu Glu Ile Val Thr Gly Gly Trp Val Met Pro Asp Glu Ala
 245 250 255

Thr Pro His Tyr Phe Ala Leu Ile Asp Gln Leu Ile Glu Gly His Gln
 260 265 270

Trp Leu Glu Asn Asn Ile Gly Val Lys Pro Arg Ser Gly Trp Ala Ile
 275 280 285

Asp Pro Phe Gly His Ser Pro Thr Met Ala Tyr Leu Leu Asn Arg Ala
 290 295 300

Gly Leu Ser His Met Leu Ile Gln Arg Val His Tyr Ala Val Lys Lys
 305 310 315 320

His Phe Ala Leu His Lys Thr Leu Glu Phe Phe Trp Arg Gln Asn Trp
 325 330 335

Asp Leu Gly Ser Val Thr Asp Ile Leu Cys His Met Met Pro Phe Tyr
 340 345 350

Ser Tyr Asp Ile Pro His Thr Cys Gly Pro Asp Pro Lys Ile Cys Cys
 355 360 365

Gln Phe Asp Phe Lys Arg Leu Pro Gly Gly Arg Phe Gly Cys Pro Trp
 370 375 380

Gly Val Pro Pro Glu Thr Ile His Pro Gly Asn Val Gln Ser Arg Ala
 385 390 395 400

Arg Met Leu Leu Asp Gln Tyr Arg Lys Lys Ser Lys Leu Phe Arg Thr
 405 410 415

Lys Val Leu Leu Ala Pro Leu Gly Asp Asp Phe Arg Tyr Cys Glu Tyr
 420 425 430

Thr Glu Trp Asp Leu Gln Phe Lys Asn Tyr Gln Gln Leu Phe Asp Tyr
 435 440 445

Met Asn Ser Gln Ser Lys Phe Lys Val Lys Ile Gln Phe Gly Thr Leu
 450 455 460

Ser Asp Phe Phe Asp Ala Leu Asp Lys Ala Asp Glu Thr Gln Arg Asp
 465 470 475 480

Lys Gly Gln Ser Met Phe Pro Val Leu Ser Gly Asp Phe Phe Thr Tyr
 485 490 495

Ala Asp Arg Asp Asp His Tyr Trp Ser Gly Tyr Phe Thr Ser Arg Pro
 500 505 510

Phe Tyr Lys Arg Met Asp Arg Ile Met Glu Ser His Leu Arg Ala Ala
 515 520 525

Glu Ile Leu Tyr Tyr Phe Ala Leu Arg Gln Ala His Lys Tyr Lys Ile
 530 535 540

Asn Lys Phe Leu Ser Ser Ser Leu Tyr Thr Ala Leu Thr Glu Ala Arg
 545 550 555 560

Arg Asn Leu Gly Leu Phe Gln His His Asp Ala Ile Thr Gly Thr Ala
 565 570 575

Lys Asp Trp Val Val Val Asp Tyr Gly Thr Arg Leu Phe His Ser Leu
 580 585 590

Met Val Leu Glu Lys Ile Ile Gly Asn Ser Ala Phe Leu Leu Ile Gly
 595 600 605

Lys Asp Lys Leu Thr Tyr Asp Ser Tyr Ser Pro Asp Thr Phe Leu Glu
 610 615 620

Met Asp Leu Lys Gln Lys Ser Gln Asp Ser Leu Pro Gln Lys Asn Ile

625		630		635		640
Ile Arg Leu Ser	Ala Glu Pro Arg Tyr Leu Val Val Tyr Asn Pro Leu					
	645		650		655	
Glu Gln Asp Arg	Ile Ser Leu Val Ser Val Tyr Val Ser Ser Pro Thr					
	660		665		670	
Val Gln Val Phe Ser Ala Ser Gly Lys Pro Val Glu Val Gln Val Ser						
	675		680		685	
Ala Val Trp Asp Thr Ala Asn Thr Ile Ser Glu Thr Ala Tyr Glu Ile						
	690		695		700	
Ser Phe Arg Ala His Ile Pro Pro Leu Gly Leu Lys Val Tyr Lys Ile						
705		710		715		720
Leu Glu Ser Ala Ser Ser Asn Ser His Leu Ala Asp Tyr Val Leu Tyr						
	725		730		735	
Lys Asn Lys Val Glu Asp Ser Gly Ile Phe Thr Ile Lys Asn Met Ile						
	740		745		750	
Asn Thr Glu Glu Gly Ile Thr Leu Glu Asn Ser Phe Val Leu Leu Arg						
	755		760		765	
Phe Asp Gln Thr Gly Leu Met Lys Gln Met Met Thr Lys Glu Asp Gly						
	770		775		780	
Lys His His Glu Val Asn Val Gln Phe Ser Trp Tyr Gly Thr Thr Ile						
785		790		795		800
Lys Arg Asp Lys Ser Gly Ala Tyr Leu Phe Leu Pro Asp Gly Asn Ala						
	805		810		815	
Lys Pro Tyr Val Tyr Thr Thr Pro Pro Phe Val Arg Val Thr His Gly						
	820		825		830	
Arg Ile Tyr Ser Glu Val Thr Cys Phe Phe Asp His Val Thr His Arg						
	835		840		845	
Val Arg Leu Tyr His Ile Gln Gly Ile Glu Gly Gln Ser Val Glu Val						
850		855		860		

Ser Asn Ile Val Asp Ile Arg Lys Val Tyr Asn Arg Glu Ile Ala Met
 865 870 875 880

Lys Ile Ser Ser Asp Ile Lys Ser Gln Asn Arg Phe Tyr Thr Asp Leu
 885 890 895

Asn Gly Tyr Gln Ile Gln Pro Arg Met Thr Leu Ser Lys Leu Pro Leu
 900 905 910

Gln Ala Asn Val Tyr Pro Met Thr Thr Met Ala Tyr Ile Gln Asp Ala
 915 920 925

Lys His Arg Leu Thr Leu Leu Ser Ala Gln Ser Leu Gly Val Ser Ser
 930 935 940

Leu Asn Ser Gly Gln Ile Glu Val Ile Met Asp Arg Arg Leu Met Gln
 945 950 955 960

Asp Asp Asn Arg Gly Leu Glu Gln Gly Ile Gln Asp Asn Lys Ile Thr
 965 970 975

Ala Asn Leu Phe Arg Ile Leu Leu Glu Lys Arg Ser Ala Val Asn Thr
 980 985 990

Glu Glu Glu Lys Lys Ser Val Ser Tyr Pro Ser Leu Leu Ser His Ile
 995 1000 1005

Thr Ser Ser Leu Met Asn His Pro Val Ile Pro Met Ala Asn Lys
 1010 1015 1020

Phe Ser Ser Pro Thr Leu Glu Leu Gln Gly Glu Phe Ser Pro Leu
 1025 1030 1035

Gln Ser Ser Leu Pro Cys Asp Ile His Leu Val Asn Leu Arg Thr
 1040 1045 1050

Ile Gln Ser Lys Val Gly Asn Gly His Ser Asn Glu Ala Ala Leu
 1055 1060 1065

Ile Leu His Arg Lys Gly Phe Asp Cys Arg Phe Ser Ser Lys Gly
 1070 1075 1080

Thr Gly Leu Phe Cys Ser Thr Thr Gln Gly Lys Ile Leu Val Gln
 1085 1090 1095

Lys Leu Leu Asn Lys Phe Ile Val Glu Ser Leu Thr Pro Ser Ser
1100 1105 1110

Leu Ser Leu Met His Ser Pro Pro Gly Thr Gln Asn Ile Ser Glu
 1115 1120 1125

Ile Asn Leu Ser Pro Met Glu Ile Ser Thr Phe Arg Ile Gln Leu
1130 1135 1140

Arg

```
<210> 10
<211> 3414
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> CDS
<222> (683)..(2026)
```

[illegible]

Ser	Asn	Gly	Arg	Gln	Arg	Lys	Asn	Glu	Ala	Leu	Ala	Pro	Pro	Leu	Leu	
		30						35				40				
gac	gcc	gaa	ccc	gcg	cgg	ggt	gcc	ggc	ggc	cgc	ggt	ggg	gac	cac	ccc	856
Asp	Ala	Glu	Pro	Ala	Arg	Gly	Ala	Gly	Gly	Arg	Gly	Gly	Asp	His	Pro	
		45					50					55				
tct	gtg	gct	gtg	ggc	atc	cgc	agg	gtc	tcc	aac	gtg	tcg	gcg	gct	tcc	904
Ser	Val	Ala	Val	Gly	Ile	Arg	Arg	Val	Ser	Asn	Val	Ser	Ala	Ala	Ser	
	60					65					70					
ctg	gtc	ccg	gcg	gtc	ccc	cag	ccc	gag	gcg	gac	aac	ctg	acg	ctg	cgg	952
Leu	Val	Pro	Ala	Val	Pro	Gln	Pro	Glu	Ala	Asp	Asn	Leu	Thr	Leu	Arg	
75					80					85					90	
tac	cgg	tcc	ctg	gtg	tac	cag	ctg	aac	ttt	gat	cag	acc	ctg	agg	aat	1000
Tyr	Arg	Ser	Leu	Val	Tyr	Gln	Leu	Asn	Phe	Asp	Gln	Thr	Leu	Arg	Asn	
			95					100					105			
gta	gat	aag	gct	ggc	acc	tgg	gcc	ccc	cgg	gag	ctg	gtg	ctg	gtg	gtc	1048
Val	Asp	Lys	Ala	Gly	Thr	Trp	Ala	Pro	Arg	Glu	Leu	Val	Leu	Val	Val	
		110					115					120				
cag	gtg	cat	aac	cgg	ccc	gaa	tac	ctc	aga	ctg	ctg	ctg	gac	tca	ctt	1096
Gln	Val	His	Asn	Arg	Pro	Glu	Tyr	Leu	Arg	Leu	Leu	Leu	Asp	Ser	Leu	
		125				130						135				
cga	aaa	gcc	cag	gga	att	gac	aac	gtc	ctc	gtc	atc	ttt	agc	cat	gac	1144
Arg	Lys	Ala	Gln	Gly	Ile	Asp	Asn	Val	Leu	Val	Ile	Phe	Ser	His	Asp	
	140					145					150					
ttc	tgg	tcg	acc	gag	atc	aat	cag	ctg	atc	gcc	ggg	gtg	aat	ttc	tgt	1192
Phe	Trp	Ser	Thr	Glu	Ile	Asn	Gln	Leu	Ile	Ala	Gly	Val	Asn	Phe	Cys	
155				160					165						170	
ccg	gtt	ctg	cag	gtg	ttc	ttt	cct	ttc	agc	att	cag	ttg	tac	cct	aac	1240
Pro	Val	Leu	Gln	Val	Phe	Phe	Pro	Phe	Ser	Ile	Gln	Leu	Tyr	Pro	Asn	
			175				180						185			
gag	ttt	cca	ggt	agt	gac	cct	aga	gat	tgt	ccc	aga	gac	ctg	ccg	aag	1288
Glu	Phe	Pro	Gly	Ser	Asp	Pro	Arg	Asp	Cys	Pro	Arg	Asp	Leu	Pro	Lys	
		190				195						200				
aat	gcc	gct	ttg	aaa	ttg	ggg	tgc	atc	aat	gct	gag	tat	ccc	gac	tcc	1336
Asn	Ala	Ala	Leu	Lys	Leu	Gly	Cys	Ile	Asn	Ala	Glu	Tyr	Pro	Asp	Ser	
	205					210					215					
ttc	ggc	cat	tat	aga	gag	gcc	aaa	ttc	tcc	cag	acc	aaa	cat	cac	tgg	1384
Phe	Gly	His	Tyr	Arg	Glu	Ala	Lys	Phe	Ser	Gln	Thr	Lys	His	His	Trp	
	220				225					230						
tgg	tgg	aag	ctg	cat	ttt	gtg	tgg	gaa	aga	gtg	aaa	att	ctt	cga	gat	1432
Trp	Trp	Lys	Leu	His	Phe	Val	Trp	Glu	Arg	Val	Lys	Ile	Leu	Arg	Asp	
235				240				245							250	
tat	gct	ggc	ctt	ata	ctt	ttc	cta	gaa	gag	gat	cac	tac	tta	gcc	cca	1480
Tyr	Ala	Gly	Leu	Ile	Leu	Phe	Leu	Glu	Glu	Asp	His	Tyr	Leu	Ala	Pro	
			255					260						265		

gac ttt tac cat gtc ttc aaa aag atg tgg aaa ctg aag cag caa gag 1528
 Asp Phe Tyr His Val Phe Lys Lys Met Trp Lys Leu Lys Gln Gln Glu
 270 275 280

tgc cct gaa tgt gat gtt ctc tcc ctg ggg acc tat agt gcc agt cgc 1576
 Cys Pro Glu Cys Asp Val Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg
 285 290 295

agt ttc tat ggc atg gct gac aag gta gat gtg aaa act tgg aaa tcc 1624
 Ser Phe Tyr Gly Met Ala Asp Lys Val Asp Val Lys Thr Trp Lys Ser
 300 305 310

aca gag cac aat atg ggt cta gcc ttg acc cgg aat gcc tat cag aag 1672
 Thr Glu His Asn Met Gly Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys
 315 320 325 330

ctg atc gag tgc aca gac act ttc tgt act tat gat gat tat aac tgg 1720
 Leu Ile Glu Cys Thr Asp Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp
 335 340 345

gac tgg act ctt caa tac ttg act gta tct tgt ctt cca aaa ttc tgg 1768
 Asp Trp Thr Leu Gln Tyr Leu Thr Val Ser Cys Leu Pro Lys Phe Trp
 350 355 360

aaa gtg ctg gtt cct caa att cct agg atc ttt cat gct gga gac tgt 1816
 Lys Val Leu Val Pro Gln Ile Pro Arg Ile Phe His Ala Gly Asp Cys
 365 370 375

ggt atg cat cac aag aaa acc tgt aga cca tcc act cag agt gcc caa 1864
 Gly Met His His Lys Lys Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln
 380 385 390

att gag tca ctc tta aat aat aac aaa caa tac atg ttt cca gaa act 1912
 Ile Glu Ser Leu Leu Asn Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr
 395 400 405 410

cta act atc agt gaa aag ttt act gtg gta gcc att tcc cca cct aga 1960
 Leu Thr Ile Ser Glu Lys Phe Thr Val Val Ala Ile Ser Pro Pro Arg
 415 420 425

aaa aat gga ggg tgg gga gat att agg gac cat gaa ctc tgt aaa agt 2008
 Lys Asn Gly Gly Trp Gly Asp Ile Arg Asp His Glu Leu Cys Lys Ser
 430 435 440

tat aga aga ctg cag tga aaatcacagt tacaaaagcg acagtcttct 2056
 Tyr Arg Arg Leu Gln
 445

atttttgata tttgtccaaa caggacatac aattgaataa aagagtttag gaactggttt 2116

ctgctttaat acaaaaacaa aatcttgtaa aagggtgtcca aatacatagt aatcttttcc 2176

agttatgtct gattaagatt taaaactgaa ggtttcattt tgggagtagg gttttaaaagc 2236

tcaatctgtt atctgctaaa attgattatt gttgatatga gagaagagg gaaattttat 2296

ttaaattgca tttattaatc tttttatctg aaactttgta cacttttcca ctttcaaaac 2356

ctattttaag tacagcaaaa tttatttaaa actgtgatag cagtaaaaag tattacgatg 2416

```

aaattgttag ggtattaatg gaacaaaccc agtttcactc tcttgacaca cttattagga 2476
agggattgct tcactgggtt aataatttaa aagttatggt tgtaaacac cctgtcagaa 2536
cagtcatttt cagtattaga ttcctgtact attgtgtttt gagtgtgttt tggaaccttc 2596
atagaacaca ctttcttttg gaatgtattt gattgataag aaagtttaaa cattgttttc 2656
acctcaatgt agaaatacag tggttttggt tttttttttc ttttagtgct gacaaaataa 2716
aataactcatt tttgcataaa aaggttccta atccttttgc agaataagtt ttgtttactc 2776
tttataccaa aattcagtga aggcatctta caagttttga gttagcatta ctttttaata 2836
tttactattg ctacattgta taattgagtt tgaaataaaa cccagcttat gacaatgcat 2896
tccctgtgca agaaactggt tggctttcaa attaccagg cattgaaaat gaatgataaa 2956
aagttgctgt gtaagggaaa tacagcctaa atgttttgaa agccagaaat gatacaaagt 3016
tcagtcatgc caaagtgaaa tactttctag tgccagcttt aacttaaadc atacgtttta 3076
aaaggacaga tacagaaaat tataggaaac aggcctaaat tttgctccat atttaatgta 3136
gacgtttata gaagtttccc ttaatttgta attgcattca accgagaatt tctcataaaa 3196
gactaatttc tgtgtaaaga tattacgggc tgggtgtggt ggctcatgtc tgtaatccag 3256
cactctggga ggttgaggca ggacgattgc ttgaactcag agtttgagac cagcctgggc 3316
aacatggcga aaaacccatc tctactaaaa ataacaaaaa attagccggg cgtagtgggtg 3376
actctgtagt cccagctact tgagaggctg aggtggga 3414

```

```

<210> 11
<211> 447
<212> PRT
<213> Homo sapiens

```

```

<400> 11

```

```

Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
1           5           10           15

```

```

Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
          20           25           30

```

```

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
          35           40           45

```

```

Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
50           55           60

```

```

Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
65                               70                               75                               80

Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
                               85                               90                               95

Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
                               100                              105                              110

Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
                               115                              120                              125

Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
130                               135                               140

Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
145                               150                               155                               160

Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
                               165                              170                              175

Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
                               180                              185                              190

Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
                               195                              200                              205

Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
210                               215                               220

Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
225                               230                              235                              240

Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
                               245                              250                              255

Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
                               260                              265                              270

Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
                               275                              280                              285

Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
290                               295                              300

```

Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320

Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335

Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350

Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365

Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380

Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400

Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415

Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430

Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

<210> 12
 <211> 394
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 12

Ser Leu Ser Leu Val Ser Tyr Arg Leu Arg Lys Asn Pro Trp Val Asn
 1 5 10 15

Ile Phe Leu Pro Val Leu Ala Ile Phe Leu Ile Tyr Ile Ile Phe Phe
 20 25 30

Gln Arg Asp Gln Ser Leu Leu Gly Leu Asn Gly Gln Ser Ile Ser Gln
 35 40 45

His Lys Trp Ala His Glu Lys Glu Asn Thr Phe Tyr Phe Pro Phe Thr
 50 55 60

Lys Lys Tyr Lys Met Pro Lys Tyr Ser Tyr Lys Lys Lys Ser Gly Trp
65 70 75 80

Leu Phe Asn Asp His Val Glu Asp Ile Ile Pro Glu Gly His Ile Ala
85 90 95

His Tyr Asp Leu Asn Lys Leu His Ser Thr Ser Glu Ala Ala Val Asn
100 105 110

Lys Glu His Ile Leu Ile Leu Thr Pro Met Gln Thr Phe His Gln Gln
115 120 125

Tyr Trp Asp Asn Leu Leu Gln Leu Asn Tyr Pro Arg Glu Leu Ile Glu
130 135 140

Leu Gly Phe Ile Thr Pro Arg Thr Ala Thr Gly Asp Leu Ala Leu Lys
145 150 155 160

Lys Leu Glu Asn Ala Ile Lys Lys Val Gln Thr Asp Lys Lys Thr Gln
165 170 175

Arg Phe Ser Lys Ile Thr Ile Leu Arg Gln Asn Ser Gln Ser Phe Asp
180 185 190

Lys Leu Met Glu Lys Glu Arg His Ala Leu Asp Val Gln Lys Glu Arg
195 200 205

Arg Ala Ala Met Ala Leu Ala Arg Asn Glu Leu Leu Phe Ser Thr Ile
210 215 220

Gly Pro His Thr Ser Trp Val Leu Trp Leu Asp Ala Asp Ile Ile Glu
225 230 235 240

Thr Pro Pro Ser Leu Ile Gln Asp Met Thr Lys His Asn Lys Ala Ile
245 250 255

Leu Ala Ala Asn Ile Tyr Gln Arg Phe Tyr Asp Glu Glu Lys Lys Gln
260 265 270

Pro Ser Ile Arg Pro Tyr Asp Phe Asn Asn Trp Gln Glu Ser Asp Thr
275 280 285

Gly Leu Glu Ile Ala Ser Gln Met Gly Asp Asp Glu Ile Ile Val Glu
290 295 300

Gly Tyr Ala Glu Ile Ala Thr Tyr Arg Pro Leu Met Ala His Phe Tyr
305 310 315 320

Asp Ala Asn Gly Val Pro Gly Glu Glu Met Ala Leu Asp Gly Val Gly
325 330 335

Gly Gly Cys Thr Leu Val Lys Ala Glu Val His Arg Asp Gly Ala Met
340 345 350

Phe Pro Asn Phe Pro Phe Tyr His Leu Ile Glu Thr Glu Gly Phe Ala
355 360 365

Lys Met Ala Lys Arg Leu Asn Tyr Asp Val Phe Gly Leu Pro Asn Tyr
370 375 380

Leu Val Tyr His Ile Glu Glu Glu Asn His
385 390

INTERNATIONAL SEARCH REPORT

International Application No

PCT/DK2005/000569

A. CLASSIFICATION OF SUBJECT MATTER

C12P21/00 C12N1/14 C12N9/24 C12N9/10 C12N5/10
C12N15/62 C12N15/80
//C12R1/645

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12P C12N

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

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

EPO-Internal, BIOSIS, MEDLINE, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HAMILTON S R ET AL: "Production of complex human glycoproteins in yeast" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 301, no. 5637, 29 August 2003 (2003-08-29), pages 1244-1246, XP002267832 ISSN: 0036-8075 cited in the application the whole document	1-16,18, 21-38, 44,45
X	EP 1 211 310 A (KIRIN BREWERY ;CHIBA YASUNORI (JP); ISHII TOMOKO (JP); KAINUMA MAM) 5 June 2002 (2002-06-05) the whole document	1-16,18, 21-38, 44,45
	----- -/--	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

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

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

Date of the actual completion of the international search

10 November 2005

Date of mailing of the international search report

29/11/2005

Name and mailing address of the ISA

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

Authorized officer

Aslund, J

INTERNATIONAL SEARCH REPORT

International Application No

PCT/DK2005/000569

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2004/003194 A (FLANDERS INTERUNIVERSITY INSTITUTE FOR BIOTECH) 8 January 2004 (2004-01-08) the whole document	1-16,18, 21-38, 44,45
X	VERVECKEN WOUTER ET AL: "In vivo synthesis of mammalian-like, hybrid-type N-glycans in Pichia pastoris." APPLIED AND ENVIRONMENTAL MICROBIOLOGY. MAY 2004, vol. 70, no. 5, May 2004 (2004-05), pages 2639-2646, XP002329570 ISSN: 0099-2240 the whole document	1-5,7-9, 18, 21-25, 28-30, 38,44,45
X	CHOI B-K ET AL: "Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 100, no. 9, 29 April 2003 (2003-04-29), pages 5022-5027, XP002267831 ISSN: 0027-8424 cited in the application the whole document	1-9, 21-25, 28-30, 38,44,45
X	WO 02/00879 A (GLYCOFI, INC) 3 January 2002 (2002-01-03) cited in the application the whole document	1-16,18, 21-38, 44,45
X	CHIBA YASUNORI ET AL: "Production of human compatible high mannose-type (Man5GlcNAc2) sugar chains in Saccharomyces cerevisiae." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 41, 9 October 1998 (1998-10-09), pages 26298-26304, XP002202331 ISSN: 0021-9258 cited in the application Figure 1 the whole document	1-5, 21-26, 28,38
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/DK2005/000569

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ICHISHIMA EIJI ET AL: "Molecular and enzymic properties of recombinant 1,2-alpha-mannosidase from Aspergillus saitoi overexpressed in Aspergillus oryzae cells" BIOCHEMICAL JOURNAL, vol. 339, no. 3, 1 May 1999 (1999-05-01), pages 589-597, XP002353325 ISSN: 0264-6021 figure 3	1,17,40
A	KALSNER INGE ET AL: "Insertion into Aspergillus nidulans of functional UDP-GlcNAc: alpha-3-D-mannoside beta-1,2-N-acetylglucosaminyl-transferase I, the enzyme catalysing the first committed step from oligomannose to hybrid and complex N-glycans." GLYCOCONJUGATE JOURNAL, vol. 12, no. 3, 1995, pages 360-370, XP001073956 ISSN: 0282-0080 the whole document	17,40
A	NAGASU T ET AL: "ISOLATION OF NEW TEMPERATURE-SENSITIVE MUTANTS OF SACCHAROMYCES CEREVISIAE DEFICIENT IN MANNOSE OUTER CHAIN ELONGATION" YEAST, CHICHESTER, SUSSEX, GB, vol. 7, July 1992 (1992-07), pages 535-547, XP000770465 ISSN: 0749-503X	18,38
A	EP 0 314 096 A (ZYMOGENETICS INC; ZYMOGENETICS, INC) 3 May 1989 (1989-05-03)	38,39
A	JARS METTE U ET AL: "N- and O-glycosylation and phosphorylation of the bar secretion leader derived from the barrier protease of Saccharomyces cerevisiae" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 42, 1995, pages 24810-24817, XP002329571 ISSN: 0021-9258	
A	YIP CARLI LAM ET AL: "Cloning and analysis of the Saccharomyces cerevisiae MNN9 and MNN1 genes required for complex glycosylation of secreted proteins" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, vol. 91, no. 7, 1994, pages 2723-2727, XP002329572 ISSN: 0027-8424 page 2726, column 1	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/DK2005/000569

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 1211310	A	05-06-2002	AU 6592800 A CA 2382184 A1 CN 1379811 A WO 0114522 A1	19-03-2001 01-03-2001 13-11-2002 01-03-2001
WO 2004003194	A	08-01-2004	AU 2003238051 A1 CA 2490268 A1 EP 1516048 A2 US 2004018588 A1	19-01-2004 08-01-2004 23-03-2005 29-01-2004
WO 0200879	A	03-01-2002	AU 7684201 A CA 2412701 A1 EP 1297172 A2 JP 2004501642 T NZ 523476 A	08-01-2002 03-01-2002 02-04-2003 22-01-2004 30-04-2004
EP 0314096	A	03-05-1989	AU 2435688 A CA 1325781 C DE 3887082 D1 DE 3887082 T2 DK 599788 A FI 884991 A JP 2000419 A	11-05-1989 04-01-1994 24-02-1994 28-04-1994 30-04-1989 30-04-1989 05-01-1990