TITLE OF THE INVENTION

IMPROVED VECTORS AND YEAST STRAINS FOR PROTEIN PRODUCTION: CA2+ ATPASE OVEREXPRESSION

BACKGROUND OF THE INVENTION

(1) Field of the Invention

The present invention relates to host cells that include one or more nucleic acid molecules encoding a Ca^{2+} ATPase, endoplasmic reticulum lectin chaperones, e.g., calreticulin (CRT) or calnexin (CRX), and/or ERp57 protein and their use for producing recombinant glycoproteins that have reduced *O*-glycosylation.

(2) Description of Related Art

Glycoproteins mediate many essential functions in humans and other mammals, including catalysis, signaling, cell-cell communication, and molecular recognition and association. Glycoproteins make up the majority of non-cytosolic proteins in eukaryotic organisms (Lis and Sharon, Eur. J. Biochem. 218: 1-27 (1993)). Many glycoproteins have been exploited for therapeutic purposes, and during the last two decades, recombinant versions of naturally-occurring glycoproteins have been a major part of the biotechnology industry. Examples of recombinant glycosylated proteins used as therapeutics include erythropoietin (EPO), therapeutic monoclonal antibodies (mAbs), tissue plasminogen activator (tPA), interferon-β (IFN-β), granulocyte-macrophage colony stimulating factor (GM-CSF)5 and human chorionic gonadotrophin (hCH) (Gumming *et al.*, Glycobiology 1:115-130 (1991)). Variations in glycosylation patterns of recombinantly produced glycoproteins have recently been the topic of much attention in the scientific community as recombinant proteins produced as potential prophylactics and therapeutics approach the clinic.

In general, the glycosylation structures of glycoprotein oligosaccharides will vary depending upon the host species of the cells used to produce them. Therapeutic proteins produced in non-human host cells are likely to contain non-human glycosylation which may elicit an immunogenic response in humans—e.g. hypermannosylation in yeast (Ballou, Methods Enzymol. 185:440-470 (1990); α(l,3)-fucose and β(l,2)-xylose in plants, (Cabanes-Macheteau *et al*, Glycobiology 9: 365-372 (1999)); N-glycolylneuraminic acid in Chinese hamster ovary cells (Noguchi *et al.*, J. Biochem. 117: 5-62 (1995); and, Galα-l,3Gal glycosylation in mice (Borrebaeck *et al.*, Immunol. Today, 14: 477-479 (1993).

Because the oligosaccharide structures of glycoproteins produced by non-human mammalian cells tend to be more closely related to those of human glycoproteins, most commercial glycoproteins are produced in mammalian cells. However, mammalian cells have several important disadvantages as host cells for protein production. Besides being costly,

processes for producing proteins in mammalian cells produce heterogeneous populations of glycoforms, have low volumetric titers, and require both ongoing viral containment and significant time to generate stable cell lines. Until about 2000, lower eukaryote host cells suitable for producing recombinant glycoproteins with human-like *N*-glycosylation patterns had not been possible. Since then, Gerngross in U.S. Patent No. 7,029,872 disclosed methods for making recombinant lower eukaryote host cells that are capable of making glycoproteins that have human-like N-glycosylation patterns. Thus, there is now considerable interest in using lower eukaryote host cells to produce recombinant glycoproteins.

While the pathway for *N*-linked glycosylation has been the subject of much analysis, the process and function of *O*-linked glycosylation is not as well understood. It is known that in contrast to *N*-linked glycosylation, *O*-glycosylation is a posttranslational event, which occurs in the cis-Golgi (Varki, Glycobiol., 3: 97-130 (1993)). While a consensus acceptor sequence for *O*-linked glycosylation like that for *N*-linked glycosylation does not appear to exist, a comparison of amino acid sequences around a large number of *O*-linked glycosylation sites of several glycoproteins show an increased frequency of proline residues at positions -1 and +3 relative to the glycosylated residues and a marked increase of serine, threonine, and alanine residues (Wilson *et al.*, Biochem. J., 275: 529- 534 (1991)). Stretches of serine and threonine residues in glycoproteins, may also be potential sites for *O*-glycosylation. It has been shown that yeast-derived recombinant proteins often bear additional unnatural *O*-glycans compared to their natural counterpart (Van den Steen, et al., Crit. Reviews in Biochem. and Mole. Biol. 33: 151-208, (1998)). These unnatural *O*-glycans can result in proteins that have unwanted immunogenicity or aberrant activity. Thus, there is a need to develop methods for producing proteins in yeast and other lower eukaryotes that have reduced or no *O*-glycosylation.

Tanner *et al.* in U. S. Patent No. 5,714,377 describes the *PMT1* and *PMT2* genes of *Saccharomyces cerevisiae* and a method for making recombinant proteins having reduced O-linked glycosylation that uses fungal cells in which one or more of *PMT* genes have been genetically modified so that recombinant proteins are produced, which have reduced *O*-linked glycosylation.

Ng *et al.* in U.S. Published Patent Application No. 20020068325 discloses inhibition of *O*-glycosylation through the use of antisense or cosuppression or through the engineering of yeast host strains that have loss of function mutations in genes associated with *O*-linked glycosylation, in particular, one or more of the *PMT* genes.

Clausen in U.S. Published Patent Application No. 20030186850 discloses the use of GalNAc-beta-benzyl to selectively inhibit lectins of polypeptide GalNAc-transferases and not serve as substrates for other glycosyltransferases involved in *O*-glycan biosyntheses, thus inhibiting *O*-glycosylation.

Orchard *et al.* in U.S. Patent No. 7,105,554 describes benzylidene thiazolidinediones and their use as antimycotic agents, e.g., antifungal agents which Bobrowicz *et al.* in WO2007061631 show can be used in a way which is not lethal to the host cells for production of recombinant proteins with reduced *O*-linked glycosylation.

Konrad *et al.* in U.S. Published Patent Application No. 20020128235 disclose a method for treating or preventing diabetes mellitus by pharmacologically inhibiting *O*-linked protein glycosylation in a tissue or cell.

Kojima *et al.* in U. S. Patent No. 5,268,364 disclose therapeutic compositions for inhibition of *O*-glycosylation using compounds such as benzyle-α-N-acetylgalactosamine, which inhibits extension of *O*-glycosylation leading to accumulation of *O*-α-GalNAc, to block expression of SLex or SLea by leukocytes or tumor cells and thereby inhibit adhesion of these cells to endothelial cells and platelets.

Boime *et al.* in U. S. Patent No. 6,103,501 disclose variants of hormones in which *O*-linked glycosylation was altered by modifying the amino acid sequence at the site of glycosylation.

However, even in light of the above attempts to produce recombinant host cells that produce proteins that have reduced or no *O*-glycosylation, there still remains a need for host cells that are capable of producing recombinant proteins that have reduced *O*-glycosylation.

BRIEF SUMMARY OF THE INVENTION

The present inventors have found that expression of recombinant proteins in a recombinant host cell with reduced *O*-glycosylation can be effected by overexpressing an endogenous or exogenous Ca^{2+} ATPase in the recombinant host cell. Host cells that overexpress an endogenous or exogenous Ca^{2+} ATPase produce recombinant proteins with reduced *O*-glycosylation compared to the same cells that do not overexpress the Ca^{2+} ATPase. As shown in the examples, recombinant host lower eukaryote host cells that included an expression cassette encoding a heterologous or endogenous Ca^{2+} ATPase were capable of producing recombinant proteins wherein the *O*-glycan occupancy was reduced by up to 4 fold compared to cells that did not overexpress an endogenous or exogenous Ca^{2+} ATPase.

Thus, the present invention provides lower eukaryotic host cells, in which a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase is introduced into and expressed in the host cell, wherein expression of the Ca^{2+} ATPase is ectopic. In particular aspects, the Ca^{2+} ATPase is encoded by an open reading frame operably linked to a heterologous regulatory sequences, which may provide constitutive or regulatable expression of the Ca^{2+} ATPase, and which is operable in the host cell. In further aspects, the lower eukaryotic host cell is a yeast or filamentous fungi host cell. In further still aspects, the host cell is a methylotrophic yeast, for example *Pichia pastoris*. In particular aspects, the Ca^{2+} ATP is

selected from the group consisting of the *Pichia pastoris PMR1* and the *Arabidopsis thaliana* ECA1.

In further aspects, the lower eukaryotic host cells of the invention are further transformed with a recombinant vector comprising regulatory nucleotide sequences derived from lower eukaryotic host cells and a coding sequence encoding a selected mammalian protein to be produced by the above host cells. In certain aspects, the selected mammalian protein is a therapeutic protein, and may be a glycoprotein, including but not limited to, an antibody.

In further embodiments, the host cell may be a yeast or filamentous fungal host cell, such as a *Pichia pastoris* cell, in which a vector encoding at least one endogenous or exogenous Ca^{2+} ATPase is introduced into and expressed in the host cell and the host cell further expresses a nucleic acid molecule comprising regulatory nucleotide sequences derived from or functional in *Pichia pastoris* cells operably linked with an open reading frame encoding a human therapeutic glycoprotein, such as an antibody, which is introduced into the host cell.

It has also been found that overexpressing a calreticulin and an ERp57 protein in the lower eukaryote host cells also effected a reduction in *O*-glycan occupancy. Thus, also provided are lower eukaryote host cells comprising one or more nucleic acid molecules encoding a calreticulin and/or an ERp57 protein wherein the proteins are ectopically expressed. In further embodiments, the host cells include a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase, wherein expression of the Ca^{2+} ATPase is ectopic. In general, the lower eukaryote host cell further includes a nucleic acid molecule encoding a recombinant protein, which in particular aspects, is a glycoprotein, which in further aspects is an antibody or fragment thereof such as Fc or Fab.

In further embodiments, any one of the above host cell is engineered to reduce or eliminate the function of at least one endogenous *Pichia pastoris* gene encoding a protein *O*-mannosyltransferase (*PMT*) protein to provide a host cell that is capable of making recombinant proteins having reduced *O*-glycosylation compared to host cells that have functional *PMT* genes. In further aspects, the *PMT* protein is selected from the group consisting of *PMT1* and *PMT4*. In further aspects, the host cells are further contacted with one or more inhibitors of *PMT* gene expression or *PMT* protein function.

In further embodiments, the gene encoding an endogenous chaperone protein is reduced, deleted, or disrupted and a nucleic acid molecule encoding a heterologous chaperone protein is introduced into the cell. In particular aspects, the chaperone protein is the PDI1 protein.

In further aspects of the above host cells, the host cell is selected from the group consisting of *Pichia pastoris, Pichia finlandica, Pichia trehalophila, Pichia koclamae, Pichia membranaefaciens, Pichia minuta (Ogataea minuta, Pichia lindneri), Pichia opuntiae, Pichia thermotolerans, Pichia salictaria, Pichia guercuum, Pichia pijperi, Pichia stipitis, Pichia*

methanolica, Pichia sp., *Saccharomyces cerevisiae, Saccharomyces* sp., *Schizosaccharomyces sp., Schizosaccharomyce pombe, Hansenula polymorpha, Kluyveromyces sp., Kluyveromyces lactis, Candida albicans, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Trichoderma reesei, Chrysosporium lucknowense, Fusarium* sp., *Fusarium gramineum, Fusarium venenatum, Physcomitrella patens* and *Neurospora crassa. Pichia sp.,* any *Saccharomyces sp., Hansenula polymorpha,* any *Kluyveromyces sp.,* any *Aspergillus sp., Trichoderma reesei, Chrysosporium lucknowense,* any *Fusarium sp.* and *Neurospora crassa.*

Further embodiments include methods for producing recombinant proteins that have reduced *O*-glycosylation or *O*-glycan occupancy compared to recombinant glycoproteins that do not include the genetic modifications disclosed herein. Recombinant proteins include proteins and glycoproteins of therapeutic relevance, including antibodies and fragments thereof.

Thus, provided is a method for producing a recombinant protein comprising: (a) providing a lower eukaryote host cell comprising a nucleic acid molecule encoding an endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic; (b) introducing a nucleic acid molecule into the host cell encoding the recombinant protein: and (c) growing the host cell under conditions suitable for producing the recombinant protein.

Further provided is a method for producing a recombinant protein comprising: (a) providing a lower eukaryote host cell comprising a nucleic acid molecule encoding at least one of CRT or ERp57, wherein expression of the CRT and/or ERp57 in the host cell is ectopic; (b) introducing a nucleic acid molecule into the host cell encoding the recombinant protein: and (c) growing the host cell under conditions suitable for producing the recombinant protein.

Further provided is a method for producing a recombinant protein comprising: (a) providing a lower eukaryote host cell comprising nucleic acid molecules encoding an endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic and at least one of CRT or ERp57, wherein expression of the Ca^{2+} ATPase, CRT and/or ERp57 in the host cell is ectopic; (b) introducing a nucleic acid molecule into the host cell encoding the recombinant protein: and (c) growing the host cell under conditions suitable for producing the recombinant protein.

In further embodiments, the function of at least one endogenous *Pichia pastoris* gene encoding a protein *O*-mannosyltransferase (*PMT*) protein to provide a host cell that is capable of making recombinant proteins having reduced *O*-glycosylation compared to host cells that have functional *PMT* genes. In further aspects, the *PMT* protein is selected from the group consisting of *PMT1* and *PMT4*. In further aspects, the host cells are further contacted with one or more inhibitors of *PMT* gene expression or *PMT* protein function.

In further embodiments, the gene encoding an endogenous chaperone protein is reduced, deleted, or disrupted and a nucleic acid molecule encoding a heterologous chaperone

protein is introduced into the cell. In particular aspects, the chaperone protein is the PDI1 protein.

In further still aspects, any one of the host cells disclosed herein can be grown in the presence of an inhibitor of a *PMT* gene.

The methods herein are particularly useful for producing proteins of therapeutic value, including but not limited to, antibodies. Thus provided is the use of any one of the host cells herein for producing a protein of therapeutic value. In particular aspects, use of any one of the host cells herein for producing an antibody.

In further aspects of the above methods, the host cell is selected from the group consisting of *Pichia pastoris, Pichia finlandica, Pichia trehalophila, Pichia koclamae, Pichia membranaefaciens, Pichia minuta (Ogataea minuta, Pichia lindneri), Pichia opuntiae, Pichia thermotolerans, Pichia salictaria, Pichia guercuum, Pichia pijperi, Pichia stipitis, Pichia methanolica, Pichia* sp., *Saccharomyces cerevisiae, Saccharomyces* sp., *Schizosaccharomyces sp., Schizosaccharomyce pombe, Hansenula polymorpha, Kluyveromyces sp., Kluyveromyces lactis, Candida albicans, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Trichoderma reesei, Chrysosporium lucknowense, Fusarium* sp., *Fusarium gramineum, Fusarium venenatum, Physcomitrella patens* and *Neurospora crassa. Pichia sp.,* any *Saccharomyces sp., Hansenula polymorpha,* any *Kluyveromyces sp.,* any *Aspergillus sp., Trichoderma reesei, Chrysosporium lucknowense,* any *Fusarium sp.* and *Neurospora crassa.*

Further provided are recombinant proteins produced by the host cells disclosed herein.

In particular embodiments, any one of the aforementioned host cells can further include genetic modifications that enable the host cells to produce glycoproteins have predominantly particular *N*-glycan structures thereon or particular mixtures of *N*-glycan structures thereon. For example, the host cells have been genetically engineered to produce *N*-glycans having a $Man_3GlcNAc_2$ or $Man_5GlcNAc_2$ core structure, which in particular aspects include one or more additional sugars such as GlcNAc, Galactose, or sialic acid on the non-reducing end, and optionally fucose on the GlcNAc at the reducing end. Thus, the *N*-glycans include both bi-antennary and multi-antennary glycoforms and glycoforms that are bisected. Examples of *N*-glycans include but are not limited to $Man_8GlcNAc_2$, $Man_7GlcNAc_2$, $Man_6GlcNAc_2$, $Man_5GlcNAc_2$, $GlcNAcMan_5GlcNAc_2$, $GalGlcNAcMan_5GlcNAc_2$, $NANAGalGlcNAcMan_5GlcNAc_2$, $Man_3GlcNAc_2$, $GlcNAc_{(1-4)}Man_3GlcNAc_2$, $Gal_{(1-4)}GlcNAc_{(1-4)}Man_3GlcNAc_2$, $NANA_{(1-4)}Gal_{(1-4)}GlcNAc_{(1-4)}Man_3GlcNAc_2$.

Definitions

Unless otherwise defined herein, scientific and technical terms and phrases used in connection with the present invention shall have the meanings that are commonly understood

by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include the plural and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of biochemistry, enzymology, molecular and cellular biology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. *See, e.g.*, Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Associates (1992, and Supplements to 2002); Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1990); Taylor and Drickamer, *Introduction to Glycobiology*, Oxford Univ. Press (2003); *Worthington Enzyme Manual*, Worthington Biochemical Corp., Freehold, NJ; *Handbook of Biochemistry: Section A Proteins*, Vol I, CRC Press (1976); *Handbook of Biochemistry: Section A Proteins*, Vol II, CRC Press (1976); *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press (1999).

All publications, patents and other references mentioned herein are hereby incorporated by reference in their entireties.

The following terms, unless otherwise indicated, shall be understood to have the following meanings:

As used herein, the terms "*N*-glycan" and "glycoform" are used interchangeably and refer to an *N*-linked oligosaccharide, e.g., one that is attached by an asparagine-*N*-acetylglucosamine linkage to an asparagine residue of a polypeptide. *N*-linked glycoproteins contain an *N*-acetylglucosamine residue linked to the amide nitrogen of an asparagine residue in the protein. The predominant sugars found on glycoproteins are glucose, galactose, mannose, fucose, *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc) and sialic acid (e.g., *N*-acetyl-neuraminic acid (NANA)). The processing of the sugar groups occurs cotranslationally in the lumen of the ER and continues in the Golgi apparatus for *N*-linked glycoproteins.

N-glycans have a common pentasaccharide core of $Man_3GlcNAc_2$ ("Man" refers to mannose; "Glc" refers to glucose; and "NAc" refers to *N*-acetyl; GlcNAc refers to *N*-acetylglucosamine). *N*-glycans differ with respect to the number of branches (antennae) comprising peripheral sugars (e.g., GlcNAc, galactose, fucose and sialic acid) that are added to the $Man_3GlcNAc_2$ ("Man3") core structure which is also referred to as the "trimannose core", the "pentasaccharide core" or the "paucimannose core". *N*-glycans are classified according to their branched constituents (e.g., high mannose, complex or hybrid). A "high mannose" type *N*-glycan has five or more mannose residues. A "complex" type *N*-glycan typically has at least one

GlcNAc attached to the 1,3 mannose arm and at least one GlcNAc attached to the 1,6 mannose arm of a "trimannose" core. Complex *N*-glycans may also have galactose ("Gal") or *N*-acetylgalactosamine ("GalNAc") residues that are optionally modified with sialic acid or derivatives (e.g., "NANA" or "NeuAc", where "Neu" refers to neuraminic acid and "Ac" refers to acetyl). Complex *N*-glycans may also have intrachain substitutions comprising "bisecting" GlcNAc and core fucose ("Fuc"). Complex *N*-glycans may also have multiple antennae on the "trimannose core," often referred to as "multiple antennary glycans." A "hybrid" *N*-glycan has at least one GlcNAc on the terminal of the 1,3 mannose arm of the trimannose core and zero or more mannoses on the 1,6 mannose arm of the trimannose core. The various *N*-glycans are also referred to as "glycoforms."

Abbreviations used herein are of common usage in the art, see, e.g., abbreviations of sugars, above. Other common abbreviations include "PNGase", or "glycanase" or "glucosidase" which all refer to peptide *N*-glycosidase F (EC 3.2.2.18).

The term "vector" as used herein is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid molecule to which it has been linked. One type of vector is a "plasmid vector", which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome (discussed in more detail below). Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, vectors having an origin of replication which functions in the host cell). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and are thereby replicated along with the host genome. Moreover, certain preferred vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "recombinant expression vectors" (or simply, "expression vectors").

As used herein, the term "sequence of interest" or "gene of interest" refers to a nucleic acid sequence, typically encoding a protein, that is not normally produced in the host cell. The methods disclosed herein allow efficient expression of one or more sequences of interest or genes of interest stably integrated into a host cell genome. Non-limiting examples of sequences of interest include sequences encoding one or more polypeptides having an enzymatic activity, *e.g.*, an enzyme which affects *N*-glycan synthesis in a host such as mannosyltransferases, *N*-acetylglucosaminyltransferases, UDP-*N*-acetylglucosamine transporters, galactosyltransferases, UDP-*N*-acetylgalactosyltransferase, sialyltransferases and fucosyltransferases.

The term "marker sequence" or "marker gene" refers to a nucleic acid sequence capable of expressing an activity that allows either positive or negative selection for the presence or absence of the sequence within a host cell. For example, the *Pichia pastoris URA5* gene is a

marker gene because its presence can be selected for by the ability of cells containing the gene to grow in the absence of uracil. Its presence can also be selected against by the inability of cells containing the gene to grow in the presence of 5-FOA. Marker sequences or genes do not necessarily need to display both positive and negative selectability. Non-limiting examples of marker sequences or genes from *Pichia pastoris* include *ADE1*, *ARG4*, *HIS4* and *URA3*. For antibiotic resistance marker genes, kanamycin, neomycin, geneticin (or G418), paromomycin and hygromycin resistance genes are commonly used to allow for growth in the presence of these antibiotics.

"Operatively linked" expression control sequences refers to a linkage in which the expression control sequence is contiguous with the gene of interest to control the gene of interest, as well as expression control sequences that act in *trans* or at a distance to control the gene of interest.

The term "expression control sequence" or "regulatory sequences" are used interchangeably and as used herein refer to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*e.g.*, ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, all components whose presence is essential for expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

The term "recombinant host cell" ("expression host cell", "expression host system", "expression system" or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein. A recombinant host cell may be an isolated cell or cell line grown in culture or may be a cell which resides in a living tissue or organism.

The term "eukaryotic" refers to a nucleated cell or organism, and includes insect cells, plant cells, mammalian cells, animal cells and lower eukaryotic cells.

The term "lower eukaryotic cells" includes yeast and filamentous fungi. Yeast and filamentous fungi include, but are not limited to: *Pichia pastoris, Pichia finlandica, Pichia trehalophila, Pichia koclamae, Pichia membranaefaciens, Pichia minuta (Ogataea minuta, Pichia lindneri), Pichia opuntiae, Pichia thermotolerans, Pichia salictaria, Pichia guercuum, Pichia pijperi, Pichia stipitis, Pichia methanolica, Pichia* sp., *Saccharomyces cerevisiae, Saccharomyces* sp., *Schizosaccharomyces sp., Schizosaccharomyce pombe, Hansenula polymorpha, Kluyveromyces sp., Kluyveromyces lactis, Candida albicans, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Trichoderma reesei, Chrysosporium lucknowense, Fusarium* sp., *Fusarium gramineum, Fusarium venenatum, Physcomitrella patens* and *Neurospora crassa. Pichia sp.,* any *Saccharomyces sp., Hansenula polymorpha,* any *Kluyveromyces sp.,* any *Aspergillus sp., Trichoderma reesei, Chrysosporium lucknowense,* any *Fusarium sp.* and *Neurospora crassa.*

The function of a gene encoding a protein is said to be 'reduced' when that gene has been modified, for example, by deletion, insertion, mutation or substitution of one or more nucleotides, such that the modified gene encodes a protein which has at least 20% to 50% lower activity, in particular aspects, at least 40% lower activity or at least 50% lower activity, when measured in a standard assay, as compared to the protein encoded by the corresponding gene without such modification. The function of a gene encoding a protein is said to be 'eliminated' when the gene has been modified, for example, by deletion, insertion, mutation or substitution of one or more nucleotides, such that the modified gene encodes a protein which has at least 90% to 99% lower activity, in particular aspects, at least 95% lower activity or at least 99% lower activity, when measured in a standard assay, as compared to the protein encoded by the corresponding gene without such modification.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice of the present invention and will be apparent to those of skill in the art. All publications and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting in any manner.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show the genealogy of yeast strains described in the examples for illustrating the invention.

Figure 2 illustrates the construction of plasmid vector pGLY642 encoding the human PDI1 (hPDI) and targeting the *Pichia pastoris PDI1* locus.

Figure 3 illustrates the construction of plasmid vector pGLY2232 encoding the human ERO1α (hERO1α) and targeting the *Pichia pastoris PrB1* locus.

Figure 4 illustrates the construction of plasmid vector pGLY2233 encoding the human GRP94 and targeting the *Pichia pastoris PEP4* locus.

Figure 5 illustrates the construction of plasmid vectors pGLY1896 and pGFI207t encoding the *T. reesei* α-1,2 mannosidase (TrMNS1) and mouse α-1,2 mannosidase IA (FB53) and targeting the *Pichia pastoris PRO* locus.

Figure 6 illustrates the construction of plasmid vector pGLY1162 encoding the *T. reesei* α-1,2 mannosidase (TrMNS1) and targeting the *Pichia pastoris PRO* locus.

Figure 7 is a map of plasmid vectors pGLY2260 and pGLY2261 encoding the anti-DKK1 antibody heavy chain (GFI710H) and light chain (GFI710L) and targeting the *Pichia pastoris TRP2* locus and targeting the *Pichia pastoris TRP2* locus.

Figure 8 is a map of plasmid vector pGLY3822 encoding the *Pichia pastoris PMR1* and targeting the *Pichia pastoris URA6* locus.

Figure 9 is a map of plasmid vector pGLY3827 encoding the *Arabidopsis thaliana* ECA1 (AtECA1) and targeting the *Pichia pastoris URA6* locus.

Figure 10 is a map of plasmid vector pGLY1234 encoding the human CRT (hCRT) and human ERp57(hERp57) and targeting the *Pichia pastoris HIS3* locus.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides recombinant host cells that are capable of producing recombinant proteins that have reduced *O*-glycosylation compared to host cells that have not been genetically engineered as disclosed herein. In general, provided are recombinant host cells comprising one or more nucleic acid molecules for ectopic expression of one or more endogenous or exogenous Ca^{2+} ATPases and the use of the recombinant host cells to produce glycoproteins that have reduced *O*-glycosylation.

We have found that overexpression of an endogenous or exogenous Ca^{2+} ATPase in recombinant host cells enabled us to produce recombinant proteins that had reduced *O*-glycosylation compared to host cells that did not overexpress an endogenous or exogenous Ca^{2+} ATPase. As shown in Examples 3 and 4, overexpression of *Pichia pastoris* Golgi Ca^{2+} ATPase (Pp*PMR1*) or *Arabidopsis thaliana* ER Ca^{2+} ATPase (AtECA1) effected greater than a 4-fold reduction in *O*-glycan occupancy compared to the host cell strains that did not express either Ca^{2+} ATPase. Thus, recombinant host cells that include one or more nucleic acid molecules encoding an endogenous or exogenous Golgi or ER Ca^{2+} ATPase, wherein the Ca^{2+} ATPase is operably linked to a heterologous promoter, will provide host cells that are capable of producing

recombinant glycoproteins that have reduced *O*-glycosylation. These host cells can be used for producing recombinant proteins in which it is desired that the amount of *O*-glycosylation on the protein is reduced. Other Ca^{2+} ATPases that are suitable include but are not limited to human SERCA2b protein (ATP2A2 ATPase, Ca^{++} transporting, cardiac muscle, slow twitch 2) and the *Pichia pastoris COD1* protein (homologue of *Saccharomyces cerevisiae* SPF1).

Calreticulin (CRT) is a multifunctional protein that acts as a major Ca(2+)-binding (storage) protein in the lumen of the endoplasmic reticulum. It is also found in the nucleus, suggesting that it may have a role in transcription regulation. Calreticulin binds to the synthetic peptide KLGFFKR (SEQ ID NO:47), which is almost identical to an amino acid sequence in the DNA-binding domain of the superfamily of nuclear receptors. Calreticulin binds to antibodies in certain sera of systemic lupus and Sjogren patients which contain anti-Ro/SSA antibodies, it is highly conserved among species, and it is located in the endoplasmic and sarcoplasmic reticulum where it may bind calcium. Calreticulin binds to misfolded proteins and prevents them from being exported from the Endoplasmic reticulum to the Golgi apparatus. Other proteins that are suitable include but are not limited to human UGGT (UDP-glucose:glycoprotein glucosyltransferase) protein and human ERp27 protein.

ERp57 is a chaperone protein of the endoplasmic reticulum that interacts with lectin chaperones calreticulin and calnexin to modulate folding of newly synthesized glycoproteins. The protein was once thought to be a phospholipase; however, it has been demonstrated that the protein actually has protein disulfide isomerase activity. Thus, the ERp57 is a lumenal protein of the endoplasmic reticulum (ER) and a member of the protein disulfide isomerase (PDI) family. It is thought that complexes of lectins and this protein mediate protein folding by promoting formation of disulfide bonds in their glycoprotein substrates. In contrast to archetypal PDI, ERp57 interacts specifically with newly synthesized glycoproteins.

We have further found that overexpression of the human CRT and human ERp57 in *Pichia pastoris* effected about a one-third reduction in *O*-glycan occupancy compared to strains which did not express the hCRT and hERp57.

Thus, further provided are recombinant host cells comprising one or more nucleic acid molecules encoding a calreticulin protein and/or ERp57 protein for ectopic expression in the host cell. These host cells can be used for producing recombinant proteins where it is desired that the amount of *O*-glycosylation on the protein is reduced. When the host cells further include one or more nucleic acid molecules encoding an endogenous or heterologous Ca^{2+} ATPase, these host cells have a further reduction in *O*-glycosylation. As shown in Example 4, providing a recombinant host cell that overexpressed either an endogenous Ca^{2+} ATPase or an exogenous Ca^{2+} ATPase and overexpressed the human calreticulin protein and human ERp57 protein had a further reduction in the *O*-glycosylation of recombinant proteins produced by the host cells. Thus, further provided are recombinant host cells comprising one or more nucleic acid molecules

encoding an endogenous or heterologous Ca^{2+} ATPase and one or more nucleic acid molecules encoding a calreticulin protein and/or an ERp57 protein. These host cells can be used to produce glycoproteins with reduced *O*-glycosylation.

Molecular chaperones play a critical role in the folding and secretion of antibodies. One chaperone protein in particular, Protein Disulfide Isomerase (PDI), functions to catalyze inter and intra disulphide bond formation that link the antibody heavy and light chains. Protein disulfide isomerase (PDI) can produce a substantial increase or a substantial decrease in the recovery of disulfide-containing proteins, when compared with the uncatalyzed reaction; a high concentration of PDI in the endoplasmic reticulum (ER) is essential for the expression of disulfide-containing proteins (Puig and Gilbert, J. Biol. Chem. 269: 7764-7771 (1994)). As shown in the Examples, cells in which the endogenous PDI1 chaperone gene has been replaced with a human PDI chaperone gene had reduced *O*-glycosylation. When these cells further include ectopic overexpression of an endogenous or exogenous Ca^{2+} ATPase and/or CRT and/or ERp57 protein, there was a further reduction in *O*-glycosylation (*See* Examples 3 and 4).

Thus, further included are host cells that ectopically express a CA^{2+} ATPase and/or CRT and/or ERp57 protein and wherein one or more genes encoding an endogenous chaperone protein has been deleted or disrupted and a nucleic acid molecule encoding a heterologous chaperone protein has been introduced for ectopic expression of the chaperone protein. Further embodiments, include the above cells wherein additional heterologous co-chaperone proteins, such as ERO-1α and/or the GRP94 proteins are also expressed in the cells.

Lower eukaryotic cells such as *Saccharomyces cerevisiae*, *Candida albicans*, and *Pichia pastoris*, contain a family of genes known as protein *O*-mannosyltransferases (*PMTs*) involved in the transfer of mannose to seryl and threonyl residues of secretory proteins. We found that *Pichia pastoris* cell lines, which have been genetically altered to express one or more humanized or chimeric chaperone genes, are better able to tolerate deletion of one or more *PMT* genes, with little or no effect on cell growth or protein expression. *PMT* genes which may be deleted include *PMT1*, *PMT2*, *PMT4*, *PMT5*, and *PMT6*. In general, *Pichia pastoris* host cells in which both the *OCH1* gene and the *PMT* gene is deleted either grow poorly or not at all. Deletion or functional knockout of the *OCH1* gene is necessary for constructing recombinant *Pichia pastoris* host cells that can make human glycoproteins that have human-like *N*-glycans. Because it is desirable to produce human glycoproteins that have no or reduced *O*-glycosylation, there has been a need to find means for reducing *O*-glycosylation in recombinant *Pichia pastoris* host cells that are also capable of producing human glycoproteins with human-like *N*-glycans. Thus, in further embodiments, provided are host cells that further include deletion or disruption of one or more *PMT* genes.

In further aspects, the overexpressed gene product is a secreted gene product. Procedures for observing whether an overexpressed gene product is secreted are readily available

to the skilled artisan. For example, Goeddel, (Ed.) 1990, Gene Expression Technology, Methods in Enzymology, Vol 185, Academic Press, and Sambrook *et al.* 1989, Molecular Cloning: A Laboratory Manual, Vols. 1-3, Cold Spring Harbor Press, N.Y., provide procedures for detecting secreted gene products.

To secrete an overexpressed gene product the host cell is cultivated under conditions sufficient for secretion of the overexpressed gene product. Such conditions include temperature, nutrient and cell density conditions that permit secretion by the cell. Moreover, such conditions are conditions under which the cell can perform basic cellular functions of transcription, translation and passage of proteins from one cellular compartment to another and are known to the skilled artisan.

Moreover, as is known to the skilled artisan a secreted gene product can be detected in the culture medium used to maintain or grow the present host cells. The culture medium can be separated from the host cells by known procedures, for example, centrifugation or filtration. The overexpressed gene product can then be detected in the cell-free culture medium by taking advantage of known properties characteristic of the overexpressed gene product. Such properties can include the distinct immunological, enzymatic or physical properties of the overexpressed gene product. For example, if an overexpressed gene product has a unique enzyme activity an assay for that activity can be performed on the culture medium used by the host cells. Moreover, when antibodies reactive against a given overexpressed gene product are available, such antibodies can be used to detect the gene product in any known immunological assay (*See* Harlowe, *et al.*, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press).

In addition, a secreted gene product can be a fusion protein wherein the gene product includes a heterologous signal or leader peptide that facilitates the secretion of the gene product. Secretion signal peptides are discrete amino acid sequences, which cause the host cell to direct a gene product through internal and external cellular membranes and into the extracellular environment. Secretion signal peptides are present at the *N*-terminus of a nascent polypeptide gene product targeted for secretion. Additional eukaryotic secretion signals can also be present along the polypeptide chain of the gene product in the form of carbohydrates attached to specific amino acids, i.e. glycosylation secretion signals.

N-terminal signal peptides include a hydrophobic domain of about 10 to about 30 amino acids which can be preceded by a short charged domain of about two to about 10 amino acids. Moreover, the signal peptide is present at the *N*-terminus of gene products destined for secretion. In general, the particular sequence of a signal sequence is not critical but signal sequences are rich in hydrophobic amino acids such as alanine (Ala) , valine (Val), leucine (Leu), isoleucine (Ile), proline (Pro), phenylalanine (Phe), tryptophan (Trp), methionine (Met) and the like.

Many signal peptides are known (Michaelis *et al.*, Ann. Rev. Microbiol. 36: 425 (1982). For example, the yeast acid phosphatase, yeast invertase, and the yeast α-factor signal peptides have been attached to heterologous polypeptide coding regions and used successfully for secretion of the heterologous polypeptide (See for example, Sato *et al.* Gene 83: 355-365 (1989); Chang *et al.* Mol. Cell. Biol. 6: 1812-1819 (1986); and Brake *et al.* Proc. Natl. Acad. Sci. USA 81: 4642-4646 (1984). Therefore, the skilled artisan can readily design or obtain a nucleic acid molecule which encodes a coding region for an overexpressed gene product which also has a signal peptide at the 5'-end.

Examples of overexpressed gene products which are preferably secreted by the present methods include mammalian gene products such as enzymes, cytokines, growth factors, hormones, vaccines, antibodies and the like. More particularly, overexpressed gene products include but are not limited to gene products such as erythropoietin, insulin, somatotropin, growth hormone releasing factor, platelet derived growth factor, epidermal growth factor, transforming growth factor α., transforming growth factor β, epidermal growth factor, fibroblast growth factor, nerve growth factor, insulin-like growth factor I, insulin-like growth factor II, clotting Factor VIII, superoxide dismutase, α-interferon, γ-interferon, interleukin-1, interleukin-2, interleukin-3, interleukin-4, interleukin-5, interleukin-6, granulocyte colony stimulating factor, multi-lineage colony stimulating activity, granulocyte-macrophage stimulating factor, macrophage colony stimulating factor, T cell growth factor, lymphotoxin, immunoglobulins, antibodies, and the like. Further included are fusion proteins, including but not limited to, peptides and polypeptides fused to the constant region of an immunoglobulin or antibody. Particularly useful overexpressed gene products are human gene products.

The terms "antibody", "antibodies", and "immunoglobulin(s)" encompass any recombinant monoclonal antibody produced by recombinant DNA technology and further is meant to include humanized and chimeric antibodies.

The present methods can readily be adapted to enhance secretion of any overexpressed gene product which can be used as a vaccine. Overexpressed gene products which can be used as vaccines include any structural, membrane-associated, membrane-bound or secreted gene product of a mammalian pathogen. Mammalian pathogens include viruses, bacteria, single-celled or multi-celled parasites which can infect or attack a mammal. For example, viral vaccines can include vaccines against viruses such as human immunodeficiency virus (HIV), *R. rickettsii*, vaccinia, *Shigella*, poliovirus, adenovirus, influenza, hepatitis A, hepatitis B, dengue virus, Japanese B encephalitis, Varicella zoster, cytomegalovirus, hepatitis A, rotavirus, as well as vaccines against viral diseases like Lyme disease, measles, yellow fever, mumps, rabies, herpes, influenza, parainfluenza and the like. Bacterial vaccines can include vaccines against bacteria such as *Vibrio cholerae*, *Salmonella typhi*, *Bordetella pertussis*, *Streptococcus pneumoniae*, *Hemophilus influenza*, *Clostridium tetani*, *Corynebacterium*

diphtheriae, *Mycobacterium leprae*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Coccidioides immitis*, and the like.

In general, the overexpressed proteins of the present invention (for example, Ca^{2+} ATPase, ERp57, calreticulin) and recombinant protein are expressed recombinantly, that is, by placing a nucleic acid molecule encoding an overexpressed protein or recombinant protein into an expression cassette. Such an expression cassette minimally contains a regulatory sequence which effects expression of the protein when the sequence is operably linked to a nucleic acid molecule encoding the protein. The expression cassette is then inserted into a vector such as a plasmid that can also contain additional elements like origins of replication, selectable markers, transcription or termination signals, centromeres, autonomous replication sequences, and the like to provide an expression vector.

An expression vector can be a replicable or a non-replicable expression vector. A replicable expression vector can replicate either independently of host cell chromosomal DNA or because such a vector has integrated into host cell chromosomal DNA. An integrating expression vector comprises a targeting sequence that targets the expression vector to a particular location in the host cell genome where the vector then integrates. Upon integration into host cell chromosomal DNA such an expression vector can lose some structural elements but retains the nucleic acid molecule encoding the overexpressed or recombinant protein and a segment which can effect expression of the overexpressed or recombinant protein. Therefore, the expression vectors herein can be chromosomally integrating or chromosomally nonintegrating expression vectors.

In a further embodiment, one or more overexpressed or recombinant proteins are overexpressed in a host cell by introduction of integrating or nonintegrating expression vectors into the host cell. Following introduction of at least one expression vector encoding at least one overexpressed or recombinant protein, the gene product is then overexpressed by inducing expression of an endogenous gene encoding the gene product, or by introducing into the host cell an expression vector encoding the gene product. In another embodiment, cell lines are established which constitutively or inducibly express at least one heterologous chaperone protein. An expression vector encoding the gene product to be overexpressed is introduced into such cell lines to achieve increased secretion of the overexpressed gene product.

The present expression vectors can be replicable in one host cell type, e.g., *Escherichia coli*, and undergo little or no replication in another host cell type, e.g., a eukaryotic host cell, so long as an expression vector permits expression of the overexpressed or recombinant proteins and thereby facilitates secretion of such gene products in a selected host cell type.

Expression vectors as described herein include DNA or RNA molecules that have been engineered for controlled expression of a desired gene, that is, a gene encoding the overexpressed or recombinant proteins. Such vectors also encode nucleic acid molecule

segments which are operably linked to nucleic acid molecules encoding the overexpressed or recombinant proteins. Operably linked in this context means that such segments can effect expression of nucleic acid molecules encoding the overexpressed or recombinant proteins. These nucleic acid sequences include promoters, enhancers, upstream control elements, transcription factors or repressor binding sites, termination signals and other elements which can control gene expression in the contemplated host cell. Preferably the vectors are vectors, bacteriophages, cosmids, or viruses.

Expression vectors of the present invention function in yeast or mammalian cells. Yeast vectors can include the yeast 2μ circle and derivatives thereof, yeast vectors encoding yeast autonomous replication sequences, yeast minichromosomes, any yeast integrating vector and the like. A comprehensive listing of many types of yeast vectors is provided in Parent *et al.* (Yeast 1: 83-138 (1985)).

Elements or nucleic acid regulatory sequences capable of effecting expression of a gene product include promoters, enhancer elements, upstream activating sequences, transcription termination signals and polyadenylation sites. All such promoter and transcriptional regulatory elements, singly or in combination, are contemplated for use in the present expression vectors. Moreover, genetically-engineered and mutated regulatory sequences are also contemplated herein.

Promoters are DNA sequence elements for controlling gene expression. In particular, promoters specify transcription initiation sites and can include a TATA box and upstream promoter elements. The promoters selected are those which would be expected to be operable in the particular host system selected. For example, yeast promoters are used in the present expression vectors when a yeast host cell such as *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, or *Pichia pastoris* is used whereas fungal promoters would be used in host cells such as *Aspergillus niger*, *Neurospora crassa*, or *Tricoderma reesei*. Examples of yeast promoters include but are not limited to the GAPDH, AOX1, GAL1, PGK, GAP, TPI, CYC1, ADH2, PHO5, CUP1, MFα1, PMA1, PDI, TEF, and GUT1 promoters. Romanos *et al.* (Yeast 8: 423-488 (1992)) provide a review of yeast promoters and expression vectors.

The promoters that are operably linked to the nucleic acid molecules disclosed herein can be constitutive promoters or inducible promoters. Inducible promoters, that is. promoters which direct transcription at an increased or decreased rate upon binding of a transcription factor. Transcription factors as used herein include any factor that can bind to a regulatory or control region of a promoter an thereby affect transcription. The synthesis or the promoter binding ability of a transcription factor within the host cell can be controlled by exposing the host to an inducer or removing an inducer from the host cell medium. Accordingly to regulate expression of an inducible promoter, an inducer is added or removed from the growth medium of the host cell. Such inducers can include sugars, phosphate, alcohol, metal ions,

hormones, heat, cold and the like. For example, commonly used inducers in yeast are glucose, galactose, and the like.

Transcription termination sequences that are selected are those that are operable in the particular host cell selected. For example, yeast transcription termination sequences are used in the present expression vectors when a yeast host cell such as *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, or *Pichia pastoris* is used whereas fungal transcription termination sequences would be used in host cells such as *Aspergillus niger*, *Neurospora crassa*, or *Tricoderma reesei*. Transcription termination sequences include but are not limited to the *Saccharomyces cerevisiae CYC* transcription termination sequence (ScCYC TT), the *Pichia pastoris ALG3* transcription termination sequence (ALG3 TT), and *Pichia pastoris PMA1* transcription termination sequence (PpPMA1 TT).

The expression vectors of the present invention can also encode selectable markers. Selectable markers are genetic functions that confer an identifiable trait upon a host cell so that cells transformed with a vector carrying the selectable marker can be distinguished from non-transformed cells. Inclusion of a selectable marker into a vector can also be used to ensure that genetic functions linked to the marker are retained in the host cell population. Such selectable markers can confer any easily identified dominant trait, e.g. drug resistance, the ability to synthesize or metabolize cellular nutrients and the like.

Yeast selectable markers include drug resistance markers and genetic functions which allow the yeast host cell to synthesize essential cellular nutrients, e.g. amino acids. Drug resistance markers which are commonly used in yeast include chloramphenicol, kanamycin, methotrexate, G418 (geneticin), Zeocin, and the like. Genetic functions which allow the yeast host cell to synthesize essential cellular nutrients are used with available yeast strains having auxotrophic mutations in the corresponding genomic function. Common yeast selectable markers provide genetic functions for synthesizing leucine (*LEU2*), tryptophan (*TRP1* and *TRP2*), proline (PRO1), uracil (*URA3*, *URA5*, *URA6*), histidine (*HIS3*), lysine (*LYS2*), adenine (*ADE1* or *ADE2*), and the like. Other yeast selectable markers include the *ARR3* gene from *S. cerevisiae*, which confers arsenite resistance to yeast cells that are grown in the presence of arsenite (Bobrowicz *et al.*, Yeast, 13:819-828 (1997); Wysocki *et al.*, J. Biol. Chem. 272:30061-30066 (1997)). A number of suitable integration sites include those enumerated in U.S. Published application No. 2007/0072262 and include homologs to loci known for *Saccharomyces cerevisiae* and other yeast or fungi. Methods for integrating vectors into yeast are well known, for example, see U.S. Patent No. 7,479,389, WO2007136865, and PCT/US2008/13719. Examples of insertion sites include, but are not limited to, *Pichia ADE* genes; *Pichia TRP* (including *TRP1* through *TRP2*) genes; *Pichia MCA* genes; *Pichia CYM* genes; *Pichia PEP* genes; *Pichia PRB* genes; and *Pichia LEU* genes. The *Pichia ADE1* and *ARG4* genes have been described in Lin Cereghino *et al.*, Gene 263:159-169 (2001) and U.S.

Patent No. 4,818,700, the *HIS3* and *TRP1* genes have been described in Cosano *et al.*, Yeast 14:861-867 (1998), *HIS4* has been described in GenBank Accession No. X56180.

Therefore the present expression vectors can encode selectable markers which are useful for identifying and maintaining vector-containing host cells within a cell population present in culture. In some circumstances selectable markers can also be used to amplify the copy number of the expression vector. After inducing transcription from the present expression vectors to produce an RNA encoding an overexpressed or recombinant protein, the RNA is translated by cellular factors to produce the overexpressed or recombinant protein.

In yeast and other eukaryotes, translation of a messenger RNA (mRNA) is initiated by ribosomal binding to the 5' cap of the mRNA and migration of the ribosome along the mRNA to the first AUG start codon where polypeptide synthesis can begin. Expression in yeast and mammalian cells generally does not require specific number of nucleotides between a ribosomal-binding site and an initiation codon, as is sometimes required in prokaryotic expression systems. However, for expression in a yeast or a mammalian host cell, the first AUG codon in an mRNA is preferably the desired translational start codon.

Moreover, when expression is performed in a yeast host cell the presence of long untranslated leader sequences, e.g. longer than 50-100 nucleotides, can diminish translation of an mRNA. Yeast mRNA leader sequences have an average length of about 50 nucleotides, are rich in adenine, have little secondary structure and almost always use the first AUG for initiation. Since leader sequences which do not have these characteristics can decrease the efficiency of protein translation, yeast leader sequences are preferably used for expression of an overexpressed gene product or a chaperone protein in a yeast host cell. The sequences of many yeast leader sequences are known and are available to the skilled artisan, for example, by reference to Cigan *et al.* (Gene 59: 1-18 (1987)).

In addition to the promoter, the ribosomal-binding site and the position of the start codon, factors which can effect the level of expression obtained include the copy number of a replicable expression vector. The copy number of a vector is generally determined by the vector's origin of replication and any cis-acting control elements associated therewith. For example, an increase in copy number of a yeast episomal vector encoding a regulated centromere can be achieved by inducing transcription from a promoter which is closely juxtaposed to the centromere. Moreover, encoding the yeast FLP function in a yeast vector can also increase the copy number of the vector.

One skilled in the art can also readily design and make expression vectors which include the above-described sequences by combining DNA fragments from available vectors, by synthesizing nucleic acid molecules encoding such regulatory elements or by cloning and placing new regulatory elements into the present vectors. Methods for making expression vectors are

well-known. Overexpressed DNA methods are found in any of the myriad of standard laboratory manuals on genetic engineering.

The expression vectors of the present invention can be made by ligating the overexpressed or recombinant protein coding regions in the proper orientation to the promoter and other sequence elements being used to control gene expression. After construction of the present expression vectors, such vectors are transformed into host cells where the overexpressed gene product and the overexpressed or recombinant protein can be expressed. Methods for transforming yeast and other lower eukaryotic cells with expression vectors are well known and readily available to the skilled artisan. For example, expression vectors can be transformed into yeast cells by any of several procedures including lithium acetate, spheroplast, electroporation, and similar procedures.

Yeast host cells which can be used with yeast replicable expression vectors include any wild type or mutant strain of yeast which is capable of secretion. Such strains can be derived from *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Pichia pastoris*, *Schizosaccharomyces pombe*, *Yarrowia lipolytica*, and related species of yeast. In general, useful mutant strains of yeast include strains which have a genetic deficiency that can be used in combination with a yeast vector encoding a selectable marker. Many types of yeast strains are available from the Yeast Genetics Stock Center (Donner Laboratory, University of California, Berkeley, Calif. 94720), the American Type Culture Collection (12301 Parklawn Drive, Rockville, Md. 20852, hereinafter ATCC), the National Collection of Yeast Cultures (Food Research Institute, Colney Lane, Norwich NR4 7UA, UK) and the Centraalbureau voor Schimmelcultures (Yeast Division, Julianalaan 67a, 2628 BC Delft, Netherlands).

In general, lower eukaryotes such as yeast are useful for expression of glycoproteins because they can be economically cultured, give high yields, and when appropriately modified are capable of suitable glycosylation. Yeast particularly offers established genetics allowing for rapid transformations, tested protein localization strategies and facile gene knock-out techniques. Suitable vectors have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or other glycolytic enzymes, and an origin of replication, termination sequences and the like as desired.

Various yeasts, such as *Kluyveromyces lactis, Pichia pastoris, Pichia methanolica,* and *Hansenula polymorpha* are useful for cell culture because they are able to grow to high cell densities and secrete large quantities of recombinant protein. Likewise, filamentous fungi, such as *Aspergillus niger, Fusarium sp, Neurospora crassa* and others can be used to produce glycoproteins of the invention at an industrial scale.

Lower eukaryotes, particularly yeast, can be genetically modified so that they express glycoproteins in which the glycosylation pattern is human-like or humanized. Such can be achieved by eliminating selected endogenous glycosylation enzymes and/or supplying

exogenous enzymes as described by Gerngross *et al.*, US 20040018590. For example, a host cell can be selected or engineered to be depleted in 1,6-mannosyl transferase activities, which would otherwise add mannose residues onto the *N*-glycan on a glycoprotein.

In one embodiment, the host cell further includes an α1,2-mannosidase catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target the α1,2-mannosidase activity to the ER or Golgi apparatus of the host cell. Passage of a recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $Man_5GlcNAc_2$ glycoform, for example, a recombinant glycoprotein composition comprising predominantly a $Man_5GlcNAc_2$ glycoform. For example, U.S. Patent No, 7,029,872 and U.S. Published Patent Application Nos. 2004/0018590 and 2005/0170452 disclose lower eukaryote host cells capable of producing a glycoprotein comprising a $Man_5GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell further includes a GlcNAc transferase I (GnT I) catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target GlcNAc transferase I activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $GlcNAcMan_5GlcNAc_2$ glycoform, for example a recombinant glycoprotein composition comprising predominantly a $GlcNAcMan_5GlcNAc_2$ glycoform. U.S. Patent No, 7,029,872 and U.S. Published Patent Application Nos. 2004/0018590 and 2005/0170452 disclose lower eukaryote host cells capable of producing a glycoprotein comprising a $GlcNAcMan_5GlcNAc_2$ glycoform. The glycoprotein produced in the above cells can be treated *in vitro* with a hexaminidase to produce a recombinant glycoprotein comprising a $Man_5GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell further includes a mannosidase II catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target mannosidase II activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $GlcNAcMan_3GlcNAc_2$ glycoform, for example a recombinant glycoprotein composition comprising predominantly a $GlcNAcMan_3GlcNAc_2$ glycoform. U.S. Patent No, 7,029,872 and U.S. Published Patent Application No. 2004/0230042 discloses lower eukaryote host cells that express mannosidase II enzymes and are capable of producing glycoproteins having predominantly a $GlcNAc_2Man_3GlcNAc_2$ glycoform. The glycoprotein produced in the above cells can be treated *in vitro* with a hexaminidase to produce a recombinant glycoprotein comprising a $Man_3GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell further includes GlcNAc transferase II (GnT II) catalytic domain fused to a cellular targeting signal peptide not

normally associated with the catalytic domain and selected to target GlcNAc transferase II activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $GlcNAc_2Man_3GlcNAc_2$ glycoform, for example a recombinant glycoprotein composition comprising predominantly a $GlcNAc_2Man_3GlcNAc_2$ glycoform. U.S. Patent No, 7,029,872 and U.S. Published Patent Application Nos. 2004/0018590 and 2005/0170452 disclose lower eukaryote host cells capable of producing a glycoprotein comprising a $GlcNAc_2Man_3GlcNAc_2$ glycoform. The glycoprotein produced in the above cells can be treated *in vitro* with a hexaminidase to produce a recombinant glycoprotein comprising a $Man_3GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell further includes a galactosyltransferase catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target galactosyltransferase activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $GalGlcNAc_2Man_3GlcNAc_2$ or $Gal_2GlcNAc_2Man_3GlcNAc_2$ glycoform, or mixture thereof for example a recombinant glycoprotein composition comprising predominantly a $GalGlcNAc_2Man_3GlcNAc_2$ glycoform or $Gal_2GlcNAc_2Man_3GlcNAc_2$ glycoform or mixture thereof. U.S. Patent No, 7,029,872 and U.S. Published Patent Application No. 2006/0040353 discloses lower eukaryote host cells capable of producing a glycoprotein comprising a $Gal_2GlcNAc_2Man_3GlcNAc_2$ glycoform. The glycoprotein produced in the above cells can be treated *in vitro* with a galactosidase to produce a recombinant glycoprotein comprising a $GlcNAc_2Man_3GlcNAc_2$ glycoform, for example a recombinant glycoprotein composition comprising predominantly a $GlcNAc_2Man_3GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell further includes a sialyltransferase catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target sialytransferase activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising predominantly a $NANA_2Gal_2GlcNAc_2Man_3GlcNAc_2$ glycoform or $NANAGal_2GlcNAc_2Man_3GlcNAc_2$ glycoform or mixture thereof. For lower eukaryote host cells such as yeast and filamentous fungi, it is useful that the host cell further include a means for providing CMP-sialic acid for transfer to the *N*-glycan. U.S. Published Patent Application No. 2005/0260729 discloses a method for genetically engineering lower eukaryotes to have a CMP-sialic acid synthesis pathway and U.S. Published Patent Application No. 2006/0286637 discloses a method for genetically engineering lower eukaryotes to produce sialylated glycoproteins. The glycoprotein produced in the above cells can be treated *in vitro* with a neuraminidase to produce a

recombinant glycoprotein comprising predominantly a $Gal_2GlcNAc_2Man_3GlcNAc_2$ glycoform or $GalGlcNAc_2Man_3GlcNAc_2$ glycoform or mixture thereof.

Any one of the preceding host cells can further include one or more GlcNAc transferase selected from the group consisting of GnT III, GnT IV, GnT V, GnT VI, and GnT IX to produce glycoproteins having bisected (GnT III) and/or multiantennary (GnT IV, V, VI, and IX) *N*-glycan structures such as disclosed in U.S. Published Patent Application Nos. 2004/074458 and 2007/0037248.

In further embodiments, the host cell that produces glycoproteins that have predominantly $GlcNAcMan_5GlcNAc_2$ *N*-glycans further includes a galactosyltransferase catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target Galactosyltransferase activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising predominantly the $GalGlcNAcMan_5GlcNAc_2$ glycoform.

In a further embodiment, the immediately preceding host cell that produced glycoproteins that have predominantly the predominantly the $GalGlcNAcMan_5GlcNAc_2$ *N*-glycans further includes a sialyltransferase catalytic domain fused to a cellular targeting signal peptide not normally associated with the catalytic domain and selected to target sialytransferase activity to the ER or Golgi apparatus of the host cell. Passage of the recombinant glycoprotein through the ER or Golgi apparatus of the host cell produces a recombinant glycoprotein comprising a $NANAGalGlcNAcMan_5GlcNAc_2$ glycoform.

Various of the preceding host cells further include one or more sugar transporters such as UDP-GlcNAc transporters (for example, *Kluyveromyces lactis* and *Mus musculus* UDP-GlcNAc transporters), UDP-galactose transporters (for example, *Drosophila melanogaster* UDP-galactose transporter), and CMP-sialic acid transporter (for example, human sialic acid transporter). Because lower eukaryote host cells such as yeast and filamentous fungi lack the above transporters, it is preferable that lower eukaryote host cells such as yeast and filamentous fungi be genetically engineered to include the above transporters.

In further embodiments of the above host cells, the host cells are further genetically engineered to eliminate glycoproteins having α-mannosidase-resistant *N*-glycans by deleting or disrupting the β-mannosyltransferase gene (*BMT2*)(*See*, U.S. Published Patent Application No. 2006/0211085) and glycoproteins having phosphomannose residues by deleting or disrupting one or both of the phosphomannosyl transferase genes *PNO1* and *MNN4B* (*See* for example, U.S. Patent Nos. 7,198,921 and 7,259,007). In further still embodiments of the above host cells, the host cells are further genetically modified to eliminate *O*-glycosylation of the glycoprotein by deleting or disrupting one or more of the protein *O*-mannosyltransferase (Dol-P-Man:Protein (Ser/Thr) Mannosyl Transferase genes) (*PMT*s) (*See* U.S. Patent No. 5,714,377) or

grown in the presence of i inhibitors such as Pmt-1, Pmti-2, and Pmti-3 as disclosed in Published International Application No. WO 2007061631, or both.

Thus, provided are host cells that have been genetically modified to produce glycoproteins wherein the predominant *N*-glycans thereon include but are not limited to $Man_8GlcNAc_2$, $Man_7GlcNAc_2$, $Man_6GlcNAc_2$, $Man_5GlcNAc_2$, $GlcNAcMan_5GlcNAc_2$, $GalGlcNAcMan_5GlcNAc_2$, $NANAGalGlcNAcMan_5GlcNAc_2$, $Man_3GlcNAc_2$, $GlcNAc(1-4)Man_3GlcNAc_2$, $Gal(1-4)GlcNAc(1-4)Man_3GlcNAc_2$, $NANA(1-4)Gal(1-4)GlcNAc(1-4)Man_3GlcNAc_2$. Further included are host cells that produce glycoproteins that have particular mixtures of the aforementioned *N*-glycans thereon.

In the following examples, heterologous human proteins are expressed in host cells of the species *Pichia pastoris*. These examples demonstrate the invention with respect to specific embodiments of the invention, and are not to be construed as limiting in any manner. The skilled artisan, having read the disclosure and examples herein, will recognize that numerous variants, modifications and improvements to the methods and materials described that are possible without deviating from the practice of the present invention.

EXAMPLE 1

This example shows the construction of a recombinant *Pichia pastoris* that produces recombinant proteins with $Man_5GlcNAc_2$ *N*-glycans.

Construction of expression/integration plasmid vector pGLY642 comprising an expression cassette encoding the human PDI protein and nucleic acid molecules to target the plasmid vector to the *Pichia pastoris PDI1* locus for replacement of the gene encoding the *Pichia pastoris PDI1* with a nucleic acid molecule encoding the human PDI was as follows and is shown in Figure 2. cDNA encoding the human PDI1 was amplified by PCR using the primers hPDI/UP1: 5' AGCGCTGACGCCCCCGAGGAGGAGGACCAC 3' (SEQ ID NO: 1) and hPDI/LP-PacI: 5' CCTTAATTAATTACAGTTCATCATGCACAGCTTTC TGATCAT 3' (SEQ ID NO: 2), Pfu turbo DNA polymerase (Stratagene, La Jolla, CA), and a human liver cDNA (BD Bioscience, San Jose, CA). The PCR conditions were 1 cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 58°C for 30 seconds, and 72°C for 1.5 minutes, and followed by one cycle of 72°C for 10 minutes. The resulting PCR product was cloned into plasmid vector pCR2.1 to make plasmid vector pGLY618. The nucleotide and amino acid sequences of the human PDI1 (SEQ ID NOs:19 and 20, respectively) are shown in Table 9.

The nucleotide and amino acid sequences of the *Pichia pastoris* PDI1 (SEQ ID NOs:21 and 22, respectively) are shown in Table 9. Isolation of nucleic acid molecules comprising the *Pichia pastoris PDI1* 5' and 3' regions was performed by PCR amplification of the regions from *Pichia pastoris* genomic DNA. The 5' region was amplified using primers PB248: 5' ATGAA TTCAG GCCAT ATCGG CCATT GTTTA CTGTG CGCCC ACAGT AG

3' (SEQ ID NO: 3); PB249: 5' ATGTT TAAAC GTGAG GATTA CTGGT GATGA AAGAC 3' (SEQ ID NO: 4). The 3' region was amplified using primers PB250: 5' AGACT AGTCT ATTTG GAGAC ATTGA CGGAT CCAC 3' (SEQ ID NO: 5); PB251: 5' ATCTC GAGAG GCCAT GCAGG CCAAC CACAA GATGA ATCAA ATTTT G-3' (SEQ ID NO: 6). *Pichia pastoris* strain NRRL-Y11430 genomic DNA was used for PCR amplification. The PCR conditions were one cycle of 95°C for two minutes, 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2.5 minutes, and followed by one cycle of 72°C for 10 minutes. The resulting PCR fragments, PpPDI1 (5') and PpPDI1 (3'), were separately cloned into plasmid vector pCR2.1 to make plasmid vectors pGLY620 and pGLY617, respectively. To construct pGLY678, DNA fragments PpARG3-5' and PpARG-3' of integration plasmid vector pGLY24, which targets the plasmid vector to *Pichia pastoris ARG3* locus, were replaced with DNA fragments PpPDI (5') and PpPDI (3'), respectively, which targets the plasmid vector pGLY678 to the *PDI1* locus and disrupts expression of the *PDI1* locus.

The nucleic acid molecule encoding the human PDI was then cloned into plasmid vector pGLY678 to produce plasmid vector pGLY642 in which the nucleic acid molecule encoding the human PDI was placed under the control of the *Pichia pastoris* GAPDH promoter (PpGAPDH). Expression/integration plasmid vector pGLY642 was constructed by ligating a nucleic acid molecule (SEQ ID NO:17) encoding the *Saccharomyces cerevisiae* alpha mating factor pre-signal peptide (ScαMFpre-signal peptide (SEQ ID NO:18) having a *Not*I restriction enzyme site at the 5' end and a blunt 3' end and the expression cassette comprising the nucleic acid molecule encoding the human PDI released from plasmid vector pGLY618 with *Afe*I and *Pac*I to produce a nucleic acid molecule having a blunt 5' end and a *Pac*I site at the 3' end into plasmid vector pGLY678 digested with *Not*I and *Pac*I. The resulting integration/expression plasmid vector pGLY642 comprises an expression cassette encoding a human PDI1/ ScαMFpre-signal peptide fusion protein operably linked to the *Pichia pastoris* promoter and nucleic acid molecule sequences to target the plasmid vector to the *Pichia pastoris PDI1* locus for disruption of the *PDI1* locus and integration of the expression cassette into the *PDI1* locus. Figure 2 illustrates the construction of plasmid vector pGLY642. The nucleotide and amino acid sequences of the ScαMFpre-signal peptide are shown in SEQ ID NOs:17 and 18, respectively.

Construction of expression/integration vector pGLY2232 encoding the human ERO1α protein was as follows and is shown in Figure 3. A nucleic acid molecule encoding the human ERO1α protein was synthesized by GeneArt AG (Regensburg, Germany) and used to construct plasmid vector pGLY2224. The nucleotide and amino acid sequences of the human ERO1α protein (SEQ ID NOs:23 and 24, respectively) are shown in Table 9. The nucleic acid molecule encoding the human ERO1α protein was released from the plasmid vector using restriction enzymes *Afe*I and *Fse*I and then ligated with a nucleic acid molecule encoding the ScαMPpre-signal peptide with 5' *Not*I and 3' blunt ends as above into plasmid vector pGLY2228

digested with *Not*I and *Fse*I. Plasmid vector pGLY2228 also included nucleic acid molecules that included the 5' and 3' regions of the *Pichia pastoris PRB1* gene (PpPRB1-5' and PpPRB1-3' regions, respectively). The resulting plasmid vector, pGLY2230 was digested with *Bgl*II and *Not*I and then ligated with a nucleic acid molecule containing the *Pichia pastoris PDI1* promoter (PpPDI promoter) which had been obtained from plasmid vector pGLY2187 digested with *B*glII and *N*otI. The nucleotide sequence of the PpPDI promoter is 5'-AACACGAACACTGTAAA TAGAATAAAAGAAAACTTGGATAGTAGAACTTCAATGTAGTGTTTCTATTGTCTTAC GCGGCTCTTTAGATTGCAATCCCCAGAATGGAATCGTCCATCTTTCTCAACCCACTC AAAGATAATCTACCAGACATACCTACGCCCTCCATCCCAGCACCACGTCGCGATCAC CCCTAAAACTTCAATAATTGAACACGTACTGATTTCCAAACCTTCTTCTTCTTCCTAT CTATAAGA-3' (SEQ ID NO:31). The resulting plasmid vector, pGLY2232, is an expression/integration vector that contains an expression cassette that encodes the human ERO1α fusion protein under control of the *Pichia pastoris PDI1* promoter and includes the 5' and 3' regions of the *Pichia pastoris PRB1* gene to target the plasmid vector to the *PRB1* locus of genome for disruption of the *PRB1* locus and integration of the expression cassette into the *PRB1* locus. Figure 3 illustrates the construction of plasmid vector pGLY2232.

Construction of expression/integration vector pGLY2233 encoding the human GRP94 protein was as follows and is shown in Figure 4. The human GRP94 was PCR amplified from human liver cDNA (BD Bioscience) with the primers hGRP94/UP1: 5'-AGCGC TGACG ATGAA GTTGA TGTGG ATGGT ACAGT AG-3'; (SEQ ID NO: 15); and hGRP94/LP1: 5'-GGCCG GCCTT ACAAT TCATC ATGTT CAGCT GTAGA TTC 3'; (SEQ ID NO: 16). The PCR conditions were one cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 55°C for 20 seconds, and 72°C for 2.5 minutes, and followed by one cycle of 72°C for 10 minutes. The PCR product was cloned into plasmid vector pCR2.1 to make plasmid vector pGLY2216. The nucleotide and amino acid sequences of the human GRP94 (SEQ ID NOs:25 and 26, respectively) are shown in Table 9.

The nucleic acid molecule encoding the human GRP94 was released from plasmid vector pGLY2216 with *Afe*I and *Fse*I. The nucleic acid molecule was then ligated to a nucleic acid molecule encoding the ScαMPpre-signal peptide having *Not*I and blunt ends as above and plasmid vector pGLY2231 digested with *Not*I and *Fse*I carrying nucleic acid molecules comprising the *Pichia pastoris PEP4* 5' and 3' regions (PpPEP4-5' and PpPEP4-3' regions, respectively) to make plasmid vector pGLY2229. Plasmid vector pGLY2229 was digested with BglII and NotI and a DNA fragment containing the PpPDI1 promoter was removed from plasmid vector pGLY2187 with BglII and NotI and the DNA fragment ligated into pGLY2229 to make plasmid vector pGLY2233. Plasmid vector pGLY2233 encodes the human GRP94 fusion protein under control of the *Pichia pastoris* PDI promoter and includes the 5' and 3' regions of the *Pichia pastoris PEP4* gene to target the plasmid vector to the *PEP4* locus of genome for

disruption of the *PEP4* locus and integration of the expression cassette into the *PEP4* locus. Figure 4 illustrates the construction of plasmid vector pGLY2233.

Construction of plasmid vectors pGLY1162, pGLY1896, and pGFI207t was as follows. All *Trichoderma reesei* α-1,2-mannosidase expression plasmid vectors were derived from pGFI165, which encodes the *T. reesei* α-1,2-mannosidase catalytic domain (*See* published International Application No. WO2007061631) fused to *S. cerevisiae* αMATpre signal peptide herein expression is under the control of the *Pichia pastoris* GAP promoter and wherein integration of the plasmid vectors is targeted to the *Pichia pastoris PRO1* locus and selection is using the *Pichia pastoris URA5* gene. A map of plasmid vector pGFI165 is shown in Figure 5.

Plasmid vector pGLY1162 was made by replacing the GAP promoter in pGFI165 with the *Pichia pastoris* AOX1 (PpAOX1) promoter. This was accomplished by isolating the PpAOX1 promoter as an *Eco*RI (made blunt)-*Bgl*II fragment from pGLY2028, and inserting into pGFI165 that was digested with NotI (made blunt) and BglII. Integration of the plasmid vector is to the *Pichia pastoris PRO1* locus and selection is using the *Pichia pastoris URA5* gene. A map of plasmid vector pGLY1162 is shown in Figure 6.

Plasmid vector pGLY1896 contains an expression cassette encoding the mouse α-1,2-mannosidase catalytic domain fused to the *S. cerevisiae MNN2* membrane insertion leader peptide fusion protein (*See* Choi *et al.*, Proc. Natl. Acad. Sci. USA 100: 5022 (2003)) inserted into plasmid vector pGFI165 (Figure 5). This was accomplished by isolating the GAPp-ScMNN2-mouse MNSI expression cassette from pGLY1433 digested with *Xho*I (and the ends made blunt) and *Pme*I, and inserting the fragment into pGFI165 that digested with *Pme*I. Integration of the plasmid vector is to the *Pichia pastoris PRO1* locus and selection is using the *Pichia pastoris URA5* gene. A map of plasmid vector pGLY1896 is shown in Figure 5.

Plasmid vector pGFI207t is similar to pGLY1896 except that the *URA5* selection marker was replaced with the *S. cerevisiae ARR3* (ScARR3) gene, which confers resistance to arsenite. This was accomplished by isolating the ScARR3 gene from pGFI166 digested with *Asc*I and the *Asc*I ends made blunt) and *Bgl*II, and inserting the fragment into pGLY1896 that digested with *Spe*I and the *Spe*I ends made blunt and *Bgl*II. Integration of the plasmid vector is to the *Pichia pastoris PRO1* locus and selection is using the *Saccharomyces cerevisiae ARR3* gene. A map of plasmid vector pGFI207t is shown in Figure 5.

Construction of anti-DKK1 antibody expression/integration plasmid vector pGLY2260 and pGLY2261 (Figure 7) was as follows. Anti-DKK1 antibodies are antibodies that recognize Dickkopf protein 1, a ligand involved in the Wnt signaling pathway. To generate expression/integration plasmid vectors pGLY2260 and pGLY2261 encoding an anti-DKK1 antibody, codon-optimized nucleic acid molecules encoding heavy chain (HC; fusion protein containing VH + IgG_2m4) and light chain (LC; fusion protein containing VL + Lλ constant region) fusion proteins, each in frame with a nucleic acid molecule encoding an α-amylase (from

Aspergillus niger) signal peptide were synthesized by GeneArt AG. The nucleotide and amino acid sequences for the a-amylase signal peptide are shown in SEQ ID NOs:48 and 49. The nucleotide sequence of the HC is shown in SEQ ID NO:27 and the amino acid sequence is shown in SEQ ID NO:28. The nucleotide sequence of the LC is shown in SEQ ID NO:29 and the amino acid sequence is shown in SEQ ID NO:30. The IgG_2m4 isotype has been disclosed in U.S. Published Application No. 2007/0148167 and U.S. Published Application No. 2006/0228349. The nucleic acid molecules encoding the HC and LC fusion proteins were separately cloned using unique 5'-*EcoR*I and 3'-*Fse*I sites into expression plasmid vector pGLY1508 to form plasmid vectors pGLY1278 and pGLY1274, respectively. These plasmid vectors contained the Zeocin-resistance marker and *TRP2* integration sites and the *Pichia pastoris* AOX1 promoter operably linked to the nucleic acid molecules encoding the HC and LC fusion proteins. The LC fusion protein expression cassette was removed from pGLY1274 with *Bgl*II and *Bam*H1 and cloned into pGLY1278 digested with *Bgl*II to generate plasmid vector pGLY2260, which encodes the HC and LC fusion proteins and targets the expression cassettes to the *TRP2* locus for integration of the expression cassettes into the *TRP2* locus. The plasmid vector pGLY2261 contains an additional LC in plasmid vector pGLY2260. (Figure 7).

Yeast transformations with the above expression/integration vectors were as follows. *Pichia pastoris* strains were grown in 50 mL YPD media (yeast extract (1%), peptone (2%), dextrose (2%)) overnight to an OD of between about 0.2 to 6. After incubation on ice for 30 minutes, cells were pelleted by centrifugation at 2500-3000 rpm for 5minutes. Media was removed and the cells washed three times with ice cold sterile 1 M sorbitol before resuspension in 0.5 ml ice cold sterile 1M sorbitol. Ten μL linearized DNA (5-20 μg) and 100 μL cell suspension was combined in an electroporation cuvette and incubated for 5 minutes on ice. Electroporation was in a Bio-Rad GenePulser Xcell following the preset *Pichia pastoris* protocol (2 kV, 25 μF, 200 Ω), immediately followed by the addition of 1 mL YPDS recovery media (YPD media plus 1 M sorbitol). The transformed cells were allowed to recover for four hours to overnight at room temperature (24°C) before plating the cells on selective media.

Generation of Cell Lines was as follows and is shown in Figure 1A and 1B. The strain **yGLY24-1** (*ura5*Δ::*MET1 och1*Δ::*lacZ bmt2*Δ::*lacZ*/*KlMNN2-2* / *mnn4L1*Δ::*lacZ*/ *MmSLC35A3 pno1*Δ*mnn4*Δ::*lacZ met16*Δ::*lacZ*), was constructed using methods described earlier (*See* for example, Nett and Gerngross, Yeast 20:1279 (2003); Choi *et al.*, Proc. Natl. Acad. Sci. USA 100:5022 (2003); Hamilton *et al.*, Science 301:1244 (2003)). The *BMT2* gene has been disclosed in Mille *et al.*, J. Biol. Chem. 283: 9724-9736 (2008) and U.S. Published Application No. 20060211085. The *PNO1* gene has been disclosed in U.S. Patent No. 7,198,921 and the *mnn4L1* gene (also referred to as *mnn4b*) has been disclosed in U.S. Patent No. 7,259,007. The *mnn4* refers to *mnn4L2* or *mnn4a*. In the genotype, KlMNN2-2 is the *Kluveromyces lactis* GlcNAc transporter and MmSLC35A3 is the *Mus musculus* GlcNAc

transporter. The *URA5* deletion renders the yGLY24-1 strain auxotrophic for uracil (*See* U.S. Published application No. 2004/0229306) and was used to construct the humanized chaperone strains that follow. While the various expression cassettes were integrated into particular loci of the Pichia pastoris genome in the examples herein, it is understood that the operation of the invention is independent of the loci used for integration. Loci other than those disclosed herein can be used for integration of the expression cassettes. Suitable integration sites include those enumerated in U.S. Published application No. 20070072262 and include homologs to loci known for *Saccharomyces cerevisiae* and other yeast or fungi.

Control strain **yGLY645** (PpPDI1) was constructed. Strain yGLY645 expresses both a *Trichoderma Reesei* mannosidase1 (TrMNS1) and a mouse mannosidase IA (MuMNS1A), each constitutively expressed under the control of a PpGAPDH promoter, with the native *Pichia pastoris PDI1* locus intact. Strain yGLY645 was generated from strain yGLY24-1 by transforming yGLY24-1 with plasmid vector pGLY1896, which targeted the plasmid vector to the Proline 1 (*PRO1*) locus in the *Pichia* genome. Plasmid vector pGLY1896 contains expression cassettes encoding the *Trichoderma Reesei* mannosidase 1 (TrMNS1) and the mouse mannosidase IA (FB53, MuMNS1A), each constitutively expressed under the control of a PpGAPDH promoter.

Strains **yGLY702** and **yGLY704** were generated in order to test the effectiveness of the human PDI expressed in *Pichia pastoris* cells in the absence of the endogenous *Pichia pastoris* PDI1 gene. Strains yGLY702 and yGLY704 (hPDI) were constructed as follows. Strain yGLY702 was generated by transforming yGLY24-1 with plasmid vector pGLY642 containing the expression cassette encoding the human PDI under control of the constitutive PpGAPDH promoter. Plasmid vector pGLY642 also contained an expression cassette encoding the *Pichia pastoris URA5*, which rendered strain yGLY702 prototrophic for uracil. The *URA5* expression cassette was removed by counterselecting yGLY702 on 5-FOA plates to produce strain yGLY704 in which, so that the *Pichia pastoris* PDI1 gene has been stably replaced by the human PDI gene and the strain is auxotrophic for uracil.

The replacement of the *Pichia pastoris PDI1* with the human PDI using plasmid vector pGLY642 was confirmed by colony PCR using the following primers specific to only the PpPDI1 ORF; PpPDI/UPi-1, 5'-GGTGA GGTTG AGGTC CCAAG TGACT ATCAA GGTC-3'; (SEQ ID NO: 7); PpPDI/LPi-1, 5'-GACCT TGATA GTCAC TTGGG ACCTC AACCT CACC-3'; (SEQ ID NO: 8); PpPDI/UPi-2, 5' CGCCA ATGAT GAGGA TGCCT CTTCA AAGGT TGTG-3'; (SEQ ID NO: 9); and PpPDI/LPi-2, 5'-CACAA CCTTT GAAGA GGCAT CCTCA TCATT GGCG-3'; (SEQ ID NO: 10). Thus, the absence of PCR product indicates the knockout of PpPDI1. The PCR conditions were one cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 58°C for 20 seconds, and 72°C for one minute, and followed by one cycle of 72°C for 10 minutes.

Additional PCR was used to confirm the double crossover of pGLY642 at the PpPDI1 locus using PCR primers; PpPDI-5'/UP, 5'-GGCGA TTGCA TTCGC GACTG TATC-3'; (SEQ ID NO: 11); and, hPDI-3'/LP 5'-CCTAG AGAGC GGTGG CCAAG ATG-3'; (SEQ ID NO: 12). PpPDI-5'/UP primes the upstream region of PpPDI1 that is absent in PpPDI1 (5') of pGY642 and hPDI-3'/LP primes human PDI ORF in pGLY642. The PCR conditions were one cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 50°C for 30 seconds, and 72°C for 2.5 minutes, and followed by one cycle of 72°C for 10 minutes.

The integration efficiency of a plasmid vector as a knockout (i.e., a double cross-over event) or as a 'roll-in' (i.e., a single integration of the plasmid vector into the genome, can be dependent upon a number of factors, including the number and length of homologous regions between vectors and the corresponding genes on host chromosomal DNA, selection markers, the role of the gene of interest, and the ability of the knocked-in gene to complement the endogenous function. The inventors found that in some instances pGLY642 was integrated as a double cross-over, resulting in replacement of the endogenous PpPDI gene with human PpPDI, while in other cases, the pGLY642 plasmid vector was integrated as a single integration, resulting in presence of both the endogenous PpPDI1 gene and a human PpPDI gene. In order to distinguish between these events, the inventors utilized PCR primers of Sequence ID Nos. 11 through 14, described herein. If the PpPDI gene has been retained after integration of the pGLY642 plasmid vector, PpPDI-5'/UP and hPDI-3'/LP, directed to the internal PpPDI coding sequence, will result in an amplification product and a corresponding band. In the event of a knockout or double cross-over, these primers will not result in any amplification product and no corresponding band will be visible.

The roll-in of pGLY642 was confirmed with the primers; PpPDI/UPi-1 (SEQ ID NO: 7) and PpPDI/LPi-1 (SEQ ID NO: 8) encoding PpPDI1, and hPDI/UP, 5'-GTGGC CACAC CAGGG GGCAT GGAAC-3'; (SEQ ID NO: 13); and hPDI-3'/LP, 5'-CCTAG AGAGC GGTGG CCAAG ATG-3'; (SEQ ID NO: 14); encoding human PDI. The PCR conditions were one cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 58°C for 20 seconds, and 72°C for one minute, and followed by 1 cycle of 72°C for 10 minutes for PpPDI1, and 1 cycle of 95°C for two minutes, 25 cycles of 95°C for 20 seconds, 50°C for 30 seconds, and 72°C for 2.5 minutes, and followed by one cycle of 72°C for 10 minutes for human PDI.

Strain **yGLY733** was generated by transforming with plasmid vector pGLY1162, which comprises an expression cassette that encodes the *Trichoderma Reesei* mannosidase (TrMNS1) operably linked to the *Pichia pastoris* AOX1 promoter (PpAOX1-TrMNS1), into the *PRO1* locus of yGLY704. This strain has the gene encoding the *Pichia pastoris* PDI1 replaced with the expression cassette encoding the human *PDI*, has the PpAOX1-TrMNS1 expression cassette integrated into the *PRO1* locus, and is a *URA5* prototroph. The PpAOX1 promoter allows overexpression when the cells are grown in the presence of methanol.

Strain **yGLY762** was constructed by integrating expression cassettes encoding TrMNS1 and mouse mannosidase IA (MuMNS1A), each operably linked to the *Pichia pastoris* GAPDH promoter in plasmid vector pGFI207t into strain yGLY733 at the 5' *PRO1* locus UTR in *Pichia pastoris* genome. This strain has the gene encoding the *Pichia pastoris* PDI1 replaced with the expression cassette encoding the human PDI, has the PpGAPDH-TrMNS1 and PpGAPDH-MuMNS1A expression cassettes integrated into the *PRO1* locus, and is a *URA5* prototroph.

Strain **yGLY2263** was generated by transforming strain yGLY645 with integration/expression plasmid pGLY2260, which targets an expression cassette encoding the anti-DKK1 antibody to the *TRP2* locus.

Strain **yGLY2674** was generated by counterselecting yGLY733 on 5-FOA plates. This strain has the gene encoding the *Pichia pastoris* PD1 replaced with the expression cassette encoding the human PDI, has the PpAOX1-TrMNS1 expression cassette integrated into the *PRO1* locus, and is a *URA5* auxotroph.

Strain **yGLY2677** was generated by counterselecting yGLY762 on 5-FOA plates. This strain has the gene encoding the *Pichia pastoris* PDI1 replaced with the expression cassette encoding the human PDI, has the PpAOX1-TrMNS1 expression cassette integrated into the *PRO1* locus, has the PpGAPH-TrMNS1 and PpGAPDH-MuMNS1A expression cassettes integrated into the *PRO1* locus, and is a *URA5* auxotroph.

Strains **yGLY2690** was generated by integrating plasmid vector pGLY2232, which encodes the human ERO1α protein, into the *PRB1* locus. This strain has the gene encoding the *Pichia pastoris* PDI1 replaced with the expression cassette encoding the human PDI, has the PpAOX1-TrMNS1 expression cassette integrated into the *PRO1* locus, the human ERO1α expression cassette integrated into the PRB1 locus, and is a *URA5* prototroph.

Strains **yGLY2696** was generated by integrating plasmid vector pGLY2233, which encodes the human GRP94 protein, into the PEP4 locus. This strain has the gene encoding the *Pichia pastoris* PDI1 replaced with the expression cassette encoding the human PDI, has the PpAOX1-TrMNS1 expression cassette integrated into the *PRO1* locus, has the PpGAPDH-TrMNS1 and PpGAPDH-MuMNS1A expression cassettes integrated into the *PRO1* locus, has the human GRP94 integrated into the *PEP4* locus, and is a *URA5* prototroph.

Strain **yGLY3628** was generated by transforming strain yGLY2696 with integration/expression plasmid pGLY2261, which targets an expression cassette encoding the anti-DKK1 antibody to the *TRP2* locus.

Strain **yGLY3647** was generated by transforming strain yGLY2690 with integration/expression plasmid pGLY2261, which targets an expression cassette encoding the anti-DKK1 antibody to the *TRP2* locus.

Table 1 shows that replacing the gene encoding the *Pichia pastoris* PDI1 with an expression cassette encoding the human PDI in yeast genetically engineered to produce glycoproteins that have predominantly $Man_5GlcNAc_2$ *N*-glycans effects a reduction in *O*-glycosylation occupancy and an increase in *N*-glycosylation.

Table 1

GS2.0 Strain	yGLY2263 (control)	yGLY3647	yGLY3628
Pichia pastoris PDI1	Wild-type	Knockout	Knockout
Human PDI	None	Overexpressed	Overexpressed
Human ERO1α	None	Expressed	None
Human GRP94	None	None	Expressed
Pichia pastoris PRB1	Intact	Knockout	Intact
Pichia pastoris PEP4	Intact	Intact	Knockout
O-glycan (Occupancy: H2L2)	23.7	9.2	10.0

EXAMPLE 2

Cell Growth conditions of the transformed strains for antibody production was generally as follows.

Protein expression for the transformed yeast strains was carried out at in shake flasks at 24°C with buffered glycerol-complex medium (BMGY) consisting of 1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer pH 6.0, 1.34% yeast nitrogen base, 4 x 10^{-5} % biotin, and 1% glycerol. The induction medium for protein expression was buffered methanol-complex medium (BMMY) consisting of 1% methanol instead of glycerol in BMGY. Pmt inhibitor Pmti-3 in methanol was added to the growth medium to a final concentration of 18.3μM at the time the induction medium was added. Cells were harvested and centrifuged at 2,000 rpm for five minutes.

SixFors Fermentor Screening Protocol followed the parameters shown in Table 2.

Table 2
SixFors Fermentor Parameters

Parameter	Set-point	Actuated Element
pH	6.5 ± 0.1	30% NH_4OH
Temperature	24 ± 0.1	Cooling Water & Heating Blanket

Dissolved O2	n/a	Initial impeller speed of 550 rpm is ramped to 1200 rpm over first 10 hr, then fixed at 1200 rpm for remainder of run

At time of about 18 hours post-inoculation, SixFors vessels containing 350 mL media A (See Table 6 below) plus 4% glycerol were inoculated with strain of interest. A small dose (0.3 mL of 0.2 mg/mL in 100% methanol) of Pmti-3 (5-[[3-(l-Phenyl-2-hydroxy)ethoxy)-4-(2- phenylethoxy)]phenyl]methylene]-4-oxo-2-thioxo-3-thiazolidineacetic Acid) (*See* Published International Application No. WO 2007061631) was added with inoculum. At time about 20 hour, a bolus of 17 mL 50% glycerol solution (Glycerol Fed-Batch Feed, *See* Table 7 below) plus a larger dose (0.3 mL of 4 mg/mL) of Pmti-3 was added per vessel. At about 26 hours, when the glycerol was consumed, as indicated by a positive spike in the dissolved oxygen (DO) concentration, a methanol feed (See Table 6 below) was initiated at 0.7 mL/hr continuously. At the same time, another dose of Pmti-3 (0.3 mL of 4 mg/mL stock) was added per vessel. At time about 48 hours, another dose (0.3 mL of 4 mg/mL) of Pmti-3 was added per vessel. Cultures were harvested and processed at time about 60 hours post-inoculation.

Table 3
Composition of Media A

Martone L-1	20	g/L
Yeast Extract	10	g/L
KH_2PO4	11.9	g/L
K_2HPO_4	2.3	g/L
Sorbitol	18.2	g/L
Glycerol	40	g/L
Antifoam Sigma 204	8	drops/L
10X YNB w/Ammonium Sulfate w/o Amino Acids (134 g/L)	100	mL/L
250X Biotin (0.4 g/L)	10	mL/L
500X Chloramphenicol (50 g/L)	2	mL/L
500X Kanamycin (50 g/L)	2	mL/L

Table 4
Glycerol Fed-Batch Feed

Glycerol	50	% m/m
PTM1 Salts (see Table IV-E below)	12.5	mL/L
250X Biotin (0.4 g/L)	12.5	mL/L

Table 5 Methanol Feed		
Methanol	100	% m/m
PTM1 Salts	12.5	mL/L
250X Biotin (0.4 g/L)	12.5	mL/L

Table 6 PTM1 Salts		
CuSO4-5H2O	6	g/L
NaI	80	mg/L
MnSO4-7H2O	3	g/L
NaMoO4-2H2O	200	mg/L
H3BO3	20	mg/L
CoCl2-6H2O	500	mg/L
ZnCl2	20	g/L
FeSO4-7H2O	65	g/L
Biotin	200	mg/L
H2SO4 (98%)	5	mL/L

O-glycan determination was performed using a Dionex-HPLC (HPAEC-PAD) as follows. To measure *O*-glycosylation reduction, protein was purified from the growth medium using protein A chromatography (Li et al. Nat. Biotechnol. 24(2):210-5 (2006)) and the *O*-glycans released from and separated from protein by alkaline elimination (beta -elimination) (Harvey, Mass Spectrometry Reviews 18: 349- 451 (1999)). This process also reduces the newly formed reducing terminus of the released *O*-glycan (either oligomannose or mannose) to mannitol. The mannitol group thus serves as a unique indicator of each *O*- glycan. 0.5 nmole or more of protein, contained within a volume of 100 µL PBS buffer, was required for beta elimination. The sample was treated with 25 µL alkaline borohydride reagent and incubated at 50°C for 16 hours. About 20 µL arabitol internal standard was added, followed by 10 µL glacial acetic acid. The sample was then centrifuged through a Millipore filter containing both SEPABEADS and AG 50W-X8 resin and washed with water. The samples, including wash, were transferred to plastic autosampler vials and evaporated to dryness in a centrifugal evaporator. 150 µL 1% AcOH/MeOH was added to the samples and the samples evaporated to dryness in a centrifugal evaporator. This last step was repeated five more times. 200 µL of water was added and 100 µL of the sample was analyzed by high pH anion-exchange chromatography coupled with pulsed

electrochemical detection-Dionex HPLC (HPAEC-PAD). Average *O*-glycan occupancy was determined based upon the amount of mannitol recovered.

EXAMPLE 3

This example demonstrates that occupancy of *O*-glycans in proteins produced in the above strains expressing the human PDI in place of the *Pichia pastoris* PDI1 can be significantly reduced when either the *Pichia pastoris* Golgi Ca^{2+} ATPase (PpPMR1) or the *Arabidopsis thaliana* ER Ca^{2+} ATPase (AtECA1) is overexpressed in the strains. In this example, the effect is illustrated using glycoengineered *Pichia pastoris* strains that produce antibodies having predominantly $Man_5GlcNAc_2$ N-glycans.

An expression cassette encoding the PpPMR1 gene was constructed as follows. The open reading frame of *P. pastoris* Golgi Ca^{2+} ATPase (PpPMR1) was PCR amplified from *P. pastoris* NRRL11430 genomic DNA using the primers (PpPMR1/UP: 5'-GAATTCATGACAGCTAATGAAAATCCTTTTGAGAATGAG-3' (SEQ ID NO:36) and PpPMR1/LP: 5'-GGCCGGCCTCAAACAGCCATGCTGTATCCATTGTATG -3' (SEQ ID NO:37). The PCR conditions were one cycle of 95°C for two minutes; five cycles of 95°C for 10 seconds, 52°C for 20 seconds, and 72°C for 3 minutes; 20 cycles of 95°C for 10 seconds, 55°C for 20 seconds, and 72°C for 3 minutes; followed by 1 cycle of 72°C for 10 minutes. The resulting PCR product was cloned into pCR2.1 and designated pGLY3811. PpPMR1 was removed from pGLY3811 by digesting with plasmid with *Pst*I and *Fse*I and the *Pst*I end had been made blunt with T4 DNA polymerase prior to digestion with *Fse*I. The DNA fragment encoding the PpPMR1 was cloned into pGFI30t digested with *Eco*RI with the ends made blunt with T4 DNA polymerase and *Fse*I to generate pGLY3822 in which the PpPMR1 is operably linked to the AOX1 promoter. Plasmid pGLY3822 targets the *Pichia pastoris URA6* locus. Plasmid pGLY3822 is shown in Figure 8. The DNA sequence of PpPMR1 is set forth in SEQ ID NO:32 and the amino acid sequence of the PpPMR1 is shown in SEQ ID NO:33.

An expression cassette encoding the *Arabidopsis thaliana* ER Ca^{2+} ATPase (AtECA1) was constructed as follows. A DNA encoding AtECA1 was synthesized from GeneArt AG (Regensburg, Germany) and cloned to make pGLY3306. The synthesized AtECA1 was removed from pGLY3306 by digesting with *Mly*I and *Fse*I and cloning the DNA fragment encoding the AtECA1 into pGFI30t digested with *Eco*RI with the ends made blunt with T4 DNA polymerase and *Fse*I to generate integration/expression plasmid pGLY3827. Plasmid pGLY3827 targets the *Pichia pastoris URA6* locus. Plasmid pGLY3827 is shown in Figure 9. The DNA sequence of the AtECA1 was codon-optimized for expression in *Pichia pastoris* and is shown in SEQ ID NO:34. The encoded AtECA1 has the amino acid sequence set forth in SEQ ID NO:35.

Integration/expression plasmid pGLY3822 (contains expression cassette encoding PpPMR1) or pGLY3827 (contains expression cassette encoding AtECA1) was linearized with *Spe*I and transformed into *Pichia pastoris* strain **yGLY3647** or **yGLY3693** at the *URA6* locus. The genomic integration of pGLY3822 or pGLY3827 at *URA6* locus was confirmed by colony PCR (cPCR) using primers, 5'AOX1 (5'-GCGACTGGTTCCAATTGACAAGCTT-3' (SEQ ID NO:38) and PpPMR1/cLP (5'-GGTTGCTCTCGTCGATACTCAAGTGGGAAG-3' (SEQ ID NO:39) for confirming PpPMR1 integration into the URA6 locus , and 5'AOX1 and AtECA1/cLP (5'-GTCGGCTGGAACCTTATCACCAACTCTCAG-3' (SEQ ID NO:40) for confirming integration of AtECA1 into the *URA6* locus. The PCR conditions were one cycle of 95°C for 2 minutes, 25 cycles of 95°C for 10 seconds, 55°C for 20 seconds, and 72°C for one minute; followed by one cycle of 72°C for 10 minutes.

Strain **yGLY8238** was generated by transforming strain yGLY3647 with integration/expression plasmid pGLY3822 encoding the PpPMR1 and targeting the *URA6* locus. In strain yGLY3647, the *Pichia pastoris* PDI1 chaperone gene has been replaced with the human PD1 gene as described in Example 1 and shown in Figures 1A and 1B.

Strain **yGLY8240** was generated by transforming strain yGLY3647 with plasmid pGLY3827 encoding the AtECA1 and targeting the *URA6* locus. The genealogy of the strains is shown in Figures 1A and 1B.

The strains were evaluated for the effect the addition of PpPMR1 or AtECA1 to the humanized chaperone strains might have on reducing *O*-glycosylation of the antibodies produced by the strains. As shown in Table 7 the addition of either PpPMR1 or AtECA1 into strain yGLY3647 effected a significant reduction in *O*-glycosylation occupancy compared to strain yGLY3647 expressing the human PDI in place of the *Pichia pastoris* PDI1 or strain yGLY2263 expressing only the endogenous PDI1 but capable of making antibodies with a $Man_5GlcNAc_2$ glycoform as strain yGLY3647. The results also suggest that yeast strains that express its endogenous PDI1 and not the human PDI1 and overexpress a Ca^{2+} ATPase will produce glycoproteins with reduced *O*-glycan occupancy.

Table 7

Strain	yGLY2263 (control)	yGLY3647	yGLY3647 + Ca^{2+} ATPase	
			yGLY8240	yGLY8238
			AtECA1	PpPMR1
O-glycan occupancy (H2+L2: anti-DKK1)	23.7	9.2	5.5	6.2
O-glycan occupancy was determined by Mannitol assay.				

EXAMPLE 4

A DNA fragment encoding the human calreticulin (hCRT) without its native signal sequence was PCR amplified from a human liver cDNA library (BD Biosciences, San Jose, CA) using primers hCRT-BstZ17I-HA/UP: 5'-GTATACCCATACGACGTCCCAGACTACGCTGAGCCCGCCGTCTACTTCAAGGAGC-3' (SEQ ID NO:45) and hCRT-PacI/LP: 5'-TTAATTAACTACAGCTCGTCATGGGCCTGGCCG GGGACATCTTCC-3' (SEQ ID NO:46). The PCR conditions were one cycle of 98°C for two min; 30 cycles of 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for two minutes, and followed by one cycle of 72°C for 10 minutes. The resulting PCR product was cloned into pCR2.1 Topo vector to make pGLY1224. The DNA encoding the hCRT further included modifications such that the encoded truncated hCRT has an HA tag at its N-terminus and HDEL at its C-terminus. The DNA encoding the hCRT was released from pGLY1224 by digestion with *Bst*Z17I and *Pac*I and the DNA fragment cloned into an expression vector pGLY579, which had been digested with *Not*I and *Pac*I, along with a DNA fragment encoding the *S. cerevisiae* alpha-mating factor pre signal sequence having *Not*I and *Pac*I compatible ends to create pGLY1230. This plasmid is an integration/expression plasmid that encodes the hCRT with the *S. cerevisiae* alpha-mating factor pre signal sequence and HA tag at the N-terminus and an HDEL sequence at its C-terminus operably linked to the *Pichia pastoris* GAPDH promoter and targeting the HIS3 locus of *Pichia pastoris*.

A DNA fragment encoding the human ERp57 (hERp57) was synthesized by GeneArt AG having *Not*I and *Pac*I compatible ends. The DNA fragment was then cloned into pGLY129 digested with *Not*I and *Pac*I to produce pGLY1231. This plasmid encodes the hERp57 operably linked to the *Pichia pastoris* PMA1 promoter.

Plasmid pGLY1231 was digested with *Swa*I and the DNA fragment encoding the hERp57 was cloned into plasmid pGLY1230 digested with *Pme*I. Thus, integration/expression plasmid pGLY1234 encodes both the hCRT and hERp57. Plasmid pGLY1234 is shown in Figure 10.

Strain **yGLY3642** was generated by counterselecting strain yGLY2690 in the presence of 5'FOA, a *URA5* auxotroph.

Strain **yGLY3668** was generated by transforming yGLY3642 with integration/expression plasmid pGLY1234 encoding the hCRT and hERp57 and which targets the *HIS3* locus.

Strain **yGLY3693** was generated by transforming strain yGLY3668 with integration/expression plasmid pGLY2261, which targets an expression cassette encoding the anti-DKK1 antibody to the *TRP2* locus.

Strain **yGLY8239** was generated by transforming strain yGLY3693 with integration/expression plasmid pGLY3822 encoding the PpPMR1 and targeting the *URA6* locus.

Strain **yGLY8241** was generated by transforming strain yGLY3693 with integration/expression plasmid pGLY3827 encoding the AtECA1 and targeting the *URA6* locus.

The genealogy of the strains described in this example are shown in Figures 1A and 1B.

The above strains were evaluated to see whether the addition of hCRT and hERp57 to the humanized chaperone strains expressing PpPMR1 or AtECA1 of the previous example might effect a further reduction in *O*-glycan occupancy of the antibodies produced. As shown in Table 8, in strain yGLY3693 expressing hCRT and hERp57 alone, there was about a 2-fold decrease in *O*-glycan occupancy, which was further decreased up to a 4-fold in strains that further expressed PpPMR1 or AtECA1. The results also suggest that yeast strains that express its endogenous PDI1 and overexpress a Ca^{2+} ATPase will produce glycoproteins with reduced *O*-glycan occupancy.

Table 8

Strain	yGLY2263 (control)	yGLY3693	yGLY3693 + Ca^{2+} ATPase	
			yGLY8241	yGLY8239
			AtECA1	PpPMR1
O-glycan occupancy (H2+L2: anti-DKK1)	23.7	10.4	5.5	7.8
O-glycan occupancy was determined by Mannitol assay.				

Table 9
BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO:	Description	Sequence
1	PCR primer hPDI/UP1	AGCGCTGACGCCCCCGAGGAGGAGGACCAC
2	PCR primer hPDI/LP-PacI	CCTTAATTAATTACAGTTCATCATGCACAGCTTTCTGATCAT
3	PCR primer PB248	ATGAATTCAGGC CATATCGGCCATTGTTTACTGTGCG CCCACAGTAG
4	PCR primer PB249	ATGTTTA AACGTGAGGATTACTGGTGATGAAAGAC
5	PCR primer PB250	AGACTAGTCTATTTGGAG ACATTGACGGATCCAC
6	PCR primer PB251	ATCTCGAGAGGCCATGCAGGCCAACCACAAGATGAATCAAAT TTTG
7	PCR primer PpPDI/UPi-1	GGTGAGGTTGAGGTCCCAAGTGACTATCAAGGTC
8	PCR primer PpPDI/LPi-1	GACCTTGATAGTCACTTGGGACCTCAACCTCACC
9	PCR primer PpPDI/UPi-2	CGCCAATGATGAGGATGCCTCTTCAAAGGTTGTG
10	PCR primer PpPDI/LPi-2	CACAACCTTTGAAGAGGCATCCTCATCATTGGCG
11	PCR primer PpPDI-5'/UP	GGCGATTGCATTCGCGAC TGTATC
12	PCR primer hPDI-3'/LP	CCTAGAGAGCGGTGG CCAAGATG
13	PCR primer hPDI/UP	GTGGCCACACCAGGGGGC ATGGAAC
14	PCR primer hPDI-3'/LP	CCTAGAGAGCGGTGG CCAAGATG
15	PCR primer hGRP94/UP1	AGCGCTGACGATGAAGTTGATGTGGATGGTACA GTAG

16	PCR primer hGRP94/LP1	GGCCGGCCTTACAATTCATCATG TTCAGCTGTAGATTC
17	*Saccharomyces cerevisiae* mating factor pre-signal peptide (DNA)	ATG AGA TTC CCA TCC ATC TTC ACT GCT GTT TTG TTC GCT GCT TCT TCT GCT TTG GCT
18	*Saccharomyces cerevisiae* mating factor pre-signal peptide (protein)	MRFPSIFTAVLFAASSALA
19	human PDI Gene (DNA)	GACGCCCCCGAGGAGGAGGACCACGTCTTGGTGCTGCGGAAA AGCAACTTCGCGGAGGCGCTGGCGGCCCACAAGTACCCGCCG GTGGAGTTCCATGCCCCCTGGTGTGGCCACTGCAAGGCTCTGG CCCCTGAGTATGCCAAAGCCGCTGGGAAGCTGAAGGCAGAAG GTTCCGAGATCAGGTTGGCCAAGGTGGACGCCACGGAGGAGT CTGACCTAGCCCAGCAGTACGGCGTGCGCGGCTATCCCACCA TCAAGTTCTTCAGGAATGGAGACACGGCTTCCCCCAAGGAAT ATACAGCTGGCAGAGAGGCTGATGACATCGTGAACTGGCTGA AGAAGCGCACGGGCCCGGCTGCCACCACCCTGCCTGACGGCG CAGCTGCAGAGTCCTTGGTGGAGTCCAGCGAGGTGGCCGTCA TCGGCTTCTTCAAGGACGTGGAGTCGGACTCTGCCAAGCAGTT TTTGCAGGCAGCAGAGGCCATCGATGACATACCATTTGGGAT CACTTCCAACAGTGACGTGTTCTCCAAATACCAGCTCGACAA AGATGGGGTTGTCCTCTTTAAGAAGTTTGATGAAGGCCGGAA CAACTTTGAAGGGGAGGTCACCAAGGAGAACCTGCTGGACTT TATCAAACACAACCAGCTGCCCCTTGTCATCGAGTTCACCGAG CAGACAGCCCCGAAGATTTTTGGAGGTGAAATCAAGACTCAC ATCCTGCTGTTCTTGCCCAAGAGTGTGTCTGACTATGACGGCA AACTGAGCAACTTCAAAACAGCAGCCGAGAGCTTCAAGGGCA AGATCCTGTTCATCTTCATCGACAGCGACCACACCGACAACC AGCGCATCCTCGAGTTCTTTGGCCTGAAGAAGGAAGAGTGCC CGGCCGTGCGCCTCATCACCTTGGAGGAGGAGATGACCAAGT ACAAGCCCGAATCGGAGGAGCTGACGGCAGAGAGGATCACA GAGTTCTGCCACCGCTTCCTGGAGGGCAAAATCAAGCCCCAC

		CTGATGAGCCAGGAGCTGCCGGAGGACTGGGACAAGCAGCCT GTCAAGGTGCTTGTTGGGAAGAACTTTGAAGACGTGGCTTTT GATGAGAAAAAAAACGTCTTTGTGGAGTTCTATGCCCCATGG TGTGGTCACTGCAAACAGTTGGCTCCCATTTGGGATAAACTGG GAGAGACGTACAAGGACCATGAGAACATCGTCATCGCCAAGA TGGACTCGACTGCCAACGAGGTGGAGGCCGTCAAAGTGCACG GCTTCCCCACACTCGGGTTCTTTCCTGCCAGTGCCGACAGGAC GGTCATTGATTACAACGGGGAACGCACGCTGGATGGTTTTAA GAAATTCCTAGAGAGCGGTGGCCAAGATGGGGCAGGGGATGT TGACGACCTCGAGGACCTCGAAGAAGCAGAGGAGCCAGACAT GGAGGAAGACGATGACCAGAAAGCTGTGAAAGATGAACTGT AA
20	human PDI Gene (protein)	DAPEEEDHVLVLRKSNFAEALAAHKYPPVEFHAPWCGHCKALA PEYAKAAGKLKAEGSEIRLAKVDATEESDLAQQYGVRGYPTIKF FRNGDTASPKEYTAGREADDIVNWLKKRTGPAATTLPDGAAAE SLVESSEVAVIGFFKDVESDSAKQFLQAAEAIDDIPFGITSNSDVFS KYQLDKDGVVLFKKFDEGRNNFEGEVTKENLLDFIKHNQLPLVI EFTEQTAPKIFGGEIKTHILLFLPKSVSDYDGKLSNFKTAAESFKG KILFIFIDSDHTDNQRILEFFGLKKEECPAVRLITLEEEMTKYKPES EELTAERITEFCHRFLEGKIKPHLMSQELPEDWDKQPVKVLVGK NFEDVAFDEKKNVFVEFYAPWCGHCKQLAPIWDKLGETYKDHE NIVIAKMDSTANEVEAVKVHGFPTLGFFPASADRTVIDYNGERTL DGFKKFLESGGQDGAGDVDDLEDLEEAEEPDMEEDDDQKAVH DEL
21	*Pichia pastoris* PDI1 Gene (DNA)	ATGCAATTCAACTGGAATATTAAAACTGTGGCAAGTATTTTGT CCGCTCTCACACTAGCACAAGCAAGTGATCAGGAGGCTATTG CTCCAGAGGACTCTCATGTCGTCAAATTGACTGAAGCCACTTT TGAGTCTTTCATCACCAGTAATCCTCACGTTTTGGCAGAGTTT TTTGCCCCTTGGTGTGGTCACTGTAAGAAGTTGGGCCCTGAAC TTGTTTCTGCTGCCGAGATCTTAAAGGACAATGAGCAGGTTA AGATTGCTCAAATTGATTGTACGGAGGAGAAGGAATTATGTC AAGGCTACGAAATTAAAGGGTATCCTACTTTGAAGGTGTTCC ATGGTGAGGTTGAGGTCCCAAGTGACTATCAAGGTCAAAGAC AGAGCCAAAGCATTGTCAGCTATATGCTAAAGCAGAGTTTAC CCCCTGTCAGTGAAATCAATGCAACCAAAGATTTAGACGACA CAATCGCCGAGGCAAAAGAGCCCGTGATTGTGCAAGTACTAC CGGAAGATGCATCCAACTTGGAATCTAACACCACATTTTACG

		GAGTTGCCGGTACTCTCAGAGAGAAATTCACTTTTGTCTCCAC TAAGTCTACTGATTATGCCAAAAAATACACTAGCGACTCGAC TCCTGCCTATTTGCTTGTCAGACCTGGCGAGGAACCTAGTGTT TACTCTGGTGAGGAGTTAGATGAGACTCATTTGGTGCACTGG ATTGATATTGAGTCCAAACCTCTATTTGGAGACATTGACGGAT CCACCTTCAAATCATATGCTGAAGCTAACATCCCTTTAGCCTA CTATTTCTATGAGAACGAAGAACAACGTGCTGCTGCTGCCGA TATTATTAAACCTTTTGCTAAAGAGCAACGTGGCAAAATTAA CTTTGTTGGCTTAGATGCCGTTAAATTCGGTAAGCATGCCAAG AACTTAAACATGGATGAAGAGAAACTCCCTCTATTTGTCATTC ATGATTTGGTGAGCAACAAGAAGTTTGGAGTTCCTCAAGACC AAGAATTGACGAACAAAGATGTGACCGAGCTGATTGAGAAAT TCATCGCAGGAGAGGCAGAACCAATTGTGAAATCAGAGCCAA TTCCAGAAATTCAAGAAGAGAAAGTCTTCAAGCTAGTCGGAA AGGCCCACGATGAAGTTGTCTTCGATGAATCTAAAGATGTTCT AGTCAAGTACTACGCCCCTTGGTGTGGTCACTGTAAGAGAAT GGCTCCTGCTTATGAGGAATTGGCTACTCTTTACGCCAATGAT GAGGATGCCTCTTCAAAGGTTGTGATTGCAAAACTTGATCAC ACTTTGAACGATGTCGACAACGTTGATATTCAAGGTTATCCTA CTTTGATCCTTTATCCAGCTGGTGATAAATCCAATCCTCAACT GTATGATGGATCTCGTGACCTAGAATCATTGGCTGAGTTTGTA AAGGAGAGAGGAACCCACAAAGTGGATGCCCTAGCACTCAG ACCAGTCGAGGAAGAAAAGGAAGCTGAAGAAGAAGCTGAAA GTGAGGCAGACGCTCACGACGAGCTTTAA
22	*Pichia pastoris* PDI1 Gene (protein)	MQFNWNIKTVASILSALTLAQASDQEAIAPEDSHVVKLTEATFES FITSNPHVLAEFFAPWCGHCKKLGPELVSAAEILKDNEQVKIAQI DCTEEKELCQGYEIKGYPTLKVFHGEVEVPSDYQGQRQSQSIVSY MLKQSLPPVSEINATKDLDDTIAEAKEPVIVQVLPEDASNLESNT TFYGVAGTLREKFTFVSTKSTDYAKKYTSDSTPAYLLVRPGEEPS VYSGEELDETHLVHWIDIESKPLFGDIDGSTFKSYAEANIPLAYYF YENEEQRAAAADIIKPFAKEQRGKINFVGLDAVKFGKHAKNLN MDEEKLPLFVIHDLVSNKKFGVPQDQELTNKDVTELIEKFIAGEA EPIVKSEPIPEIQEEKVFKLVGKAHDEVVFDESKDVLVKYYAPWC GHCKRMAPAYEELATLYANDEDASSKVVIAKLDHTLNDVDNVD IQGYPTLILYPAGDKSNPQLYDGSRDLESLAEFVKERGTHKVDAL ALRPVEEEKEAEEEAESEADAHDEL
23	human ERO1α	GAAGAACAACCACCAGAGACTGCTGCTCAGAGATGCTTCTGT

	Gene (DNA)	CAGGTTTCCGGTTACTTGGACGACTGTACTTGTGACGTTGAGA CTATCGACAGATTCAACAACTACAGATTGTTCCCAAGATTGCA GAAGTTGTTGGAGTCCGACTACTTCAGATACTACAAGGTTAA CTTGAAGAGACCATGTCCATTCTGGAACGACATTTCCCAGTGT GGTAGAAGAGACTGTGCTGTTAAGCCATGTCAATCCGACGAA GTTCCAGACGGTATTAAGTCCGCTTCCTACAAGTACTCTGAAG AGGCTAACAACTTGATCGAAGAGTGTGAGCAAGCTGAAAGAT TGGGTGCTGTTGACGAATCTTTGTCCGAGAGACTCAGAAGGC TGTTTTGCAGTGGACTAAGCACGATGATTCCTCCGACAACTTC TGTGAAGCTGACGACATTCAATCTCCAGAGGCTGAGTACGTT GACTTGTTGTTGAACCCAGAGAGATACACTGGTTACAAGGGT CCAGACGCTTGGAAGATTTGGAACGTTATCTACGAAGAGAAC TGTTTCAAGCCACAGACTATCAAGAGACCATTGAACCCATTG GCTTCCGGACAGGGAACTTCTGAAGAGAACACTTTCTACTCTT GGTTGGAGGGTTTGTGTGTTGAGAAGAGAGCTTTCTACAGAT TGATCTCCGGATTGCACGCTTCTATCAACGTTCACTTGTCCGC TAGATACTTGTTGCAAGAGACTTGGTTGGAAAAGAAGTGGGG TCACAACATTACTGAGTTCCAGCAGAGATTCGACGGTATTTTG ACTGAAGGTGAAGGTCCAAGAAGATTGAAGAACTTGTACTTT TTGTACTTGATCGAGTTGAGAGCTTTGTCCAAGGTTTTGCCAT TCTTCGAGAGACCAGACTTCCAATTGTTCACTGGTAACAAGAT CCAGGACGAAGAGAACAAGATGTTGTTGTTGGAGATTTTGCA CGAGATCAAGTCCTTTCCATTGCACTTCGACGAGAACTCATTT TTCGCTGGTGACAAGAAAGAAGCTCACAAGTTGAAAGAGGAC TTCAGATTGCACTTCAGAAATATCTCCAGAATCATGGACTGTG TTGGTTGTTTCAAGTGTAGATTGTGGGGTAAGTTGCAGACTCA AGGATTGGGTACTGCTTTGAAGATTTTGTTCTCCGAGAAGTTG ATCGCTAACATGCCTGAATCTGGTCCATCTTACGAGTTCCACT TGACTAGACAAGAGATCGTTTCCTTGTTCAACGCTTTCGGTAG AATCTCCACTTCCGTTAAAGAGTTGGAGAACTTCAGAAACTTG TTGCAGAACATCCACTAA
24	human ERO1α Gene (protein)	EEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLFPRLQKL LESDYFRYYKVNLKRPCPFWNDISQCGRRDCAVKPCQSDEVPDG IKSASYKYSEEANNLIEECEQAERLGAVDESLSEETQKAVLQWT KHDDSSDNFCEADDIQSPEAEYVDLLLNPERYTGYKGPDAWKIW NVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKR AFYRLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFD

		GILTEGEGPRRLKNLYFLYLIELRALSKVLPFFERPDFQLFTGNKI QDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAHKLKEDFRL HFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMP ESGPSYEFHLTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH
25	human GRP94 Gene (DNA)	GATGATGAAGTTGACGTTGACGGTACTGTTGAAGAGGACTTG GGAAAGTCTAGAGAGGGTTCCAGAACTGACGACGAAGTTGTT CAGAGAGAGGAAGAGGCTATTCAGTTGGACGGATTGAACGCT TCCCAAATCAGAGAGTTGAGAGAGAAGTCCGAGAAGTTCGCT TTCCAAGCTGAGGTTAACAGAATGATGAAATTGATTATCAAC TCCTTGTACAAGAACAAAGAGATTTTCTTGAGAGAGTTGATCT CTAACGCTTCTGACGCTTTGGACAAGATCAGATTGATCTCCTT GACTGACGAAAACGCTTTGTCCGGTAACGAAGAGTTGACTGT TAAGATCAAGTGTGACAAAGAGAAGAACTTGTTGCACGTTAC TGACACTGGTGTTGGAATGACTAGAGAAGAGTTGGTTAAGAA CTTGGGTACTATCGCTAAGTCTGGTACTTCCGAGTTCTTGAAC AAGATGACTGAGGCTCAAGAAGATGGTCAATCCACTTCCGAG TTGATTGGTCAGTTCGGTGTTGGTTTCTACTCCGCTTTCTTGGT TGCTGACAAGGTTATCGTTACTTCCAAGCACAACAACGACAC TCAACACATTTGGGAATCCGATTCCAACGAGTTCTCCGTTATT GCTGACCCAAGAGGTAACACTTTGGGTAGAGGTACTACTATC ACTTTGGTTTTGAAAGAAGAGGCTTCCGACTACTTGGAGTTGG ACACTATCAAGAACTTGGTTAAGAAGTACTCCCAGTTCATCA ACTTCCCAATCTATGTTTGGTCCTCCAAGACTGAGAC TGTTGAGGAACCAATGGAAGAAGAAGAGGCTGCTAAAGAAG AGAAAGAGGAATCTGACGACGAGGCTGCTGTTGAAGAAGAG GAAGAAGAAAAGAAGCCAAAGACTAAGAAGGTTGAAAAGAC TGTTTGGGACTGGGAGCTTATGAACGACATCAAGCCAATTTG GCAGAGACCATCCAAAGAGGTTGAGGAGGACGAGTACAAGG CTTTCTACAAGTCCTTCTCCAAAGAATCCGATGACCCAATGGC TTACATCCACTTCACTGCTGAGGGTGAAGTTACTTTCAAGTCC ATCTTGTTCGTTCCAACTTCTGCTCCAAGAGGATTGTTCGACG AGTACGGTTCTAAGAAGTCCGACTACATCAAACTTTATGTTAG AAGAGTTTTCATCACTGACGACTTCCACGATATGATGCCAAA GTACTTGAACTTCGTTAAGGGTGTTGTTGATTCCGATGACTTG CCATTGAACGTTTCCAGAGAGACTTTGCAGCAGCACAAGTTG TTGAAGGTTATCAGAAAGAAACTTGTTAGAAAGACTTTGGAC ATGATCAAGAAGATCGCTGACGACAAGTACAACGACACTTTC

		TGGAAAGAGTTCGGAACTAACATCAAGTTGGGTGTTATTGAG GACCACTCCAACAGAACTAGATTGGCTAAGTTGTTGAGATTC CAGTCCTCTCATCACCCAACTGACATCACTTCCTTGGACCAGT ACGTTGAGAGAATGAAAGAGAAGCAGGACAAAATCTACTTCA TGGCTGGTTCCTCTAGAAAAGAGGCTGAATCCTCCCCATTCGT TGAGAGATTGTTGAAGAAGGGTTACGAGGTTATCTACTTGAC TGAGCCAGTTGACGAGTACTGTATCCAGGCTTTGCCAGAGTTT GACGGAAAGAGATTCCAGAACGTTGCTAAAGAGGGTGTTAAG TTCGACGAATCCGAAAAGACTAAAGAATCCAGAGAGGCTGTT GAGAAAGAGTTCGAGCCATTGTTGAACTGGATGAAGGACAAG GCTTTGAAGGACAAGATCGAGAAGGCTGTTGTTTCCCAGAGA TTGACTGAATCCCCATGTGCTTTGGTTGCTTCCCAATACGGAT GGAGTGGTAACATGGAAAGAATCATGAAGGCTCAGGCTTACC AAACTGGAAAGGACATCTCCACTAACTACTACGCTTCCCAGA AGAAAACTTTCGAGATCAACCCAAGACACCCATTGATCAGAG ACATGTTGAGAAGAATCAAAGAGGACGAGGACGACAAGACT GTTTTGGATTTGGCTGTTGTTTTGTTCGAGACTGCTACTTTGA GATCCGGTTACTTGTTGCCAGACACTAAGGCTTACGGTGACA GAATCGAGAGAATGTTGAGATTGTCCTTGAACATTGACCCAG ACGCTAAGGTTGAAGAAGAACCAGAAGAAGAGCCAGAGGAA ACTGCTGAAGATACTACTGAGGACACTGAACAAGACGAGGAC GAAGAGATGGATGTTGGTACTGACGAAGAGGAAGAGACAGC AAAGGAATCCACTGCTGAACACGACGAGTTGTAA
26	human GRP94 Gene (protein)	DDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDGLNASQ IRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDA LDKIRLISLTDENALSGNEELTVKIKCDKEKNLLHVTDTGVGMTR EELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSELIGQFGVGFYS AFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGT TITLVLKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTETVEE PMEEEEAAKEEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWE LMNDIKPIWQRPSKEVEEDEYKAFYKSFSKESDDPMAYIHFTAEG EVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITDDFHD MMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRK TLDMIKKIADDKYNDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRF QSSHHPTDITSLDQYVERMKEKQDKIYFMAGSSRKEAESSPFVER LLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDES EKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPC

		ALVASQYGWSGNMERIMKAQAYQTGKDISTNYYASQKKTFEIN PRHPLIRDMLRRIKEDEDDKTVLDLAVVLFETATLRSGYLLPDTK AYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDE DEEMDVGTDEEEETAKESTAEHDEL
27	anti-DKK1 Heavy chain (VH + IgG2m4) (α-amylase encoding sequences underlined) (DNA)	ACGATGGTCGCTTGGTGGTCTTTGTTTCTGTACGGTCTTCAGG TCGCTGCACCTGCTTTGGCTGAGGTTCAGTTGGTTCAATCTGG TGCTGAGGTTAAGAAACCTGGTGCTTCCGTTAAGGTTTCCTGT AAGGCTTCCGGTTACACTTTCACTGACTACTACATCCACTGGG TTAGACAAGCTCCAGGTCAAGGATTGGAATGGATGGGATGGA TTCACTCTAACTCCGGTGCTACTACTTACGCTCAGAAGTTCCA GGCTAGAGTTACTATGTCCAGAGACACTTCTTCTTCCACTGCT TACATGGAATTGTCCAGATTGGAATCCGATGACACTGCTATGT ACTTTTGTTCCAGAGAGGACTACTGGGGACAGGGAACTTTGG TTACTGTTTCCTCCGCTTCTACTAAAGGGCCCTCTGTTTTTCCA TTGGCTCCATGTTCTAGATCCACTTCCGAATCCACTGCTGCTT TGGGATGTTTGGTTAAGGACTACTTCCCAGAGCCAGTTACTGT TTCTTGGAACTCCGGTGCTTTGACTTCTGGTGTTCACACTTTCC CAGCTGTTTTGCAATCTTCCGGTTTGTACTCCTTGTCCTCCGTT GTTACTGTTACTTCCTCCAACTTCGGTACTCAGACTTACACTT GTAACGTTGACCACAAGCCATCCAACACTAAGGTTGACAAGA CTGTTGAGAGAAAGTGTTGTGTTGAGTGTCCACCATGTCCAGC TCCACCAGTTGCTGGTCCATCCGTTTTTTTGTTCCCACCAAAG CCAAAGGACACTTTGATGATCTCCAGAACTCCAGAGGTTACA TGTGTTGTTGTTGACGTTTCCCAAGAGGACCCAGAGGTTCAAT TCAACTGGTACGTTGACGGTGTTGAAGTTCACAACGCTAAGA CTAAGCCAAGAGAAGAGCAGTTCAACTCCACTTTCAGAGTTG TTTCCGTTTTGACTGTTTTGCACCAGGATTGGTTGAACGGTAA AGAATACAAGTGTAAGGTTTCCAACAAGGGATTGCCATCCTC CATCGAAAAGACTATCTCCAAGACTAAGGGACAACCAAGAGA GCCACAGGTTTACACTTTGCCACCATCCAGAGAAGAGATGAC TAAGAACCAGGTTTCCTTGACTTGTTTGGTTAAAGGATTCTAC CCATCCGACATTGCTGTTGAGTGGGAATCTAACGGTCAACCA GAGAACAACTACAAGACTACTCCACCAATGTTGGATTCTGAC GGTTCCTTCTTCTTGTACTCCAAGTTGACTGTTGACAAGTCCA GATGGCAACAGGGTAACGTTTTCTCCTGTTCCGTTATGCATGA GGCTTTGCACAACCACTACACTCAAAAGTCCTTGTCTTTGTCC CCTGGTAAGTAA

28	anti-DKK1 Heavy chain (VH + IgG2m4) (protein)	EVQLVQSGAEVKKPGASVKVSCKASGYTFTDYYIHWVRQAPGQ GLEWMGWIHSNSGATTYAQKFQARVTMSRDTSSSTAYMELSRL ESDDTAMYFCSREDYWGQGTLVTVSSASTKGPSVFPLAPCSRST SESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL YSLSSVVTVTSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVEC PPCPAPPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPE VQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVLHQDWL NGKEYKCKVSNKGLPSSIEKTISKTKGQPREPQVYTLPPSREEMT KNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGS FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG K
29	anti-DKK1 Light chain (VL + lambda constant regions) (α-amylase encoding sequences underlined) (DNA)	ACGATGGTCGCTTGGTGGTCTTTGTTTCTGTACGGTCTTCAGG TCGCTGCACCTGCTTTGGCTCAGTCCGTTTTGACACAACCACC ATCTGTTTCTGGTGCTCCAGGACAGAGAGTTACTATCTCCTGT ACTGGTTCCTCTTCCAACATTGGTGCTGGTTACGATGTTCACT GGTATCAACAGTTGCCAGGTACTGCTCCAAAGTTGTTGATCTA CGGTTACTCCAACAGACCATCTGGTGTTCCAGACAGATTCTCT GGTTCTAAGTCTGGTGCTTCTGCTTCCTTGGCTATCACTGGAT TGAGACCAGATGACGAGGCTGACTACTACTGTCAATCCTACG ACAACTCCTTGTCCTCTTACGTTTTCGGTGGTGGTACTCAGTT GACTGTTTTGTCCCAGCCAAAGGCTAATCCAACTGTTACTTTG TTCCCACCATCTTCCGAAGAACTGCAGGCTAATAAGGCTACTT TGGTTTGTTTGATCTCCGACTTCTACCCAGGTGCTGTTACTGTT GCTTGGAAGGCTGATGGTTCTCCAGTTAAGGCTGGTGTTGAG ACTACTAAGCCATCCAAGCAGTCCAATAACAAGTACGCTGCT AGCTCTTACTTGTCCTTGACACCAGAACAATGGAAGTCCCACA GATCCTACTCTTGTCAGGTTACACACGAGGGTTCTACTGTTGA AAAGACTGTTGCTCCAACTGAGTGTTCCTAA
30	anti-DKK1 Light chain (VL + lambda constant regions) (protein)	QSVLTQPPSVSGAPGQRVTISCTGSSSNIGAGYDVHWYQQLPGT APKLLIYGYSNRPSGVPDRFSGSKSGASASLAITGLRPDDEADYY CQSYDNSLSSYVFGGGTQLTVLSQPKANPTVTLFPPSSEELQANK ATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKY AASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
31	PpPDI1 promoter	AACACGAACACTGTAAATAGAATAAAAGAAAACTTGGATAGT AGAACTTCAATGTAGTGTTTCTATTGTCTTACGCGGCT

		CTTTAGATTGCAATCCCCAGAATGGAATCGTCCATCTTTCTCA ACCCACTCAAAGATAATCTACCAGACATACCTACGCC CTCCATCCCAGCACCACGTCGCGATCACCCCTAAAACTTCAAT AATTGAACACGTACTGATTTCCAAACCTTCTTCTTCT TCCTATCTATAAGA
32	PpPMR1	ATGACAGCTAATGAAAATCCTTTTGAGAATGAGCTGACAGGA TCTTCTGAATCTGCCCCCCCTGCATTGGAATCGAAGACTGGAG AGTCTCTTAAGTATTGCAAATATACCGTGGATCAGGTCATAG AAGAGTTTCAAACGGATGGTCTCAAAGGATTGTGCAATTCCC AGGACATCGTATATCGGAGGTCTGTTCATGGGCCAAATGAAA TGGAAGTCGAAGAGGAAGAGAGTCTTTTTTCGAAATTCTTGT CAAGTTTCTACAGCGATCCATTGATTCTGTTACTGATGGGTTC CGCTGTGATTAGCTTTTTGATGTCTAACATTGATGATGCGATA TCTATCACTATGGCAATTACGATCGTTGTCACAGTTGGATTTG TTCAAGAGTATCGATCCGAGAAATCATTGGAGGCATTGAACA AGTTAGTCCCTGCCGAAGCTCATCTAACTAGGAATGGGAACA CTGAAACTGTTCTTGCTGCCAACCTAGTCCCAGGAGACTTGGT GGATTTTTCGGTTGGTGACAGAATTCCGGCTGATGTGAGAATT ATTCACGCTTCCCACTTGAGTATCGACGAGAGCAACCTAACTG GTGAAAATGAACCAGTTTCTAAAGACAGCAAACCTGTTGAAA GTGATGACCCAAACATTCCCTTGAACAGCCGTTCATGTATTGG GTATATGGGCACTTTAGTTCGTGATGGTAATGGCAAAGGTATT GTCATCGGAACAGCCAAAAACACAGCTTTTGGCTCTGTTTTCG AAATGATGAGCTCTATTGAGAAACCAAAGACTCCTCTTCAAC AGGCTATGGATAAACTTGGTAAGGATTTGTCTGCTTTTTCCTT CGGAATCATCGGCCTTATTTGCTTGGTTGGTGTTTTTCAAGGT AGACCCTGGTTGGAAATGTTCCAGATCTCTGTATCCTTGGCTG TTGCTGCGATTCCAGAAGGTCTTCCTATTATTGTGACTGTGAC TCTTGCTCTTGGTGTGTTGCGTATGGCTAAACAGAGGGCCATC GTCAAAAGACTGCCTAGTGTTGAAACTTTGGGATCCGTCAAT GTTATCTGTAGTGATAAGACGGGAACATTGACCCAAAATCAT ATGACCGTTAACAGATTATGGACTGTGGATATGGGCGATGAA TTCTTGAAAATTGAACAAGGGGAGTCCTATGCCAATTATCTCA AACCCGATACGCTAAAAGTTCTGCAAACTGGTAATATAGTCA ACAATGCCAAATATTCAAATGAAAAGGAAAAATACCTCGGAA ACCCAACTGATATTGCAATTATTGAATCTTTAGAAAAATTTGA TTTGCAGGACATTAGAGCAACAAAGGAAAGAATGTTGGAGAT

		TCCATTTTCTTCGTCCAAGAAATATCAGGCCGTCAGTGTTCAC TCTGGAGACAAAAGCAAATCTGAAATTTTTGTTAAAGGCGCT CTGAACAAAGTTTTGGAAAGATGTTCAAGATATTACAATGCT GAAGGTATCGCCACTCCACTCACAGATGAAATTAGAAGAAAA TCCTTGCAAATGGCCGATACGTTAGCATCTTCAGGATTGAGAA TACTGTCGTTTGCTTACGACAAAGGCAATTTTGAAGAAACTG GCGATGGACCATCGGATATGATCTTTTGTGGTCTTTTAGGTAT GAACGATCCTCCTAGACCATCTGTAAGTAAATCAATTTTGAAA TTCATGAGAGGTGGGGTTCACATTATTATGATTACAGGAGATT CAGAATCCACGGCCGTAGCCGTTGCCAAACAGGTCGGAATGG TAATTGACAATTCAAAATATGCTGTCCTCAGTGGAGACGATA TAGATGCTATGAGTACAGAGCAACTGTCTCAGGCGATCTCAC ATTGTTCTGTATTTGCCCGGACTACTCCAAAACATAAGGTGTC CATTGTAAGAGCACTACAGGCCAGAGGAGATATTGTTGCAAT GACTGGTGACGGTGTCAATGATGCCCCAGCTCTAAAACTGGC CGACATCGGAATTGCCATGGGTAATATGGGGACCGATGTTGC CAAAGAGGCAGCCGACATGGTTTTGACTGATGATGACTTTTCT ACAATCTTATCTGCAATCCAGGAGGGTAAAGGTATTTTCTACA ACATCCAGAACTTTTTAACGTTCCAACTTTCTACTTCAATTGC TGCTCTTTCGTTAATTGCTCTGAGTACTGCTTTCAACCTGCCA AATCCATTGAATGCCATGCAGATTTTGTGGATCAATATTATCA TGGATGGACCTCCAGCTCAGTCTTTGGGTGTTGAGCCAGTTGA TAAAGCTGTGATGAACAAACCACCAAGAAAGCGAAATGATAA AATTCTGACAGGTAAGGTGATTCAAAGGGTAGTACAAAGTAG TTTTATCATTGTTTGTGGTACTCTGTACGTATACATGCATGAG ATCAAAGATAATGAGGTCACAGCAAGAGACACTACGATGACC TTTACATGCTTTGTATTCTTTGACATGTTCAACGCATTAACGA CAAGACACCATTCTAAAAGTATTGCAGAACTTGGATGGAATA ATACTATGTTCAACTTTTCCGTTGCAGCTTCTATTTTGGGTCA ACTAGGAGCTATTTACATTCCATTTTTGCAGTCTATTTTCCAG ACTGAACCTCTGAGCCTCAAAGATTTGGTCCATTTATTGTTGT TATCGAGTTCAGTATGGATTGTAGACGAGCTTCGAAAACTCT ACGTCAGGAGACGTGACGCATCCCCATACAATGGATACAGCA TGGCTGTTTGA
33	PpPMR1	MTANENPFENELTGSSESAPPALESKTGESLKYCKYTVDQVIEEF QTDGLKGLCNSQDIVYRRSVHGPNEMEVEEEESLFSKFLSSFYSD PLILLLMGSAVISFLMSNIDDAISITMAITIVVTVGFVQEYRSEKSL

		EALNKLVPAEAHLTRNGNTETVLAANLVPGDLVDFSVGDRIPAD VRIIHASHLSIDESNLTGENEPVSKDSKPVESDDPNIPLNSRSCIGY MGTLVRDGNGKGIVIGTAKNTAFGSVFEMMSSIEKPKTPLQQAM DKLGKDLSAFSFGIIGLICLVGVFQGRPWLEMFQISVSLAVAAIPE GLPIIVTVTLALGVLRMAKQRAIVKRLPSVETLGSVNVICSDKTG TLTQNHMTVNRLWTVDMGDEFLKIEQGESYANYLKPDTLKVLQ TGNIVNNAKYSNEKEKYLGNPTDIAIIESLEKFDLQDIRATKERM LEIPFSSSKKYQAVSVHSGDKSKSEIFVKGALNKVLERCSRYYNA EGIATPLTDEIRRKSLQMADTLASSGLRILSFAYDKGNFEETGDGP SDMIFCGLLGMNDPPRPSVSKSILKFMRGGVHIIMITGDSESTAVA VAKQVGMVIDNSKYAVLSGDDIDAMSTEQLSQAISHCSVFARTT PKHKVSIVRALQARGDIVAMTGDGVNDAPALKLADIGIAMGNM GTDVAKEAADMVLTDDDFSTILSAIQEGKGIFYNIQNFLTFQLSTS IAALSLIALSTAFNLPNPLNAMQILWINIIMDGPPAQSLGVEPVDK AVMNKPPRKRNDKILTGKVIQRVVQSSFIIVCGTLYVYMHEIKDN EVTARDTTMTFTCFVFFDMFNALTTRHHSKSIAELGWNNTMFNF SVAASILGQLGAIYIPFLQSIFQTEPLSLKDLVHLLLLSSSVWIVDE LRKLYVRRRDASPYNGYSMAV
34	*Arabidopsis Thaliana* AtECA1 (codon optimized for *Pichia pastoris*)	ATGGGAAAGGGTTCCGAGGACCTGGTTAAGAAAGAATCCCTG AACTCCACTCCAGTTAACTCTGACACTTTCCCAGCTTGGGCTA AGGATGTTGCTGAGTGCGAAGAGCACTTCGTTGTTTCCAGAG AGAAGGGTTTGTCCTCCGACGAAGTCTTGAAGAGACACCAAA TCTACGGACTGAACGAGTTGGAAAAGCCAGAGGGAACCTCCA TCTTCAAGCTGATCTTGGAGCAGTTCAACGACACCCTTGTCAG AATTTTGTTGGCTGCCGCTGTTATTTCCTTCGTCCTGGCTTTTT TTGATGGTGACGAGGGTGGTGAAATGGGTATCACTGCCTTCG TTGAGCCTTTGGTCATCTTCCTGATCTTGATCGTTAACGCCAT CGTTGGTATCTGGCAAGAGACTAACGCTGAAAAGGCTTTGGA GGCCTTGAAAGAGATTCAATCCCAGCAGGCTACCGTTATGAG AGATGGTACTAAGGTTTCCTCCTTGCCAGCTAAAGAATTGGTT CCAGGTGACATCGTTGAGCTGAGAGTTGGTGATAAGGTTCCA GCCGACATGAGAGTTGTTGCTTTGATCTCCTCCACCTTGAGAG TTGAACAAGGTTCCCTGACTGGTGAATCTGAGGCTGTTTCCAA GACTACTAAGCACGTTGACGAGAACGCTGACATCCAGGGTAA AAAGTGCATGGTTTTCGCCGGTACTACCGTTGTTAACGGTAAC TGCATCTGTTTGGTCACTGACACTGGAATGAACACCGAGATC GGTAGAGTTCACTCCCAAATCCAAGAAGCTGCTCAACACGAA

		GAGGACACCCCATTGAAGAAGAAGCTGAACGAGTTCGGAGA GGTCTTGACCATGATCATCGGATTGATCTGTGCCCTGGTCTGG TTGATCAACGTCAAGTACTTCTTGTCCTGGGAATACGTTGATG GATGGCCAAGAAACTTCAAGTTCTCCTTCGAGAAGTGCACCT ACTACTTCGAGATCGCTGTTGCTTTGGCTGTTGCTGCTATTCC AGAGGGATTGCCAGCTGTTATCACCACTTGCTTGGCCTTGGGT ACTAGAAAGATGGCTCAGAAGAACGCCCTTGTTAGAAAGTTG CCATCCGTTGAGACTTTGGGTTGTACTACCGTCATCTGTTCCG ACAAGACTGGTACTTTGACTACCAACCAGATGGCCGTTTCCA AATTGGTTGCCATGGGTTCCAGAATCGGTACTCTGAGATCCTT CAACGTCGAGGGAACTTCTTTTGACCCAAGAGATGGAAAGAT TGAGGACTGGCCAATGGGTAGAATGGACGCCAACTTGCAGAT GATTGCTAAGATCGCCGCTATCTGTAACGACGCTAACGTTGA GCAATCCGACCAACAGTTCGTTTCCAGAGGAATGCCAACTGA GGCTGCCTTGAAGGTTTTGGTCGAGAAGATGGGTTTCCCAGA AGGATTGAACGAGGCTTCTTCCGATGGTGACGTCTTGAGATG TTGCAGACTGTGGAGTGAGTTGGAGCAGAGAATCGCTACTTT GGAGTTCGACAGAGATAGAAAGTCCATGGGTGTCATGGTTGA TTCTTCCTCCGGTAACAAGTTGTTGTTGGTCAAAGGAGCAGTT GAAAACGTTTTGGAGAGATCCACCCACATTCAATTGCTGGAC GGTTCCAAGAGAGAATTGGACCAGTACTCCAGAGACTTGATC TTGCAGTCCTTGAGAGACATGTCCTTGTCCGCCTTGAGATGTT TGGGTTTCGCTTACTCTGACGTTCCATCCGATTTCGCTACTTA CGATGGTTCTGAGGATCATCCAGCTCACCAACAGTTGCTGAA CCCATCCAACTACTCCTCCATCGAATCCAACCTGATCTTCGTT GGTTTCGTCGGTCTTAGAGACCCACCAAGAAAAGAAGTTAGA CAGGCCATCGCTGATTGTAGAACCGCCGGTATCAGAGTTATG GTCATCACCGGAGATAACAAGTCCACTGCCGAGGCTATTTGT AGAGAGATCGGAGTTTTCGAGGCTGACGAGGACATTTCTTCC AGATCCCTGACCGGTATTGAGTTCATGGACGTCCAAGACCAG AAGAACCACTTGAGACAGACCGGTGGTTTGTTGTTCTCCAGA GCCGAACCAAAGCACAAGCAAGAGATTGTCAGACTGCTGAAA GAGGACGGAGAAGTTGTTGCTATGACCGGTGATGGTGTTAAT GACGCCCCAGCTTTGAAGTTGGCTGACATCGGTGTTGCTATGG GAATTTCCGGTACTGAAGTTGCTAAGGAAGCCTCCGATATGG TTTTGGCTGACGACAACTTTTCAACTATCGTTGCTGCTGTCGG AGAAGGTAGAAGTATCTACAACAACATGAAAGCCTTTATCAG

		ATACATGATTTCCTCCAACATCGGTGAAGTTGCCTCCATTTTC TTGACTGCTGCCTTGGGTATTCCTGAGGGAATGATCCCAGTTC AGTTGTTGTGGGTTAACTTGGTTACTGACGGTCCACCTGCTAC TGCTTTGGGTTTCAACCCACCAGACAAAGACATTATGAAGAA GCCACCAAGAAGATCCGACGATTCCTTGATCACCGCCTGGAT CTTGTTCAGATACATGGTCATCGGTCTTTATGTTGGTGTTGCC ACCGTCGGTGTTTTCATCATCTGGTACACCCACTCTTCCTTCAT GGGTATTGACTTGTCTCAAGATGGTCATTCTTTGGTTTCCTAC TCCCAATTGGCTCATTGGGGACAATGTTCTTCCTGGGAGGGTT TCAAGGTTTCCCCATTCACTGCTGGTTCCCAGACTTTCTCCTTC GATTCCAACCCATGTGACTACTTCCAGCAGGGAAAGATCAAG GCTTCCACCTTGTCTTTGTCCGTTTTGGTCGCCATTGAGATGTT CAACTCCCTGAACGCTTTGTCTGAGGACGGTTCCTTGGTTACT ATGCCACCTTGGGTGAACCCATGGTTGTTGTTGGCTATGGCTG TTTCCTTCGGATTGCACTTCGTCATCCTGTACGTTCCATTCTTG GCCCAGGTTTTCGGTATTGTTCCACTGTCCTTGAACGAGTGGT TGTTGGTCTTGGCCGTTTCTTTGCCAGTTATCCTGATCGACGA GGTTTTGAAGTTCGTTGGTAGATGCACCTCTGGTTACAGATAC TCCCCAAGAACTCTGTCCACCAAGCAGAAAGAAGAGTAA
35	AtECA1	MGKGSEDLVKKESLNSTPVNSDTFPAWAKDVAECEEHFVVSRE KGLSSDEVLKRHQIYGLNELEKPEGTSIFKLILEQFNDTLVRILLA AAVISFVLAFFDGDEGGEMGITAFVEPLVIFLILIVNAIVGIWQETN AEKALEALKEIQSQQATVMRDGTKVSSLPAKELVPGDIVELRVG DKVPADMRVVALISSTLRVEQGSLTGESEAVSKTTKHVDENADI QGKKCMVFAGTTVVNGNCICLVTDTGMNTEIGRVHSQIQEAAQ HEEDTPLKKKLNEFGEVLTMIIGLICALVWLINVKYFLSWEYVDG WPRNFKFSFEKCTYYFEIAVALAVAAIPEGLPAVITTCLALGTRK MAQKNALVRKLPSVETLGCTTVICSDKTGTLTTNQMAVSKLVA MGSRIGTLRSFNVEGTSFDPRDGKIEDWPMGRMDANLQMIAKIA AICNDANVEQSDQQFVSRGMPTEAALKVLVEKMGFPEGLNEAS SDGDVLRCCRLWSELEQRIATLEFDRDRKSMGVMVDSSSGNKL LLVKGAVENVLERSTHIQLLDGSKRELDQYSRDLILQSLRDMSLS ALRCLGFAYSDVPSDFATYDGSEDHPAHQQLLNPSNYSSIESNLIF VGFVGLRDPPRKEVRQAIADCRTAGIRVMVITGDNKSTAEAICRE IGVFEADEDISSRSLTGIEFMDVQDQKNHLRQTGGLLFSRAEPKH KQEIVRLLKEDGEVVAMTGDGVNDAPALKLADIGVAMGISGTE VAKEASDMVLADDNFSTIVAAVGEGRSIYNNMKAFIRYMISSNIG

		EVASIFLTAALGIPEGMIPVQLLWVNLVTDGPPATALGFNPPDKD IMKKPPRRSDDSLITAWILFRYMVIGLYVGVATVGVFIIWYTHSS FMGIDLSQDGHSLVSYSQLAHWGQCSSWEGFKVSPFTAGSQTFS FDSNPCDYFQQGKIKASTLSLSVLVAIEMFNSLNALSEDGSLVTM PPWVNPWLLLAMAVSFGLHFVILYVPFLAQVFGIVPLSLNEWLL VLAVSLPVILIDEVLKFVGRCTSGYRYSPRTLSTKQKEE
36	PpPMR1/UP	GAATTCATGACAGCTAATGAAAATCCTTTTGAGAATGAG
37	PpPMR1/LP	GGCCGGCCTCAAACAGCCATGCTGTATCCATTGTATG
38	5'AOX1	GCGACTGGTTCCAATTGACAAGCTT
39	PpPMR1/cLP	GGTTGCTCTCGTCGATACTCAAGTGGGAAG
40	AtECA1/cLP	GTCGGCTGGAACCTTATCACCAACTCTCAG
41	Human calreticulin (hCRT)	ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTCGCAGCATC CTCCGCATTAGCTTACCCATACGACGTCCCAGACTACGCTTAC CCATACGACGTCCCAGACTACGCTGAGCCCGCCGTCTACTTCA AGGAGCAGTTTCTGGACGGAGACGGGTGGACTTCCCGCTGGA TCGAATCCAAACACAAGTCAGATTTTGGCAAATTCGTTCTCAG TTCCGGCAAGTTCTACGGTGACGAGGAGAAAGATAAAGGTTT GCAGACAAGCCAGGATGCACGCTTTTATGCTCTGTCGGCCAG TTTCGAGCCTTTCAGCAACAAAGGCCAGACGCTGGTGGTGCA GTTCACGGTGAAACATGAGCAGAACATCGACTGTGGGGGCGG CTATGTGAAGCTGTTTCCTAATAGTTTGGACCAGACAGACATG CACGGAGACTCAGAATACAACATCATGTTTGGTCCCGACATC TGTGGCCCTGGCACCAAGAAGGTTCATGTCATCTTCAACTACA AGGGCAAGAACGTGCTGATCAACAAGGACATCCGTTGCAAGG ATGATGAGTTTACACACCTGTACACACTGATTGTGCGGCCAG ACAACACCTATGAGGTGAAGATTGACAACAGCCAGGTGGAGT CCGGCTCCTTGGAAGACGATTGGGACTTCCTGCCACCCAAGA AGATAAAGGATCCTGATGCTTCAAAACCGGAAGACTGGGATG AGCGGGCCAAGATCGATGATCCCACAGACTCCAAGCCTGAGG ACTGGGACAAGCCCGAGCATATCCCTGACCCTGATGCTAAGA AGCCCGAGGACTGGGATGAAGAGATGGACGGAGAGTGGGAA CCCCCAGTGATTCAGAACCCTGAGTACAAGGGTGAGTGGAAG CCCCGGCAGATCGACAACCCAGATTACAAGGGCACTTGGATC CACCCAGAAATTGACAACCCCGAGTATTCTCCCGATCCCAGT ATCTATGCCTATGATAACTTTGGCGTGCTGGGCCTGGACCTCT GGCAGGTCAAGTCTGGCACCATCTTTGACAACTTCCTCATCAC CAACGATGAGGCATACGCTGAGGAGTTTGGCAACGAGACGTG

		GGGCGTAACAAAGGCAGCAGAGAAACAAATGAAGGACAAAC AGGACGAGGAGCAGAGGCTTAAGGAGGAGGAAGAAGACAAG AAACGCAAAGAGGAGGAGGAGGCAGAGGACAAGGAGGATGA TGAGGACAAAGATGAGGATGAGGAGGATGAGGAGGACAAGG AGGAAGATGAGGAGGAAGATGTCCCCGGCCAGGCCCATGAC GAGCTGTAG
42	Human calreticulin (hCRT)	MRFPSIFTAVLFAASSALAYPYDVPDYAYPYDVPDYAEPAVYFK EQFLDGDGWTSRWIESKHKSDFGKFVLSSGKFYGDEEKDKGLQ TSQDARFYALSASFEPFSNKGQTLVVQFTVKHEQNIDCGGGYVK LFPNSLDQTDMHGDSEYNIMFGPDICGPGTKKVHVIFNYKGKNV LINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLEDDW DFLPPKKIKDPDASKPEDWDERAKIDDPTDSKPEDWDKPEHIPDP DAKKPEDWDEEMDGEWEPPVIQNPEYKGEWKPRQIDNPDYKGT WIHPEIDNPEYSPDPSIYAYDNFGVLGLDLWQVKSGTIFDNFLITN DEAYAEEFGNETWGVTKAAEKQMKDKQDEEQRLKEEEEDKKR KEEEEAEDKEDDEDKDEDEEDEEDKEEDEEEDVPGQAHDEL
43	Human ERp57	ATGCAATTCAACTGGAACATCAAGACTGTTGCTTCCATCTTGT CCGCTTTGACTTTGGCTCAAGCTTCTGACGTTTTGGAGTTGAC TGACGACAACTTCGAGTCCAGAATTTCTGACACTGGTTCCGCT GGATTGATGTTGGTTGAGTTCTTCGCTCCATGGTGTGGTCATT GTAAGAGATTGGCTCCAGAATACGAAGCTGCTGCTACTAGAT TGAAGGGTATCGTTCCATTGGCTAAGGTTGACTGTACTGCTAA CACTAACACTTGTAACAAGTACGGTGTTTCCGGTTACCCAACT TTGAAGATCTTCAGAGATGGTGAAGAAGCTGGAGCTTACGAC GGTCCAAGAACTGCTGACGGTATCGTTTCCCACTTGAAGAAG CAAGCTGGTCCAGCTTCTGTTCCATTGAGAACTGAGGAGGAG TTCAAGAAGTTCATCTCCGACAAGGACGCTTCTATCGTTGGTT TCTTCGACGATTCTTTCTCTGAAGCTCACTCCGAATTCTTGAA GGCTGCTTCCAACTTGAGAGACAACTACAGATTCGCTCACACT AACGTTGAGTCCTTGGTTAACGAGTACGACGATAACGGTGAA GGTATCATCTTGTTCAGACCATCCCACTTGACTAACAAGTTCG AGGACAAGACAGTTGCTTACACTGAGCAGAAGATGACTTCCG GAAAGATCAAGAAGTTTATCCAAGAGAACATCTTCGGTATCT GTCCACACATGACTGAGGACAACAAGGACTTGATTCAGGGAA AGGACTTGTTGATCGCTTACTACGACGTTGACTACGAGAAGA ACGCTAAGGGTTCCAACTACTGGAGAAACAGAGTTATGATGG TTGCTAAGAAGTTCTTGGACGCTGGTCACAAGTTGAACTTCGC

		TGTTGCTTCTAGAAAGACTTTCTCCCACGAGTTGTCTGATTTC GGATTGGAATCCACTGCTGGAGAGATTCCAGTTGTTGCTATCA GAACTGCTAAGGGAGAGAAGTTCGTTATGCAAGAGGAGTTCT CCAGAGATGGAAAGGCTTTGGAGAGATTCTTGCAGGATTACT TCGACGGTAACTTGAAGAGATACTTGAAGTCCGAGCCAATTC CAGAATCTAACGACGGTCCAGTTAAAGTTGTTGTTGCTGAGA ACTTCGACGAGATCGTTAACAACGAGAACAAGGACGTTTTGA TCGAGTTTTACGCTCCTTGGTGTGGACACTGTAAAAACTTGGA GCCAAAGTACAAGGAATTGGGTGAAAAGTTGTCCAAGGACCC AAACATCGTTATCGCTAAGATGGACGCTACTGCTAACGATGTT CCATCCCCATACGAAGTTAGAGGTTTCCCAACTATCTACTTCT CCCCAGCTAACAAGAAGTTGAACCCAAAGAAGTACGAGGGA GGTAGAGAATTGTCCGACTTCATCTCCTACTTGCAGAGAGAG GCTACTAATCCACCAGTTATCCAAGAGGAGAAGCCAAAGAAG AAGAAGAAAGCTCACGACGAGTTGTAG
44	Human ERp57	MQFNWNIKTVASILSALTLAQASDVLELTDDNFESRISDTGSAGL MLVEFFAPWCGHCKRLAPEYEAAATRLKGIVPLAKVDCTANTN TCNKYGVSGYPTLKIFRDGEEAGAYDGPRTADGIVSHLKKQAGP ASVPLRTEEEFKKFISDKDASIVGFFDDSFSEAHSEFLKAASNLRD NYRFAHTNVESLVNEYDDNGEGIILFRPSHLTNKFEDKTVAYTEQ KMTSGKIKKFIQENIFGICPHMTEDNKDLIQGKDLLIAYYDVDYE KNAKGSNYWRNRVMMVAKKFLDAGHKLNFAVASRKTFSHELS DFGLESTAGEIPVVAIRTAKGEKFVMQEEFSRDGKALERFLQDYF DGNLKRYLKSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFY APWCGHCKNLEPKYKELGEKLSKDPNIVIAKMDATANDVPSPYE VRGFPTIYFSPANKKLNPKKYEGGRELSDFISYLQREATNPPVIQE EKPKKKKKAHDEL
45	hCRT-BstZ17I-HA/UP	GTATACCCATACGACGTCCCAGACTACGCTGAGCCCGCCGTCT ACTTCAAGGAGC
46	hCRT-PacI/LP	TTAATTAACTACAGCTCGTCATGGGCCTGGCCGGGGACATCTT CC
47	Synthetic peptide that binds CRT	KLGFFKR
48	Alpha amylase signal peptide	ATGGTTGCTT GGTGGTCCTT GTTCTTGTAC GGATTGCAAG TTGCTGCTCC AGCTTTGGCT

	(from *Aspergillus niger* α-amylase) (DNA)	
49	Alpha amylase signal peptide (from *Aspergillus niger* α-amylase)	MVAWWSLFLY GLQVAAPALA

While the present invention is described herein with reference to illustrated embodiments, it should be understood that the invention is not limited hereto. Those having ordinary skill in the art and access to the teachings herein will recognize additional modifications and embodiments within the scope thereof. Therefore, the present invention is limited only by the claims attached herein.

WHAT IS CLAIMED:

1. A lower eukaryote host cell comprising a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic.

2. The lower eukaryote host cell of Claim 1, wherein the nucleic acid molecule comprises an open reading frame encoding the Ca^{2+} ATPase operably linked to a heterologous promoter.

3. The lower eukaryote host cell of Claim 1, wherein the host cell further includes a nucleic acid molecule encoding a recombinant protein.

4. The lower eukaryote host cell of Claim 3, wherein the recombinant protein is an antibody.

5. The lower eukaryote host cell of Claim 1, wherein the function of at least one endogenous gene encoding a chaperone protein has been reduced, disrupted, or deleted; and a nucleic acid molecule encoding at least one mammalian homolog of the chaperone protein is expressed in the host cell.

6. The lower eukaryote host cell of Claim 1, wherein the host cell further includes a nucleic acid molecule encoding an ERp57 protein and/or a nucleic acid molecule encoding a calreticulin protein.

7. A method for producing a recombinant protein having reduced *O*-glycosylation comprising:

(a) providing a lower eukaryote host comprising a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic;

(b) introducing a nucleic acid molecule into the host cell encoding the recombinant protein: and

(c) growing the host cell under conditions suitable for producing the recombinant protein.

8. The method of Claim 7, wherein the nucleic acid molecule comprises an open reading frame encoding the Ca^{2+} ATPase operably linked to a heterologous promoter.

9. The method of Claim 8, wherein the recombinant protein is an antibody.

10. The method of Claim 7, wherein the function of at least one endogenous gene encoding a chaperone protein has been reduced, disrupted, or deleted; and a nucleic acid molecule encoding at least one mammalian homolog of the chaperone protein is expressed in the host cell.

11. The method of Claim 7, wherein the host cell further includes a nucleic acid molecule encoding an ERp57 protein and a nucleic acid molecule encoding a calreticulin protein.

12. A lower eukaryote host cell comprising a nucleic acid molecule encoding an ERp57 protein and/or a nucleic acid molecule encoding a calreticulin protein wherein expression of the proteins in the host cell is ectopic.

13. The lower eukaryote host cell of Claim 12, wherein the host cells further includes a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic.

14. The lower eukaryote host cell of Claim 12, wherein the host cell further includes a nucleic acid molecule encoding a recombinant protein.

15. The lower eukaryote host cell of Claim 14, wherein the recombinant protein is an antibody.

16. The lower eukaryote host cell of Claim 12, wherein the function of at least one endogenous gene encoding a chaperone protein has been reduced, disrupted, or deleted; and a nucleic acid molecule encoding at least one mammalian homolog of the chaperone protein is expressed in the host cell.

17. A method for producing a recombinant protein having reduced *O*-glycosylation comprising:

(a) providing a lower eukaryote host comprising a nucleic acid molecule encoding an ERp57 protein and/or a nucleic acid molecule encoding a calreticulin protein wherein expression of the proteins in the host cell is ectopic;

(b) introducing a nucleic acid molecule into the host cell encoding the recombinant protein: and

(c) growing the host cell under conditions suitable for producing the recombinant protein.

18. The method of Claim 17, wherein the host cells further includes a nucleic acid molecule encoding at least one endogenous or exogenous Ca^{2+} ATPase wherein expression of the Ca^{2+} ATPase in the host cell is ectopic.

19. The method of Claim 17, wherein the recombinant protein is an antibody.

20. The method of Claim 17, wherein the function of at least one endogenous gene encoding a chaperone protein has been reduced, disrupted, or deleted; and a nucleic acid molecule encoding at least one mammalian homolog of the chaperone protein is expressed in the host cell.