APPLIED AND ENVIRONMENTAL MICROBIOLOGY, May 2004, p. 2639–2646 Vol. 70, No. 5
0099-2240/04/$08.00+0 DOI: 10.1128/AEM.70.5.2639–2646.2004

In Vivo Synthesis of Mammalian-Like, Hybrid-Type N-Glycans in *Pichia pastoris*

Wouter Vervecken,† Vladimir Kaigorodov,† Nico Callewaert,†‡ Steven Geysens, Kristof De Vusser, and Roland Contreras*

Fundamental and Applied Molecular Biology, Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology, Ghent, Belgium

Received 18 November 2003/Accepted 29 January 2004

The *Pichia pastoris* N-glycosylation pathway is only partially homologous to the pathway in human cells. In the Golgi apparatus, human cells synthesize complex oligosaccharides, whereas *Pichia* cells form mannose structures that can contain up to 40 mannose residues. This hypermannosylation of secreted glycoproteins hampers the downstream processing of heterologously expressed glycoproteins and leads to the production of protein-based therapeutic agents that are rapidly cleared from the blood because of the presence of terminal mannose residues. Here, we describe engineering of the *P. pastoris* N-glycosylation pathway to produce nonhyperglycosylated hybrid glycans. This was accomplished by inactivation of *OCH1* and overexpression of an α-1,2-mannosidase retained in the endoplasmic reticulum and *N*-acetylglucosaminyltransferase I and β-1,4-galactosyltransferase retained in the Golgi apparatus. The engineered strain synthesized a nonsialylated hybrid-type N-linked oligosaccharide structure on its glycoproteins. The procedures which we developed allow glycan engineering of any *P. pastoris* expression strain and can yield up to 90% homogeneous protein-linked oligosaccharides.

Most protein-based therapeutic agents produced in heterologous expression systems are glycosylated, a modification that is crucial for correct folding, stability, and bioactivity of the protein and influences its pharmacokinetic properties, such as tissue distribution and blood clearance. Glycoproteins with terminal sialic acids on their glycans persist longer in the blood than glycoproteins with terminal galactose, *N*-acetylglycosamine, or mannose residues because the latter compounds are cleared rapidly via receptors in the liver and on reticuloendothelial cells (e.g., the asialoglycoprotein receptor and the mannose receptor) (10, 20, 30, 31). In addition, glycan structures produced in nonhuman cells can cause immune reactions, as exemplified by the reaction against xenografts of porcine origin; these reactions are primarily caused by the presence of α-galactose on the glycoproteins (7). Another example is the immune reaction against glycoproteins from yeast, which results from the presence of α-1,3-mannose, β-linked mannose, and/or phosphate residues in either a phosphomonoester or phosphodiester linkage (1, 32). Consequently, recombinant glycoproteins produced for therapeutic applications should be expressed in heterologous hosts that produce protein-linked oligosaccharides that closely resemble those of humans.

One of the advantages of using a yeast system for production of glycoproteins is the ability of such a system to perform Asn-linked glycosylation, a process that does not occur in *Escherichia coli*. The N-glycosylation pathway of *Pichia pastoris* has been reviewed recently by Bretthauer and Castellino (3). This methylotrophic yeast is an excellent candidate for production of protein-based biopharmaceutical agents (12).

Yeast N glycosylation has been studied mostly in baker's yeast mutants defective in glycosylation. The studies revealed that the yeast N-glycosylation pathway mirrors the pathway in typical mammalian cells up to the point where $Man_8GlcNAc_2$ N-glycosylated proteins exit from the endoplasmic reticulum (ER) and are transported to the early Golgi apparatus. In the mammalian Golgi apparatus, the $Man_8GlcNAc_2$ glycans are trimmed to $Man_5GlcNAc_2$, and then they are further modified into complex glycans (18). The first elongation step is the addition of GlcNAc, which is catalyzed by β-*N*-acetylglucosaminyltransferase I (GnTI). Further processing is performed by a number of glycan-processing enzymes. This includes trimming of two mannose residues by mannosidase II, addition of GlcNAc by several GlcNAc transferases, addition of galactose by galactosyltranferases, and sialylation by sialyltransferases. These modifications result in a large variety of possible N-glycan structures in mammalian cells, as described by Fukuda (15).

In contrast, $Man_8GlcNAc_2$ glycans in yeast are not trimmed but are generally elongated further, sometimes forming hypermannosyl glycans that constitute the bulk of the glycoprotein's molecular mass. Hypermannosylation is initiated by the α-1,6-mannosyltransferase Och1p. The α-1,6-mannose residue is then elongated by several mannosyltransferases and phosphomannosyltransferases to produce the large hypermannosyl structures (14).

As described above, the presence of nonhuman glycan structures on therapeutic glycoproteins can be detrimental to bioactivity and safety due to the faster blood clearance of glycoproteins with terminal mannoses and the potential immunogenicity of nonhuman glycans. Furthermore, hyperman-

* Corresponding author. Mailing address: Fundamental and Applied Molecular Biology, Technologiepark 927, B-9052 Ghent-Zwijnaarde, Belgium. Phone: 32 9 331 3631. Fax: 32 9 331 3502. E-mail: Roland.Contreras@dmbr.UGent.be.

† W.V., V.K., and N.C. contributed equally to this work.

‡ Present address: Institute of Microbiology, Department of Biology, ETH Zürich, CH-8092 Zürich, Switzerland.

nosylation leads to a high degree of heterogeneity that can hamper downstream processing of recombinant glycoproteins and may complicate complete molecular characterization of the molecules.

As expected, inactivation of the *OCH1* gene in *Saccharomyces cerevisiae* eliminates hypermannosylation, but unfortunately it is also severely detrimental to yeast viability (9). Moreover, inactivation of the *S. cerevisiae OCH1* gene does not lead to a homogeneous N-glycan pattern. The phosphomannosyltransferase Mnn6p and the α-1,3-mannosyltransferase Mnn1p both modify the $Man_8GlcNAc_2$ core glycan. Thus, a triple-knockout strain is needed to obtain a homogeneous $Man_8GlcNAc_2$ pattern, and the viability of such a strain is severely impaired. Further introduction of an ER-retained α-1,2-mannosidase led to production of only about 20% $Man_5GlcNAc_2$ on the yeast glycoprotein carboxypeptidase Y (8). Thus, so far N-glycan homogenization in baker's yeast has been an inefficient process, and the strains obtained might be too weak for general use in glycan engineering and subsequent glycoprotein production.

Nevertheless, the previous studies paved the way for N-glycan engineering in fungi. With the information described above in hand, we focused on developing a simple and effective tool for homogenizing and humanizing N glycosylation in the methylotrophic yeast *P. pastoris*, which is widely used in biotechnology. The steps used included inactivation of *OCH1* and overexpression of an ER-retained HDEL-tagged α-1,2-mannosidase and two chimeric glycosyltransferases, Kre2-GnTI and Kre2–β-1,4-galactosyltransferase (Fig. 1).

MATERIALS AND METHODS

Vector construction and transformation. All recombinant DNA procedures were performed by using standard methods (27). Chemicals and enzymes were obtained from major suppliers.

A *Pichia* open reading frame (ORF) homologous to *OCH1*, available from the GenBank database (accession no. E12456), was PCR cloned in pUC18 from genomic DNA of strain GS115 (Invitrogen) by using the following primers: 5′AGATCTGCCTGACAGCCTTAA3′ and 5′GTCGACGTTCTTGTCAATTC G3′. The identity of the amplified DNA fragment was confirmed by sequencing.

Plasmid pGlycoSwitchM8 (Fig. 2) contains a fragment of the *P. pastoris OCH1* ORF corresponding to the protein sequence from Ala25 to Ala155 that is inserted between the BglII and HindIII sites of pPICZB (Invitrogen). Two stop codons are located in frame just before codon Ala25 to prevent the possible synthesis of a truncated protein resulting from a potential start of transcription from a potential cryptic promoter activity located upstream of the Ala25 codon. The BstBI site of the polylinker of pPICZB was eliminated by digestion, filling in, and self-ligation. The unique BstBI site located in the cloned *OCH1* fragment can be used for linearization (Fig. 2 shows an overview of the *OCH1* inactivation strategy).

pPIC9MFManHDEL (9,535 bp) was constructed by inserting an EcoRI-NotI fragment of pGAPZMFManHDEL (5) into pPIC9 (Invitrogen) opened with the same restriction enzymes.

pGlycoSwitchM5 (5,485 bp) (Fig. 3A) was constructed as follows: (i) an XbaI/ClaI fragment of pPIC9, containing the *P. pastoris HIS4* transcriptional terminator sequence, was inserted between the HindIII and EcoRI sites of pGlycoSwitchM8; and (ii) the 2.3-kb BglII-NotI fragment of pGAPZMFManHDEL (5), containing the *GAP* promoter and the preMFmannosidaseHDEL cassette, was inserted between the HindIII and NotI sites. All restriction sites used for this construction (except the NotI site) were filled in with the Klenow DNA polymerase. The unique BstBI site in pGAPZMFmanHDEL had been eliminated as described above.

In order to target the human GnTI to the Golgi apparatus, the GnTI N-terminal portion was replaced by the *S. cerevisiae* Kre2 N-terminal portion that is responsible for localization in the yeast Golgi apparatus (21). Plasmid YEp352Kre2 (kindly provided by Howard Bussey, McGill University, Montreal, Canada) contained the *KRE2* gene cloned with SacI and PvuII in a SalI (Klenow polymerase blunted)- and SacI-opened YEp352 vector. This plasmid was digested with SacI and PvuI and was blunted with T4 polymerase. The 5′ end of the gene was isolated and cloned in a Klenow polymerase-blunted SgrAI- and XbaI-opened pUChGnTI (23) vector. The site of ligation of the two DNA fragments was sequenced by standard procedures. The resulting Kre2-GnTI ORF, containing the N-terminal part of the *KRE2* gene (the first 100 amino acids) and the catalytic domain of GnTI (the last 327 amino acids), was isolated by EcoRV/HindIII double digestion and ligated into a SalI- and EcoRI-linearized pPIC6A vector (Invitrogen) after both the vector and the insert were blunted with Klenow polymerase. The resulting plasmid was designated pPIC6Kre2GnTI (Fig. 3B). This plasmid contained the Kre2-GnTI ORF under control of the methanol-inducible *AOX1* promoter and the *BSD* gene from *Aspergillus terreus* for resistance to the antibiotic blasticidin (Fig. 3B).

Human β-1,4-galactosyltransferase was localized by fusion of the catalytic domain to the N-terminal portion of Kre2p in the same way that was used to target GnTI. The human β-1,4-galactosyltransferase 1 cDNA was amplified from a HepG2 cDNA library by using oligonucleotides 5′TTCGAAGCTTCGCTAG CTCGGTGTCCCGATGTC3′ and 5′GAATTCGAAGGGAAGATGAGGCTT CGGGAGCC3′ as starter sequences. The amplified fragment was cloned with HindIII and EcoRI into pUC18. To remove the N-terminal portion of β-1,4-galactosyltransferase, a PCR was performed with oligonucleotide primers 5′TT CGAAGCTTCGCTAGCTCGGTGTCCCGATGTC3′ and 5′CGTTCGCGAC CGGAGGGGCCCGGCCGCC3′. The amplified fragment was cut with NruI and HindIII and ligated into the SgrAI (Klenow blunted)- and HindIII-opened pUCKre2GnTI vector. The resulting Kre2–β-1,4-galactosyltransferase fusion construct was amplified by performing PCR with primers 5′TCGATATCAAG CTTAGCTCGGTGTCCCGATGTC3′ and 5′GAATTCGAACTTAAGATGG CCCTCTTTCTCAGTAAG3′. The amplified fragment was cloned with EcoRV/BstBI into the pBLURA IX vector (6) (kindly provided by James Cregg, Keck Graduate Institute of Applied Life Sciences, Claremont, Calif.). Finally, the *URA3* gene was replaced by a kanamycin resistance cassette by ligating a SpeI-SmaI fragment from the vector pFA6a-KanMX4 into the SpeI- and SspI-opened plasmid. The final plasmid was designated pBlKanMX4Kre2hGalT (Fig. 3C).

The plasmids were transformed in *Pichia* strain GS115 (Invitrogen) as described previously (13). Correct genomic integration at the *P. pastoris OCH1* locus was confirmed by performing PCR with genomic DNA.

Protein preparation. The secreted version of the *Trichoderma reesei* α-1,2-mannosidase was purified by using a combination of hydrophobic interaction chromatography, anion-exchange chromatography, and gel filtration chromatography, as described previously (22, 29). All sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) experiments were performed with 10% polyacrylamide gels under standard running conditions. Yeast cell wall mannoproteins were released as described by Ballou (1). The protocol involved extensively washing yeast cells with 0.9% NaCl and water, autoclaving them for 90 min in 20 mM sodium citrate (pH 7), and then precipitating the released proteins with 4 volumes of methanol. Due to the simplicity of this protocol, we found that the derived mannoprotein preparations were well suited for phenotypic glycan analysis of at least 15 yeast clones obtained after each glycoengineering transformation. In this way, fast and direct data for determination of the efficiency of each manipulation step were obtained.

N-glycan analysis. We performed laser-induced, DNA sequencer-assisted (DSA), fluorophore-assisted carbohydrate electrophoresis (FACE) with an ABI 377 DNA sequencer as described previously (4). Glycoproteins were denatured in RCM buffer (8 M urea, 360 mM Tris [pH 8.6], 3.2 mM EDTA) at 50°C and were subsequently immobilized on a multiscreen Immobilon-P plate. This plate contained at the bottom of each well a polyvinylidene difluoride membrane to which the glycoprotein bound. Liquid could be withdrawn by applying a mild vacuum to the bottom with a multiscreen vacuum manifold. The glycoproteins were reduced with 0.1 M dithiothreitol, and the SH groups were carboxymethylated by using 0.1 M iodoacetic acid. The unoccupied places on the Immobilon-P plate were saturated with 1% polyvinylpyrrolidone 360. N-glycans were released from the glycoproteins by treatment with peptide: *N*-glycosidase F (100 U in 20 μl 10 mM Tris-acetate [pH 8.3]; New England BioLabs). N-glycans were recovered and derivatized with the fluorophore 8-aminopyrene-1,3,6-trisulfonate by reductive amination. Excess label was removed by size fractionation on Sephadex G-10 resin by using a multiscreen format procedure. After concentration of the 1-aminopyrene-3,6,8-trisulfonate-labeled oligosaccharides by evaporation, a ROX-labeled Genescan 500 standard mixture (Applied Biosystems) was added for internal standardization. This mixture was examined with an ABI 377A DNA sequencer (Applied Biosystems) by using a 12% polyacrylamide gel and a buffer containing 89 mM Tris, 89 mM borate, and 2.2 mM EDTA. Glycans were detected at the end of the gel with an Ar laser. On each gel, N-glycans of bovine RNase B and a maltodextrose ladder were included as references. Data analysis was performed by

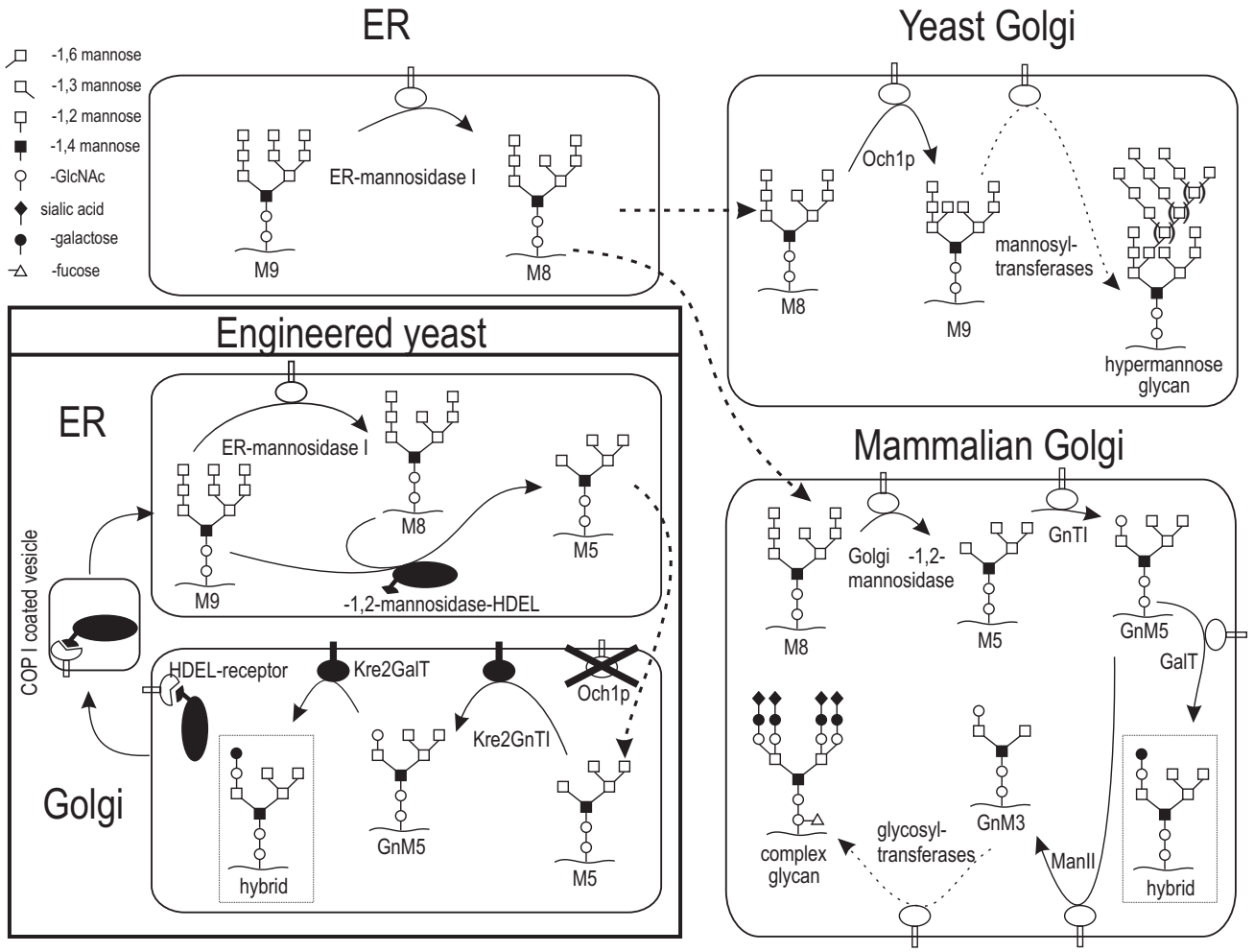

FIG. 1. Overview of yeast and mammalian N-linked glycosylation and the strategy used for humanization of *P. pastoris* glycosylation. In the nonglycoengineered ER $Man_9GlcNAc_2$ (M9) is trimmed to $Man_8GlcNAc_2$ (M8) by removal of one α-1,2-mannose residue, after which the glycoproteins are transferred to the *cis*-Golgi apparatus. In the nonengineered yeast Golgi apparatus N-glycans are further elongated by transfer of one α-1,6-mannose residue by the initiating α-1,6-mannosyltransferase (Och1p). These N-glycans can be further enlarged by several mannosyl- and phosphomannosyltransferases to obtain hypermannose structures. In the mammalian Golgi apparatus, the $Man_8GlcNAc_2$ structure is shortened to $Man_5GlcNAc_2$ before addition of a GlcNAc residue by GnTI. Subsequently, two mannose residues are removed by mannosidase II (ManII). The resulting $GlcNAcMan_3GlcNAc_2$ structure is further modified by several glycosyltransferases, resulting in a complex-type N-glycan. Alternatively, $GlcNAcMan_5GlcNAc_2$ (GnM5) can be elongated by β-1,4-galactosyltransferase (GalT) to produce a hybrid structure. The lower left panel shows the strategy used and the objectives for glycoengineering of *P. pastoris*. The enzymes introduced are indicated by solid ellipses. An HDEL-tagged α-1,2-mannosidase is targeted to the ER, trimming all α-1,2-mannose residues. This mannosidase is recognized by the HDEL receptor in the Golgi apparatus and is transported back to the ER by COP I-coated vesicles. Inactivation of *OCH1* prevents elongation of the oligosaccharides to hypermannose structures, as the presence of an initiating α-1,6-mannose residue is a prerequisite for further elongation. Further modification occurs by overexpression of Golgi apparatus-localized GnTI and β-1,4-galactosyltransferase. This localization is accomplished by fusion of the catalytic parts of the two enzymes to the sequence that is responsible for targeting of the α-1,2-mannosyltransferase (Kre2p) to the Golgi apparatus of *S. cerevisiae.* The hybrid structure that is finally obtained is enclosed in a box.

using the GENESCAN 3.1 software (Applied Biosystems). Exoglycosidase treatment of labeled glycans with β-*N*-acetylhexosaminidase (Glyko) and β-galactosidase (Prozyme) was performed overnight at 37°C in 20 mM sodium acetate (pH 5.0).

Classical FACE (8-aminonaphthalene-1,3,6-trisulfonate labeling of N-glycans and electrophoresis on 30% polyacrylamide minigels) was performed as described by Jackson (17). The DSA-FACE method is characterized by very high resolution and sensitivity, whereas classical FACE is well suited for detecting very complex mixtures of high-molecular-weight N-glycans (hyperglycosylation), which are not resolved and therefore form an easily recognizable smear. Thus, a combination of the two methods gives a more complete picture of the characteristics of yeast-produced glycoproteins.

Growth curve determination. Logarithmically growing yeast cultures were diluted with fresh YPD medium to an optical density at 600 nm (OD_{600}) of 0.02 and grown overnight at 30°C with shaking at 250 rpm (12 h; OD_{600}, <3.0). To start the experiment, 10-ml portions of fresh YPD medium in 50-ml polypropylene tubes were inoculated with overnight yeast cultures to obtain a starting OD_{600} of 0.5, and the OD_{600} was determined every 2 h. All yeast strains were examined in parallel in the same experiment.

RESULTS

Inactivation of *OCH1*. A *P. pastoris* gene homologous to *S. cerevisiae OCH1* has been described in a Japanese patent. It is unfortunate that the results described in the patent have not appeared in the peer-reviewed literature. Nevertheless, we used the sequence that is available in the GenBank database in

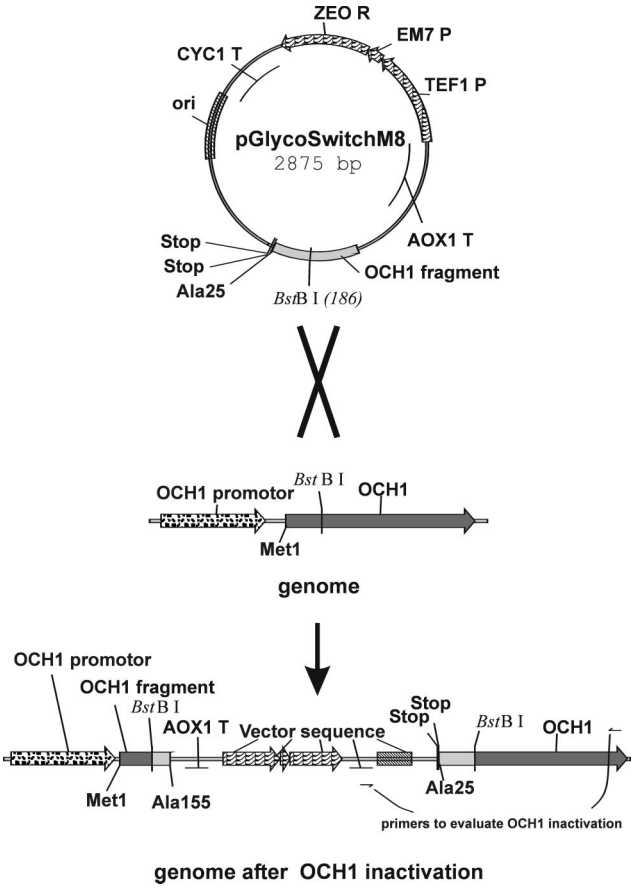

FIG. 2. *OCH1* inactivation vector. Upon digestion of pGlycoSwitchM8 with BstBI and transformation in *P. pastoris*, the construct integrates at the *OCH1* locus. This results in a short *OCH1* fragment that does not translate to a functional *OCH1* gene and a promotorless fragment that cannot be translated because of the absence of a promoter and the presence of two in-frame nonsense codons.

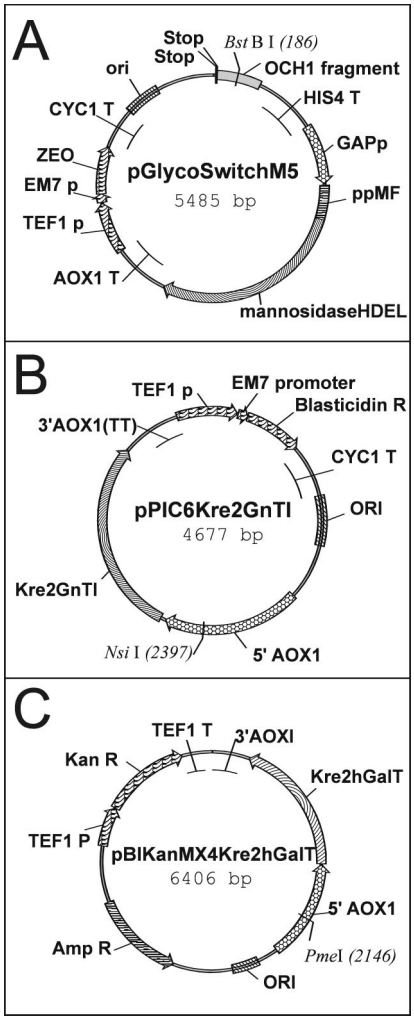

FIG. 3. Plasmids used for glycan engineering of *P. pastoris*.

our attempts to knock out the gene by double homologous recombination. However, it was apparent from our results that double homologous recombination at this *Pichia* locus is at best a rare event. We were unable to obtain a strain with the targeting construct correctly integrated, even with homologous regions that contained up to 3 kb of isogenic genomic DNA. The targeting strategies tested included prescreening for temperature sensitivity at 37°C and screening for vanadate resistance (both hallmarks of severe N-glycosylation impairment in *S. cerevisiae*), as well as the use of different positive selection markers (Zeocin, G418, *S. cerevisiae* invertase). Several hundred clones were screened without success. We concluded that double homologous recombination targeting of the *Pichia OCH1* locus is a highly inefficient procedure, if it is possible at all. This inefficiency would hamper the use of any resulting N-glycan engineering vector for routine introduction into optimized heterologous protein-expressing strains.

We therefore decided to use the known stability of genomic arrangements resulting from knockin strategies in *Pichia*. Knockin of protein expression vectors in *Pichia* is routinely done at the *AOX1* or *HIS4* locus, and the result is duplication of the homologous regions on the expression vectors. The fact that this genomic arrangement can be sufficiently stable in *Pichia* is exemplified by the expression vector for human serum albumin, which has been stably integrated for over 10 years under selective conditions on the laboratory scale, and in the absence of selection in up to 100,000 liters of culture (Bennett Cohen, personal communication). Therefore, we generated a vector designated pGlycoSwitchM8 (Fig. 2) that included bp 73 to 467 of the *P. pastoris* ORF homologous to *OCH1*, preceded by two in-frame nonsense codons to avoid read-through from potential earlier translation start sites in the vector. This fragment contained a centrally located BstBI site which was convenient for vector linearization before transformation and was 3′ linked to the *AOX1* transcription terminator sequence. Upon integration into the genomic *OCH1* locus of *Pichia*, this vector duplicated the *OCH1* sequence present in the vector. The resulting first copy of *OCH1* gave a translation product that was at most 161 amino acids long (6 amino acids of which were from vector sequences). This translation product did not contain the catalytic domain of this type II transmembrane protein. The second copy of *OCH1* lacked the coding sequence for the first 25 amino acids, and two in-frame stop codons

prevented any read-through from potential upstream translation initiation sites. This strategy should have eliminated the activity of the *OCH1* gene.

The vector was transformed in strain GS115, and the selected transformant was designated GlycoSwitchM8 or the *och1* strain. PCR performed with genomic DNA and the primer combinations shown in Fig. 2 showed that there was correct integration of this construct at the expected genomic locus in about 50% of the Zeocin-resistant transformants (average of three independent experiments). This frequency resembled that of knockin of expression vectors in either *AOX1* or *HIS4* and was certainly high enough for practical routine use.

Analysis of the cell wall mannoprotein N-glycans revealed a change in the glycosylation pattern. Whereas the $Man_9GlcNAc_2$ peak was the predominant peak for wild-type strain GS115, for the GlycoSwitchM8 strain the main peak was the $Man_8GlcNAc_2$ peak (Fig. 4, panels 2 and 3). This phenotype reverted after retransformation of the strain with the original *OCH1* ORF (results not shown).

To evaluate whether the heterogeneity of secreted glycoproteins was decreased in the GlycoSwitchM8 strain, we heterologously expressed the hyperglycosylated *T. reesei* α-1,2-mannosidase as a secreted protein in wild-type strain GS115 and the GlycoSwitchM8 strain, as described by Maras et al. (22). The supernatant was separated by SDS-PAGE (Fig. 5A). The gel revealed that the smear was absent from the proteins produced in the GlycoSwitchM8 strain. In parallel, the glycoprotein was deglycosylated by PNGase F treatment, and the glycans were analyzed by FACE on minigels. In this system the large hyperglycosyl structures were not resolved and appeared as one smeared band (Fig. 5B). This smeared band was absent from the *och1* strain, showing that heterogeneity decreased.

Expression of ER-retained mannosidase-HDEL. To further humanize the N-glycans of *P. pastoris*, ER-retained *T. reesei* α-1,2-mannosidase-HDEL was expressed in the *och1* strain. Construction of expression vectors and preliminary analyses have been described previously by us.

For easy conversion of a *P. pastoris* expression strain, the mannosidase-HDEL was introduced into the *och1* inactivation vector. This combination vector was designated pGlyco SwitchM5. Correct integration of the vector was evaluated by PCR analysis, and mannoprotein N-glycans were analyzed by DSA-FACE. The glycan profile revealed a homogeneous $Man_5GlcNAc_2$ peak (Fig. 4, panel 4). Integration of the $Man_5GlcNAc_2$ peak and of all the small peaks above the detection limit of this method (signal/noise, >3) in the size range from 5 to 25 glucose units revealed that this higher-eukaryote type of high-mannose glycan accounted for at least 90% of the total N-glycan pool present in the mixture.

In an alternative approach, mannosidase-HDEL was expressed under control of the methanol-inducible promoter *AOX1*. No apparent differences in the N-glycan profiles of the two mannosidase-expressing strains (i.e., constitutive and inducible) could be detected.

To confirm N-glycan modification on a heterologously expressed glycoprotein, we transformed the pGlycoSwitchM5 plasmid into a *Pichia* strain expressing the *trans*-sialidase of *Trypanosoma cruzi* (19). In this case, $Man_5GlcNAc_2$ was detected on the purified protein and accounted for more than 95% of the total N-glycans (results not shown).

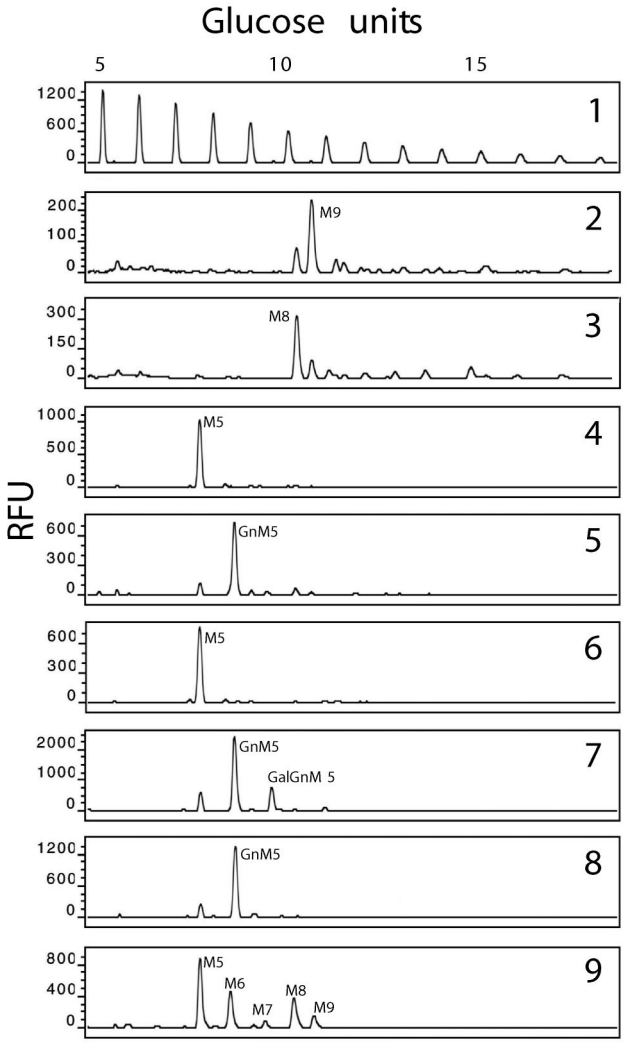

FIG. 4. DSA-FACE analysis of N-glycans from different glycan-engineered *P. pastoris* strains. Panel 1 shows the results for a maltodextrose reference. Panels 2 through 9 show the results for N-glycans, as follows: panel 2, wild-type strain GS115 (the main peak is $Man_9GlcNAc_2$ [M9]); panel 3, strain with *och1* inactivated (the main peak is $Man_8GlcNAc_2$ [M8]); panel 4, strain with *och1* inactivated expressing mannosidase-HDEL (the main peak is $Man_5GlcNAc_2$ [M5]); panel 5, strain with *och1* inactivated expressing mannosidase-HDEL and Kre-GnTI (the main peak is GlcNAc $Man_5GlcNAc_2$ [GnM5]); panel 6, same as panel 5 but the glycans were treated with β-N-acetylhexosaminidase (the GlcNAc $Man_5GlcNAc_2$ peak shifted to the $Man_5GlcNAc_2$ position, indicating that terminal GlcNAc was present); panels 7 and 8, strain with *och1* inactivated expressing mannosidase-HDEL, Kre2-GnTI, and Kre2-β-1,4-galactosyltransferase (the additional peak is GalGlcNAc $Man_5GlcNAc_2$ [GalGnM5]); when the culture was treated with β-galactosidase, the small peak disappeared, as shown in panel 8); panel 9, reference glycans from bovine RNase B ($Man_{5-9}GlcNAc_2$ [M5-M9]). RFU, relative fluorescence units.

Growth curve analysis of the engineered strain in a shake flask culture demonstrated that its doubling time was very similar to that of the wild-type strain. However, the glycoengi-

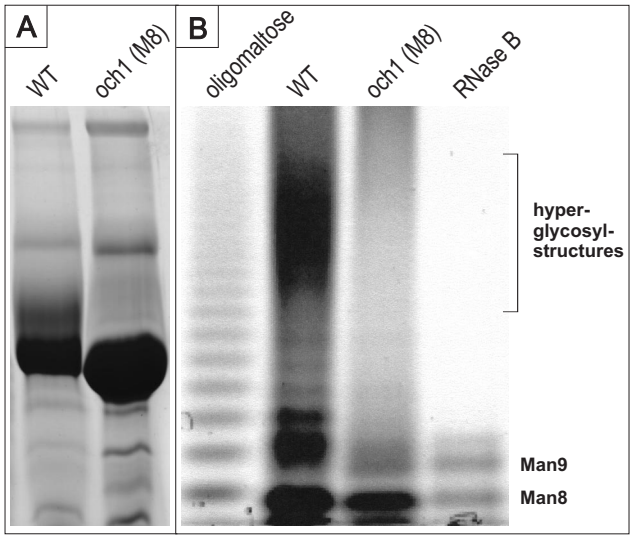

FIG. 5. Evaluation of hyperglycosylation after inactivation of *P. pastoris OCH1*. (A) Coomassie brilliant blue-stained SDS-PAGE gel containing supernatants of *P. pastoris* strains secreting *T. reesei* mannosidase. For the nonengineered strain (WT) a clear smear is visible, whereas this smear is not present for the strain with *och1* inactivated [och1 (M8)]. (B) FACE analysis of N-glycans derived from mannosidase secreted by a nonengineered strain (WT) and a strain with *och1* inactivated [och1 (M8)]. The bands with greater electrophoretic mobility are the Man8 and Man9 bands and represent core N-glycan structures. The hyperglycosyl structures are slowly migrating sugars. They are not present in the strain with *och1* inactivated.

neered strain reached the stationary phase at an optical density that was about 20% lower than that of the wild-type strain, possibly due to somewhat higher sensitivity to the stress of high cell density. Nevertheless, this phenotypic difference was much less pronounced in *Pichia* than in the *S. cerevisiae och1* strain.

Expression of Golgi apparatus-localized GnTI (Kre2-GnTI). The first step in the maturation of the N-glycans into complex oligosaccharides is the addition of a GlcNAc residue by GnTI. To act on endogenously synthesized glycans, GnTI must be targeted to the Golgi apparatus. Previous results for expression of rat GnTI in *S. cerevisiae* indicated that the mammalian Golgi retention signal does not function in yeast, since GnTI activity was observed all over the cell (33). Lussier et al. (21) studied the localization of the α-1,2-mannosyltransferase Kre2p in *S. cerevisiae*. This type II membrane protein is localized in the early to medial Golgi compartments. The protein's N terminus (which is in the cytosol), its transmembrane region, and parts of the luminal stem region together are responsible for this localization. Therefore, we replaced the N-terminal part of GnTI, including the cytosolic part, the transmembrane region, and part of the luminal stem region, with this yeast's Kre2 signal sequence. This resulted in a chimeric protein whose first 100 amino acids were from Kre2p and last 327 amino acids were from GnTI. Indeed, recent results indicated that a chimera of Kre2p and human β-1,4-galactosyltransferase was retained as an active enzyme in the yeast Golgi apparatus (28).

For expression in *P. pastoris*, the Kre2-GnTI fusion construct was placed under control of the strong methanol-induced *AOX1* promoter in a plasmid containing the blasticidin resistance marker. The construct was transformed after linearization in the *AOX1* sequence with NsiI in a GS115 *och1* strain, expressing α-1,2-mannosidase-HDEL. The presence of the construct was confirmed by PCR performed with genomic DNA and the *AOX1* 3′ and 5′ primers.

N-glycans of mannoproteins of several transformants were analyzed by DSA-FACE. The dominant peak was about one glucose unit larger than the $Man_5GlcNAc_2$ peak (Fig. 3, panel 5). To determine whether this peak had a terminal GlcNAc, exoglycosidase digestion was performed with β-N-acetylhexosaminidase, an enzyme that hydrolyzed β-GlcNAc linkages and caused the peak to shift back to the $Man_5GlcNAc_2$ peak position (Fig. 3, panel 6). The shift indicated that the original peak represented $GlcNAcMan_5GlcNAc_2$ and thus the correct in vivo activity of the chimeric GnTI enzyme.

Interestingly, overexpression of the Kre2-GnTI chimera led to almost complete conversion of $Man_5GlcNAc_2$ to GlcNAc $Man_5GlcNAc_2$, indicating that the Golgi UDP-GlcNAc concentration is probably not limiting.

Expression of Golgi apparatus-retained β-1,4-galactosyltransferase. The next step in obtaining a hybrid human-like N-linked glycan profile was overexpression of β-1,4-galactosyltransferase. For this enzyme to be active in the secretory pathway, it must be localized in the yeast Golgi apparatus. The success of Golgi apparatus-targeted expression of GnTI in which the yeast Kre2p N-terminal domain was used as a localization signal prompted us to use the same N-terminal part of Kre2p to target β-1,4-galactosyltransferase to the Golgi apparatus. Thus, the N-terminal part of human β-1,4-galactosyltransferase 1 (the first 77 amino acids), including the transmembrane domain and the cytosolic part of the enzyme, was replaced by the same part of Kre2p as described above. The resulting chimera was placed under control of the *AOX1* promoter. The 3′ end of *AOX1* served as a terminator. Kanamycin was used as a selection marker for transformation of the PmeI-linearized plasmid.

It has been suggested that UDP-Gal (the donor substrate for β-1,4-galactosyltransferase) is present in the Golgi apparatus of *S. cerevisiae* (26). The possible galactosyl transfer to the glycans of *Pichia* could provide indirect proof that the same thing is true for this yeast.

Mannoprotein N-glycans from several transformants were analyzed. An additional peak that was about one glucose unit larger than the $GlcNAcMan_5GlcNAc_2$ peak was detected in the transformants but not in the nontransformed strain (Fig. 4, panel 7). In parallel with attempts to prove that GlcNAc $Man_5GlcNAc_2$ was present in the Kre2-GnTI-expressing strain, the N-glycans were digested with β-galactosidase to ensure that the additional peak contained terminal β-galactose. After digestion of the glycans, the additional peak was indeed shifted back to the position of $GlcNAcMan_5GlcNAc_2$ (Fig. 4, compare panels 7 and 8). The amount of GalGlcNAc $Man_5GlcNAc_2$ could be determined by integrating the $GlcNAcMan_5GlcNAc_2$ peak before and after β-galactosidase digestion. Subtraction analysis of these two peaks revealed that about 10% of $GlcNAcMan_5GlcNAc_2$ was converted to $GalGlcNAcMan_5GlcNAc_2$. An attempt to increase this amount by supplementing the medium with 0.2% galactose did not increase the amount of Gal-containing oligosaccharides.

DISCUSSION

In this paper, we describe four steps for humanization of the N-glycans of *P. pastoris*. In the first step, we developed a vector that, after a simple and highly efficient transformation step, inactivated the *OCH1* gene and led to elimination of hyperglycosylation. The *OCH1* inactivation vector which we used differed from those described previously in that the inactivation of *OCH1* was accomplished via an efficient single homologous recombination event, whereas Choi et al. (11) and Murakami and Sugio (25) used a DNA construct that involved a more difficult double homologous recombination to inactivate the *OCH1* gene. Second, an HDEL-tagged α-1,2-mannosidase, which has been described previously by us (5), was overexpressed in the strain with inactivated *OCH1*. This led to almost complete conversion to $Man_5GlcNAc_2$. To facilitate glycoengineering, we developed a vector that combined the cassettes for *OCH1* inactivation and mannosidase expression in one plasmid, which allowed production of a homogeneous $Man_5GlcNAc_2$ structure after one transformation step.

It has been shown previously by us (24) that coexpression of the *T. reesei* α-1,2-mannosidase fused to the N-terminal part of *S. cerevisiae* Mns1p and the hyperglycosylated neuraminidase did not reduce hyperglycosylation of neuraminidase. In contrast, the relative amount of high-molecular-weight hypermannose structures increased. It was speculated that this was due to trimming of the high-mannose structures, which would have resulted in the availability of more substrate for elongation of mannosyltransferases. However, this explanation is speculative, as Choi et al. (11) used the same Mns1p signal sequence to direct a *Caenorhabditis elegans* mannosidase successfully to the ER, resulting in $Man_5GlcNAc_2$ on the glycoproteins. The two situations, however, are not completely comparable, because Choi et al. worked with a strain with a defective *OCH1* gene.

For the strains described here, we used the same HDEL-tagged mannosidase which was described previously (5). However, Callewaert et al. (5) reported incomplete mannose trimming, whereas in this study all N-glycans were trimmed to $Man_5GlcNAc_2$ by the same mannosidase construct. It has been suggested that in a strategy in which HDEL-tagged mannosidase is used, trimming of α-1,2-mannose residues is performed extracellularly (2), as some mannosidase was found in the medium as a result of oversaturation of the HDEL receptor. This possibility is not supported by the observation that there was complete conversion of $Man_5GlcNAc_2$ to GlcNAc $Man_5GlcNAc_2$ upon overexpression of GnTI. The latter step cannot occur extracellularly in the medium, because it is not sufficient to have only the oligosaccharide substrate and the catalyzing enzyme; the donor substrate UDP-GlcNAc also must be present, and sufficient amounts of this compound have never been observed in the medium. This demonstrates that the mannose trimming occurs intracellularly.

In two very recent studies by a group at Dartmouth College (11, 16), a similar method for glycan engineering in *P. pastoris* was developed. A strain with an inactivated *OCH1* gene was obtained by using a double-homologous-recombination strategy. In accordance with our results, no severe growth defects were observed. However, analysis of the N-glycans of a heterologously expressed protein revealed only a minor $Man_8GlcNAc_2$ peak, whereas in our system this structure was very abundant. Surprisingly, there were also quantitative and qualitative differences in the oligosaccharide compositions of the nonmodified strains. In our untransformed strain (GS115) mainly $Man_{8\text{-}9}GlcNAc_2$ and small amounts of $Man_{10\text{-}15}GlcNAc_2$ were present (in the size range of 5 to 20 glucose units), whereas Choi et al. (11) reported that $Man_{10}GlcNAc_2$ is the smallest glycan. The differences in oligosaccharide composition may have been due to the strains used (GS115 and JC308) or the nature of the expressed glycoprotein.

The strain with the inactivated *och1* gene expressing mannosidase-HDEL (M5 strain) is a starting point for further humanization of the N-glycosylation pathway, by introducing human GnTI and human β-1,4-galactosyltransferase. In experiments with *S. cerevisiae* described by Yoshida et al. (33), a heterologously expressed rat GnTI was found all over the cell. This prompted us to use a yeast-specific localization signal to target GnTI and β-1,4-galactosyltransferase to the Golgi apparatus.

We modified the *Pichia* N-glycosylation pathway by overexpression of Kre2-GnTI to produce almost homogeneous $GlcNAcMan_5GlcNAc_2$, as deduced from the sensitivity of the structure to β-*N*-acetylhexosaminidase. The complete conversion to $GlcNAcMan_5GlcNAc_2$ is somewhat in contrast to what was found by Choi et al. (11). In the latter study, human GnTI was fused to the N-terminal part of yeast Mnn9p and was expressed in an *och1* knockout strain that concomitantly expressed an α-1,2-mannosidase targeted to the Golgi apparatus by the *S. cerevisiae* Mnn10p N-terminal sequence. Complete conversion was observed only after introduction of a *Kluyveromyces lactis* UDP-GlcNAc transporter. The difference could be due to the fact that different localization signals were used for mannosidase and GnTI activities, and strain differences could also have played a role.

We expressed a Kre2–β-1,4-galactosyltranferase chimera to obtain a hybrid structure. However, conversion of the $GlcNAcMan_5GlcNAc_2$ acceptor to its galactosylated derivative was incomplete. This could have been due to a number of factors. The Kre2p localization sequence may not have been optimal for the galactosyltransferase. It has been shown that in mammalian cells galactosyltranferase is localized later in the secretory pathway than GnTI is localized. It is also possible that the amount of UDP-Gal in the Golgi apparatus is limiting, either as a result of a low rate of UDP-Gal biosynthesis or as a result of inefficient cytosol-to-Golgi transport. Future research will be focused on optimization of galactose transfer.

In the studies of Choi et al. (11) and Hamilton et al. (16) mentioned above, fusion libraries were made by fusing the N-terminal parts of proteins that are targeted to the secretory pathway of yeast with different glycosidases and glycosyltransferases from different species. These libraries were screened for the best-performing combination of enzyme and targeting signal. The detailed results of the screening analysis have not been published yet, so we cannot evaluate whether the Kre2p leader sequence was used or whether it performed well in these studies. Apparently, more than one combination performs well. For example, the researchers at Dartmouth College used two different α-1,2-mannosidase–leader combinations in a strain to create $GlcNAcMan_5GlcNAc_2$. Choi et al. (11) used an Mns1p-*C. elegans* α-1,2-mannosidase fusion, whereas Hamil-

ton et al. (16) fused a Sec12p leader sequence to mouse mannosidase IA to reach the same goal. Furthermore, in the study of Choi et al. (11) an Mnn10p-*C. elegans* mannosidase fusion also resulted in the presence of $Man_5GlcNAc_2$ on the glycoprotein. From these results and from our results we concluded that a well-considered choice for a leader sequence might be as effective as high-throughput screening.

In conclusion, we developed a *Pichia* N-glycan engineering tool that allows, in a simple single-transformation step, efficient conversion of hyperglycosylation into homogeneous $Man_5GlcNAc_2$ glycosylation that results in up to 90% homogeneity. Furthermore, by consecutive introduction of two glycosyltransferases, we obtained a hybrid-type N-glycan structure. Very importantly, none of the glycosylation manipulation steps mentioned above severely affected the yeast's viability. These tools allow production of glycoproteins with glycoforms of choice in the engineered *P. pastoris* strains and glycan engineering of existing strains with good expression levels for the purpose of synthesizing glycoproteins with homogeneous N-linked glycosylation. We believe that these tools will greatly advance the use of *P. pastoris* in biomanufacturing of recombinant glycoproteins.

ACKNOWLEDGMENTS

W.V., S.G., and K.D.V. hold IWT fellowships, and N.C. is a postdoctoral researcher of FWO-Flanders. This research was funded by RCT and by GOA grant 12051099 from Ghent University.

We thank Wouter Laroy for critical reading of the manuscript. The technical support of Annelies Van Hecke is very much appreciated.

REFERENCES

1. **Ballou, C. E.** 1990. Isolation, characterization, and properties of *Saccharomyces cerevisiae mnn* mutants with nonconditional protein glycosylation defects. Methods Enzymol. **185:**440–470.
2. **Bretthauer, R. K.** 2003. Genetic engineering of *Pichia pastoris* to humanize N-glycosylation of proteins. Trends Biotechnol. **21:**459–462.
3. **Bretthauer, R. K., and F. J. Castellino.** 1999. Glycosylation of *Pichia pastoris*-derived proteins. Biotechnol. Appl. Biochem. **30:**193–200.
4. **Callewaert, N., S. Geysens, F. Molemans, and R. Contreras.** 2001. Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment. Glycobiology **11:**275–281.
5. **Callewaert, N., W. Laroy, H. Cadirgi, S. Geysens, X. Saelens, W. Min Jou, and R. Contreras.** 2001. Use of HDEL-tagged *Trichoderma reesei* mannosyl oligosaccharide 1,2-alpha-D-mannosidase for N-glycan engineering in *Pichia pastoris*. FEBS Lett. **503:**173–178.
6. **Cereghino, G. L., J. L. Cereghino, A. J. Sunga, M. A. Johnson, M. Lim, M. A. G. Gleeson, and J. M. Cregg.** 2001. New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*. Gene **263:**159–169.
7. **Chen, X., P. R. Andreana, and P. G. Wang.** 1999. Carbohydrates in transplantation. Curr. Opin. Chem. Biol. **3:**650–658.
8. **Chiba, Y., M. Suzuki, S. Yoshida, A. Yoshida, H. Ikenaga, M. Takeuchi, Y. Jigami, and E. Ichishima.** 1998. Production of human compatible high mannose-type (Man5GlcNAc2) sugar chains in *Saccharomyces cerevisiae*. J. Biol. Chem. **273:**26298–26304.
9. **Chiba, Y., Y. Yamagata, S. Iijima, T. Nakajima, and E. Ichishima.** 1993. The carbohydrate moiety of the acid carboxypeptidase from *Aspergillus saitoi*. Curr. Microbiol. **27:**281–288.
10. **Chitlaru, T., C. Kronman, M. Zeevi, M. Kam, A. Harel, A. Ordentlich, B. Velan, and A. Shafferman.** 1998. Modulation of circulatory residence of recombinant acetylcholinesterase through biochemical or genetic manipulation of sialylation levels. Biochem. J. **336:**647–658.
11. **Choi, B. K., P. Bobrowicz, R. C. Davidson, S. R. Hamilton, D. H. Kung, H. Li, R. G. Miele, J. H. Nett, S. Wildt, and T. U. Gerngross.** 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc. Natl. Acad. Sci. USA. **100:**5022–5027.
12. **Cregg, J. M., J. L. Cereghino, J. Shi, and D. R. Higgins.** 2000. Recombinant protein expression in *Pichia pastoris*. Mol. Biotechnol. **16:**23–52.
13. **Cregg, J. M., and K. A. Russell.** 1998. Transformation, p. 27–39. *In* D. R. Higgins and J. M. Cregg (ed.), *Pichia* protocols, vol. 103. Humana Press, Totowa, N.J.
14. **Dean, N.** 1999. Asparagine-linked glycosylation in the yeast Golgi. Biochim. Biophys. Acta **1426:**309–322.
15. **Fukuda, M.** 2000. Cell surface carbohydrates: cell type-specific expression, p. 12–61. *In* M. Fukuda and O. Hindgaul (ed.), Molecular and cellular glycobiology, vol. 30. Oxford University Press, New York, N.Y.
16. **Hamilton, S. R., P. Bobrowicz, B. Bobrowicz, R. C. Davidson, H. Li, T. Mitchell, J. H. Nett, S. Rausch, T. A. Stadheim, H. Wischnewski, S. Wildt, and T. U. Gerngross.** 2003. Production of complex human glycoproteins in yeast. Science **301:**1244–1246.
17. **Jackson, P.** 1990. The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled charge-coupled device. Biochem. J. **270:**705–713.
18. **Kornfeld, R., and S. Kornfeld.** 1985. Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. **54:**631–664.
19. **Laroy, W., and R. Contreras.** 2000. Cloning of *Trypanosoma cruzi* transsialidase and expression in *Pichia pastoris*. Protein Expr. Purif. **20:**389–393.
20. **Lee, S. J., S. Evers, D. Roeder, A. F. Parlow, J. Risteli, L. Risteli, Y. C. Lee, T. Feizi, H. Langen, and M. C. Nussenzweig.** 2002. Mannose receptor-mediated regulation of serum glycoprotein homeostasis. Science **295:**1898–1901.
21. **Lussier, M., A. M. Sdicu, T. Ketela, and H. Bussey.** 1995. Localization and targeting of the *Saccharomyces cerevisiae* Kre2p/Mnt1p alpha-1,2-mannosyltransferase to a medial-Golgi compartment. J. Cell Biol. **131:**913–927.
22. **Maras, M., N. Callewaert, K. Piens, M. Claeyssens, W. Martinet, S. Dewaele, H. Contreras, I. Dewerte, M. Penttila, and R. Contreras.** 2000. Molecular cloning and enzymatic characterization of a *Trichoderma reesei* 1,2-alpha-D-mannosidase. J. Biotechnol. **77:**255–263.
23. **Maras, M., X. Saelens, W. Laroy, K. Piens, M. Claeyssens, W. Fiers, and R. Contreras.** 1997. In vitro conversion of the carbohydrate moiety of fungal glycoproteins to mammalian-type oligosaccharides—evidence for N-acetylglucosaminyltransferase-I-accepting glycans from *Trichoderma reesei*. Eur. J. Biochem. **249:**701–707.
24. **Martinet, W., M. Maras, X. Saelens, W. M. Jou, and R. Contreras.** 1998. Modification of the protein glycosylation pathway in the methylotrophic yeast *Pichia pastoris*. Bio/Technol. Lett. **20:**1171–1177.
25. **Murakami, K., and N. Sugio.** 7 January 1997. Production of glycoprotein. Japan patent JP 1997000261-A1.
26. **Roy, S. K., T. Yoko-o, H. Ikenaga, and Y. Jigami.** 1998. Functional evidence for UDP-galactose transporter in *Saccharomyces cerevisiae* through the *in vivo* galactosylation and *in vitro* transport assay. J. Biol. Chem. **273:**2583–2590.
27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
28. **Schwientek, T., H. Narimatsu, and J. F. Ernst.** 1996. Golgi localization and *in vivo* activity of a mammalian glycosyltransferase (human beta-1,4-galactosyltransferase) in yeast. J. Biol. Chem. **271:**3398–3405.
29. **Van Petegem, F., H. Contreras, R. Contreras, and J. Van Beeumen.** 2001. *Trichoderma reesei* alpha-1,2-mannosidase: structural basis for the cleavage of four consecutive mannose residues. J. Mol. Biol. **312:**157–165.
30. **Weigel, P. H.** 1994. Galactosyl and N-acetylgalactosaminyl homeostasis: a function for mammalian asialoglycoprotein receptors. Bioessays **16:**519–524.
31. **Winkelhake, J. L., and G. L. Nicolson.** 1976. Aglycosyl antibody. Effects of exoglycosidase treatments on autochthonous antibody survival time in the circulation. J. Biol. Chem. **251:**1074–1080.
32. **Yip, C. L., S. K. Welch, F. Klebl, T. Gilbert, P. Seidel, F. J. Grant, P. J. O'Hara, and V. L. MacKay.** 1994. Cloning and analysis of the *Saccharomyces cerevisiae* MNN9 and MNN1 genes required for complex glycosylation of secreted proteins. Proc. Natl. Acad. Sci. USA **91:**2723–2727.
33. **Yoshida, S., M. Suzuki, S. Yamano, M. Takeuchi, H. Ikenaga, N. Kioka, H. Sakai, and T. Komano.** 1999. Expression and characterization of rat UDP-N-acetylglucosamine:alpha-3-D-mannoside beta-1,2-N-acetylglucosaminyltransferase I in *Saccharomyces cerevisiae*. Glycobiology **9:**53–58.