Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose

Piotr Bobrowicz¹, Robert C. Davidson¹, Huijuan Li, Thomas I. Potgieter, Juergen H. Nett, Stephen R. Hamilton, Terrance A. Stadheim, Robert G. Miele, Beata Bobrowicz, Teresa Mitchell, Sebastian Rausch, Eduard Renfer, and Stefan Wildt²

GlycoFi, Inc., 21 Lafayette Street, Suite 200, Lebanon, NH 03766

Received on January 20, 2004; revised on May 25, 2004; accepted on May 26, 2004

A significant percentage of eukaryotic proteins contain post-translational modifications, including glycosylation, which are required for biological functions. However, the understanding of the structure–function relationships of N-glycans has lagged significantly due to the microheterogeneity of glycosylation in mammalian produced proteins. Recently we reported on the cellular engineering of yeast to replicate human N-glycosylation for the production of glycoproteins. Here we report the engineering of an artificial glycosylation pathway in *Pichia pastoris* blocked in dolichol oligosaccharide assembly. The *PpALG3* gene encoding Dol-P-Man₉GlcNAc₂-PP-Dol mannosyltransferase was deleted in a strain that was previously engineered to produce hybrid GlcNAcMan₅GlcNAc₂ human N-glycans. Employing this approach, combined with the use of combinatorial genetic libraries, we engineered *P. pastoris* strains that synthesize complex GlcNAc₂Man₅GlcNAc₂ N-glycans with striking homogeneity. Furthermore, through expression of a Golgi-localized fusion protein comprising UDP-glucose 4-epimerase and β-1,4-galactosyltransferase activities we demonstrate that this structure is a substrate for highly efficient in vivo galactose addition. Taken together, these data demonstrate that the artificial in vivo glycoengineering of yeast represents a major advance in the production of glycoproteins and will emerge as a practical tool to systematically elucidate the structure–function relationship of N-glycans.

**Key words:** *Pichia*/glycoprotein/N-linked glycosylation/ALG3/oligosaccharide

**Introduction**

Therapeutic proteins are the fastest-growing source of new therapies in the biopharmaceutical industry (Walsh, 2000). The resulting demand for more efficient protein expression systems is leading the industry to seek alternatives to existing expression hosts. In contrast to mammalian cell lines, which are the preferred vehicle for the expression of most glycoproteins, yeast-based expression systems offer a variety of advantages including shorter fermentation times, cheaper operating costs, chemically defined media, no viral contamination, higher protein titers, and much shorter development times from gene to protein. However, the inability to replicate human glycosylation has prevented yeast and other fungal hosts from becoming a suitable expression host for therapeutic glycoprotein production.

The processing of asparagine-linked glycans transferred onto a protein differs significantly between higher and lower eukaryotes. The initial stage, highly conserved throughout eukarya, comprises the assembly of a lipid-linked oligosaccharide, Glc₃Man₉GlcNAc₂, followed by its transfer to the nascent protein and the removal of three glucose sugars and one mannose to yield Man₅GlcNAc₂ (Figure 1A; Burda and Aebi, 1999; Kornfeld and Kornfeld, 1985). In the second stage, localized in the Golgi, the protein-linked N-glycans are further modified along pathways that diverge significantly across species. Yeast and other fungi typically produce high-mannose type N-glycans by adding up to 100 mannose sugars, whereas the formation of mammalian glycans generally involves the trimming of mannose followed by the addition of N-acetylglucosamine, galactose, fucose, and sialic acid (Figure 1; Dean, 1999; Gemmill and Trimble, 1999; Kukuruzinska and Lennon, 1998; Moremen et al., 1994).

The replication of the human N-linked glycosylation machinery in a fungal host requires the cloning, functional expression, and proper targeting of several mannosidase and glycosyltransferase enzymes with each one operating at high efficiency. Previously we have described the genetic engineering of N-glycosylation in the yeast *Pichia pastoris*, developing a strain that produces the hybrid N-glycan GlcNAcMan₅GlcNAc₂ on a secreted reporter protein (Choi et al., 2003, fig. 3B therein). This was accomplished by creating combinatorial genetic libraries of protein fusions of fungal type II membrane leaders and catalytic domains of glycosylation enzymes. After transforming these genetic fusion libraries into a *P. pastoris* strain containing a UDP-GlcNAc transporter and lacking fungal specific α₁,6-mannosyltransferase activity (Och1p), we were able to identify specific combinations of α₁,2-mannosidases (MNS I) and N-acetylglucosaminyltransferases I (GnT I) that...
efficiently replicated early human N-glycan processing (GlcNAcMan$_3$ in Figure 1B and Figure 2).

In humans, the subsequent steps of glycosylation involve the removal of the terminal $\alpha$-1,3- and $\alpha$-1,6-linked mannosides from GlcNAcMan$_5$GlcNAc$_2$ by $\alpha$-mannosidase II (Tulsiani et al., 1982). This exposes the $\alpha$-1,6 mannose residue of the trimannose core, which is then extended by GnT II (Figure 1A; Bendiak and Schachter, 1987). This biosynthetic pathway, leading to complex N-glycans, was recently replicated in P. pastoris (Hamilton et al., 2003).
In the present work, rather than expressing and targeting mannosidase II, we report an alternative strategy to produce a yeast strain capable of assembling complex N-glycans. By interfering with the assembly of the oligosaccharide at a dolichol-linked stage, we engineered an artificial glycosylation pathway that generates complex human N-glycans on a secreted reporter protein.

In *Saccharomyces cerevisiae*, ALG3 has been shown to encode the Dol-P-Man:Man₅GlcNAc₂-PP-Dol α-1,3-mannosyltransferase, which is responsible for the first Dol-P-Man-dependent mannosyl transfer step converting Man₅GlcNAc₂-PP-Dol to Man₆GlcNAc₂-PP-Dol in the lumen of the endoplasmic reticulum (Sharma et al., 2001). The deletion of ALG3 in *S. cerevisiae* and *P. pastoris* results in the transfer of Glc₃Man₅GlcNAc₂ to nascent poly-peptides, albeit at a somewhat reduced efficiency (Aebi et al., 1996; Davidson et al., 2004).

It was demonstrated previously in some mutant mammalian cell lines that the synthesis of dolichol-linked oligosaccharides is blocked at the Man₅GlcNAc₂-PP-Dol stage (resembling the lack of Alg3p activity). Despite the inability of those cells to make the mature Glc₃Man₅GlcNAc₂-PP-Dol, they retained the ability to produce complex-type N-glycans (Kornfeld et al., 1979; Lehrman and Zeng, 1989; Stoll et al., 1982). We reasoned that a yeast alg3 mutant strain expressing a properly targeted and active α-1,2-mannosidase should be able to generate a Man₃, GlcNAc₂ structure, an artificial glycosylation intermediate that could be further modified to form a complex N-glycan. Previous work by Schachter and co-workers, which demonstrated that Man₃GlcNAc₂ can serve as an acceptor for N-acetylglucosamine-transfer by GnT I in vitro, led us to attempt the replication of this reaction in vivo in a yeast (Narasimham et al., 1977; Vella et al., 1984).

Here we demonstrate that blocking the assembly of the yeast core oligosaccharide, combined with the subsequent use of combinatorial genetic libraries to target mammalian glycosylation enzymes to the secretory pathway, enabled us to engineer a *P. pastoris* strain that synthesizes complex human N-glycans of high uniformity. Moreover, we reveal that this alg3 mutant-generated GlcNAc₂Man₃GlcNAc₂ structure is the substrate for quantitative galactose addition upon proper engineering and targeting of a pool of UDP-galactose and a transferase enzyme.

**Results**

**Deletion of PpALG3 in a strain producing hybrid type N-glycans**

In our previous work, we described the deletion of OCH1 in *P. pastoris*, resulting in the loss of outer chain initiating α-1,6-mannosyltransferase activity (Figure 1A and Choi et al., 2003). Replication of hybrid-type human glycosylation required the proper targeting of several active glycosylation enzymes, which was achieved via the use of extensive combinatorial libraries of fungal endoplasmic reticulum and Golgi targeting domains fused to catalytic domains from mannosidases and glycosyl transferases (Choi et al., 2003). The functional expression of properly targeted MNS I, UDP-GlcNAc transporter, and GnT I was monitored by analyzing the glycosylation pattern of the secreted reporter protein K3 by matrix-assisted laser desorption/ionization time-of-flight (MALTI-TOF) mass spectrometry (MS). Screening several hundred strains resulted in the isolation of a few strains secreting proteins with N-glycans of the hybrid-type, GlcNAc₃Man₃GlcNAc₂ (Choi et al., 2003).

In mammals, GlcNAc₃Man₃GlcNAc₂ is modified by the removal of the two terminal mannoses from the 1,6-arm by mannosidase II, we report an alternative strategy to produce a yeast strain capable of assembling complex N-glycans. By interfering with the assembly of the oligosaccharide at a dolichol-linked stage, we engineered an artificial glycosylation pathway that generates complex human N-glycans on a secreted reporter protein. In *Saccharomyces cerevisiae*, ALG3 has been shown to encode the Dol-P-Man:Man₅GlcNAc₂-PP-Dol α-1,3-mannosyltransferase, which is responsible for the first Dol-P-Man-dependent mannosyl transfer step converting Man₅GlcNAc₂-PP-Dol to Man₆GlcNAc₂-PP-Dol in the lumen of the endoplasmic reticulum (Sharma et al., 2001). The deletion of ALG3 in *S. cerevisiae* and *P. pastoris* results in the transfer of Glc₃Man₅GlcNAc₂ to nascent poly-peptides, albeit at a somewhat reduced efficiency (Aebi et al., 1996; Davidson et al., 2004).

It was demonstrated previously in some mutant mammalian cell lines that the synthesis of dolichol-linked oligosaccharides is blocked at the Man₅GlcNAc₂-PP-Dol stage (resembling the lack of Alg3p activity). Despite the inability of those cells to make the mature Glc₃Man₅GlcNAc₂-PP-Dol, they retained the ability to produce complex-type N-glycans (Kornfeld et al., 1979; Lehrman and Zeng, 1989; Stoll et al., 1982). We reasoned that a yeast alg3 mutant strain expressing a properly targeted and active α-1,2-mannosidase should be able to generate a Man₃, GlcNAc₂ structure, an artificial glycosylation intermediate that could be further modified to form a complex N-glycan. Previous work by Schachter and co-workers, which demonstrated that Man₃GlcNAc₂ can serve as an acceptor for N-acetylglucosamine-transfer by GnT I in vitro, led us to attempt the replication of this reaction in vivo in a yeast (Narasimham et al., 1977; Vella et al., 1984).

Here we demonstrate that blocking the assembly of the yeast core oligosaccharide, combined with the subsequent use of combinatorial genetic libraries to target mammalian glycosylation enzymes to the secretory pathway, enabled us to engineer a *P. pastoris* strain that synthesizes complex human N-glycans of high uniformity. Moreover, we reveal that this alg3 mutant-generated GlcNAc₂Man₃GlcNAc₂ structure is the substrate for quantitative galactose addition upon proper engineering and targeting of a pool of UDP-galactose and a transferase enzyme.

**Results**

**Deletion of PpALG3 in a strain producing hybrid type N-glycans**

In our previous work, we described the deletion of OCH1 in *P. pastoris*, resulting in the loss of outer chain initiating α-1,6-mannosyltransferase activity (Figure 1A and Choi et al., 2003). Replication of hybrid-type human glycosylation required the proper targeting of several active glycosylation enzymes, which was achieved via the use of extensive combinatorial libraries of fungal endoplasmic reticulum and Golgi targeting domains fused to catalytic domains from mannosidases and glycosyl transferases (Choi et al., 2003). The functional expression of properly targeted MNS I, UDP-GlcNAc transporter, and GnT I was monitored by analyzing the glycosylation pattern of the secreted reporter protein K3 by matrix-assisted laser desorption/ionization time-of-flight (MALTI-TOF) mass spectrometry (MS). Screening several hundred strains resulted in the isolation of a few strains secreting proteins with N-glycans of the hybrid-type, GlcNAc₃Man₃GlcNAc₂ (Choi et al., 2003).

In mammals, GlcNAc₃Man₃GlcNAc₂ is modified by the removal of the two terminal mannoses from the 1,6-arm by mannosidase II, we report an alternative strategy to produce a yeast strain capable of assembling complex N-glycans. By interfering with the assembly of the oligosaccharide at a dolichol-linked stage, we engineered an artificial glycosylation pathway that generates complex human N-glycans on a secreted reporter protein. In *Saccharomyces cerevisiae*, ALG3 has been shown to encode the Dol-P-Man:Man₅GlcNAc₂-PP-Dol α-1,3-mannosyltransferase, which is responsible for the first Dol-P-Man-dependent mannosyl transfer step converting Man₅GlcNAc₂-PP-Dol to Man₆GlcNAc₂-PP-Dol in the lumen of the endoplasmic reticulum (Sharma et al., 2001). The deletion of ALG3 in *S. cerevisiae* and *P. pastoris* results in the transfer of Glc₃Man₅GlcNAc₂ to nascent poly-peptides, albeit at a somewhat reduced efficiency (Aebi et al., 1996; Davidson et al., 2004).

It was demonstrated previously in some mutant mammalian cell lines that the synthesis of dolichol-linked oligosaccharides is blocked at the Man₅GlcNAc₂-PP-Dol stage (resembling the lack of Alg3p activity). Despite the inability of those cells to make the mature Glc₃Man₅GlcNAc₂-PP-Dol, they retained the ability to produce complex-type N-glycans (Kornfeld et al., 1979; Lehrman and Zeng, 1989; Stoll et al., 1982). We reasoned that a yeast alg3 mutant strain expressing a properly targeted and active α-1,2-mannosidase should be able to generate a Man₃, GlcNAc₂ structure, an artificial glycosylation intermediate that could be further modified to form a complex N-glycan. Previous work by Schachter and co-workers, which demonstrated that Man₃GlcNAc₂ can serve as an acceptor for N-acetylglucosamine-transfer by GnT I in vitro, led us to attempt the replication of this reaction in vivo in a yeast (Narasimham et al., 1977; Vella et al., 1984).

Here we demonstrate that blocking the assembly of the yeast core oligosaccharide, combined with the subsequent use of combinatorial genetic libraries to target mammalian glycosylation enzymes to the secretory pathway, enabled us to engineer a *P. pastoris* strain that synthesizes complex human N-glycans of high uniformity. Moreover, we reveal that this alg3 mutant-generated GlcNAc₂Man₃GlcNAc₂ structure is the substrate for quantitative galactose addition upon proper engineering and targeting of a pool of UDP-galactose and a transferase enzyme.
glycans would result in a strain producing the artificially generated N-glycan intermediate GlcNAcMan₃GlcNAc₂ and obviate the need for mannosidase II (Figure 2). Previously we reported the cloning and deletion of PpALG3 in an och1 mutant strain of P. pastoris (Davidson et al., 2004). From a variety of K3 secreting strains generated in our laboratory, we selected strain PBP3, an och1 deletion that contains catalytic domains of mouse MNS I and human GnT I fused to yeast localization signals from Sec12p and Mnn9p, respectively, and a UDP-GlcNAc transporter from Kluyveromyces lactis (Choi et al., 2003; Hamilton et al., 2003).

The PpALG3 gene was deleted in the PBP3 strain background using a polymerase chain reaction (PCR)–generated alg3::G418R deletion allele to create strain RDP27 (Figure 2). MALDI-TOF MS analysis of N-glycans released from K3 demonstrated that the major glycoform was of a mass consistent with GlcNAcMan₃GlcNAc₂ (Figure 3) with a few additional N-glycans of higher molecular mass. MALDI-TOF analysis following in vitro β-N-acetylhexosaminidase and α-1,2-mannosidase digest indicate that these glycans contain one terminal β-linked GlcNAc and up to four α-1,2-mannose sugars (data not shown). The addition of α-1,2-linked mannose in the Golgi has previously been shown to occur in both P. pastoris and S. cerevisiae (Dean, 1999; Hamilton et al., 2003; Verostek and Trimble, 1995). Our data suggest that the alg3 mutant strain produces an alg3 Man₃ structure (Figure 1B) that serves as a substrate for MNS I, resulting in a synthetic Man₃GlcNAc₂ intermediate. This glycan is then further modified by the GnT I (ScMnn9-hGnT I fusion) and endogenous Golgi-residing α-1,2-mannosyltransferase(s), which adds up to four mannoses (Figure 3C). These results indicate that human GnT I, if efficiently targeted to the yeast Golgi, can quantitatively transfer GlcNAc to Man₃GlcNAc₂, resulting in the same GlcNAcMan₃GlcNAc₂ intermediate that would have been generated in the human pathway using mannosidase II (Figure 1A human versus Figure 1B).

**Figure 3. MALDI-TOF MS analysis of N-linked glycans released from recombinant K3.** K3 was produced in wild-type (BK64), PBP3, RDP27, PBP6-5, and RDP93 strains of P. pastoris and purified from culture supernatants by Ni-affinity chromatography. The N-linked glycans were released from the purified protein by PNGase F treatment and analyzed by MALDI-TOF MS. In the positive ion mode shown here, the N-glycans appear as sodium or potassium adducts. (A) N-glycans from BK64. (B) N-glycans from PBP3. (C) N-glycans from RDP27. (D) N-glycans from PBP6-5. (E) N-glycans from RDP93. M, mannose; GlcNAc, N-acetylgalactosamine; Gal, galactose; Mns I, α-1,2-mannosidase I; GnT I, β-1,2-N-acetylgalactosaminyltransferase I; GnT II, β-1,2-N-acetylgalactosaminyltransferase II; GalT, β-1,4-galactosyltransferase. The two core GlcNAc residues, though present in all glycans, have been omitted to simplify labeling.

**Complex N-glycan production: functional expression of GnT II**

In mammalian cells, the GlcNAcMan₃GlcNAc₂ structure generated by mannosidase II serves as a substrate for GnT II. The transfer of GlcNAc to the 1,6 arm by GnT II is an essential step in the biosynthesis of complex-type N-glycans. Although RDP27 generates mostly GlcNAcMan₃GlcNAc₂ structures, the aforementioned experiments also suggest that endogenous α-1,2-mannosyltransferases are able to convert some GlcNAcMan₃GlcNAc₂ to GlcNAcMan₄GlcNAc₂ and GlcNAcMan₃GlcNAc₄. Postulated that localization of an active GnT II catalytic domain early in the Golgi could preempt this addition of α-1,2-mannoses, leading to GlcNAcMan₃GlcNAc₂, which would not be a substrate for subsequent α-1,2-mannose transfer.

Using a strategy previously reported (Choi et al., 2003), several dozen chimeric constructs consisting of cDNA fragments encoding a rat GnT II catalytic domain were generated by fusion to a library of yeast Golgi targeting sequences and transformed into P. pastoris strain RDP27. MALDI-TOF MS analysis of N-glycans released from secreted K3 identified several strains able to produce secreted K3.
N-glycans of a mass consistent with GlcNAc_2Man_3GlcNAc_2 (Figure 3). One of these strains (PBP6-5), containing a leader sequence from the \textit{S. cerevisiae} MNN9 gene, produced glycans with a nearly uniform GlcNAc_2Man_3GlcNAc_2 structure (Figure 3D). The identity of the GlcNAc_2Man_3GlcNAc_2 structure was confirmed by nuclear magnetic resonance analysis and digestion with β-N-acetylhexosaminidase (data not shown). Furthermore, the purified K3 protein from strain PBP6-5 was incubated with β-1,4-galactosyltransferase (GalT) in the presence of UDP-galactose. MALDI-TOF MS analysis demonstrates that the GlcNAc_2Man_3GlcNAc_2 N-glycans on this protein had received two galactose residues and thus, as expected, can serve as a substrate for subsequent mammalian-type modification (data not shown).

In addition, we previously reported that the deletion of \textit{Pp}ALG3 in an ochl mutant background leads to the occurrence of several larger glycans ranging from Hex_5GlcNAc_2 to Hex_7GlcNAc_2, which are recailetant to an array of mannosidase digests (Davidson \textit{et al.}, 2004). The presence of structures larger than Man_5GlcNAc_2 suggests the action of additional yet unidentified transferases capable of transferring sugars to the alg3 Man_5 glycan. Interestingly, the high uniformity of the GlcNAc_2Man_3GlcNAc_2 N-glycans observed here suggests that these unknown transferases are now depleted of their substrate, and thus are no longer capable of adding hexoses.

In vivo transfer of galactose in an alg3 mutant strain producing complex N-glycans

We sought to confirm the identity of the GlcNAc_2Man_3GlcNAc_2 structure and show that quantitative \textit{in vivo} galactose transfer is possible on a complex structure produced in an \textit{alg3} mutant strain. Previous experiments demonstrated that galactose transfer in \textit{P. pastoris} required not only the presence of a galactosyl transferase but also the generation of a pool of precursor (UDP-galactose) and the transport of the precursor into the Golgi (Davidson \textit{et al.}, in press). Recently, an active secreted fusion protein of \textit{Escherichia coli} UDP-galactose 4-epimerase and bovine α-1,3-galactosyltransferase has been reported (Chen \textit{et al.}, 2000). We reasoned that if there was sufficient UDP-glucose in the yeast Golgi, an analogous fusion protein might function to supply a localized pool of UDP-galactose, eliminating the need for a membrane transporter.

Thus a three-part fusion gene was constructed using engineered restriction sites, which encodes the first 36 amino acids of ScMan2, the full-length protein predicted to be a UDP-galactose 4-epimerase from \textit{Schizosaccharomyces pombe}, and human GaIT I lacking the first 43 amino acids. This fusion gene was cloned downstream of the \textit{PpGAPDH} promoter and introduced into strain PBP6-5. Transformants were analyzed by MALDI-TOF MS, and several resulting strains were obtained in which the majority of N-glycans were of a mass consistent with Gal_3GlcNAc_2Man_3GlcNAc_2 or addition of two galactose residues to the N-glycans produced in PBP6-5. The MALDI-TOF MS spectrum of N-glycans resulting from one of these strains, designated RDP93, is shown (Figure 3E). The mass of the N-glycans resulting from strain RDP93 exactly corresponds with that obtained from \textit{in vitro} addition of galactose to N-glycans purified from strain PBP6-5 (data not shown). Furthermore, \textit{in vitro} β-galactosidase digest resulted in N-glycans of a mass consistent with GlcNAc_2Man_3GlcNAc_2, confirming the identity of the two additional hexose residues as galactose (data not shown).

Performance comparison of genetically reengineered strains

To determine the impact of these genetic modifications on heterologous protein expression, several shake flask and fermentation experiments were performed. A standardized shake flask protocol was used to determine K3 production levels in wild-type \textit{P. pastoris} (BK64), ochl mutant (BK3-1), PBP3, RDP27, and PBP6-5 (Figure 2). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of supernatants from these shake flasks revealed similar protein production levels for all five strains (Figure 4A). Fermentation experiments were conducted to assess the scale-up characteristics of PBP6-5. The specific productivity of PBP6-5 observed in shake flask experiments (2.2 mg/OD_{600} day) was conserved in the bioreactor (2.2 mg/OD_{600} day). Furthermore, the specific productivities observed in bioreactor experiments between PBP6-5 (2.2 mg/OD_{600} day) and BK64 (2.6 mg/OD_{600} day) were similar.
SDS–PAGE analysis also revealed the appearance of an unglycosylated fraction of K3 in alg3 strains, consistent with previous observations in alg3 mutants of *S. cerevisiae* (Figure 4A; Cueva et al., 1996). Although the unglycosylated fraction is a relatively small portion of the total K3 produced, reduced occupancy in these strains will have to be monitored on several proteins to determine the severity of this effect. MALDI-TOF analysis of the glycans released from K3 produced by PBP6-5 in the bioreactor demonstrates that the glycan structures are essentially identical to that obtained from shake flasks (Figure 4B versus Figure 3D). We thus conclude that the heterologous glycosylation enzymes remain functional and stable when the strain is cultivated under typical industrial fermentation conditions.

**Discussion**

We have demonstrated that a yeast strain of commercial interest can be altered to produce heterologous proteins with complex human N-glycans by blocking endogenous core N-glycan biosynthesis and expressing heterologous mannosidase and glycosyl transferases. Specifically, the elimination of the Alg3α-1,3-mannosyltransferase prevents the α-1,3 and α-1,6 mannose additions to the 1,6 arm and, in combination with properly targeted MNS1, GnT I, and GnT II enzymes, results in remarkably homogeneous N-glycan formation of the complex type. Through *in vitro* analysis and addition of β-1,4-galactose, we confirmed the GlcNAc2- Man3GlcNAc2 structure and demonstrate its potential for further mammalian-type modifications. Furthermore, through expression and proper localization of a novel fusion protein, comprising UDP-galactose 4-epimerase and β-1,4-galactosyltransferase activities, we demonstrate the ability to extend this alg3 mutant-derived complex N-glycan to Gal2GlcNAc2Man3GlcNAc2 in *vivo*.

In previous experiments, we observed that expression of human GaIT I alone was insufficient to extend complex GlcNAc2Man3GlcNAc2 N-glycans to Gal2GlcNAc2Man3GlcNAc2 because of the apparent lack of a pool of UDP-galactose precursor in the *P. pastoris* Golgi (Davidson *et al.*, in press). However, engineered expression of a UDP-galactose 4-epimerase from *S. pombe*, which converts UDP-glucose to UDP-galactose, successfully generated a pool of UDP-galactose precursor. Combined with expression of a UDP-galactose transporter to move the UDP-galactose into the Golgi and the GaIT, this enabled us to achieve transfer of galactose onto complex N-glycans in *P. pastoris*. Instead of independent expression of the three necessary genes, we created a fusion gene of the UDP-galactose 4-epimerase and human GaIT I and fused this to a sequence encoding a yeast Golgi localization domain to direct the entire tripartite polypeptide to the Golgi. Successful expression of this fusion protein colocalizes the production of UDP-galactose precursor and transfer of galactose onto the N-glycan and eliminates the need for a Golgi transporter. Moreover it reduces the necessary repertoire of genes for metabolic engineering and secretory pathway engineering to a single open reading frame.

As with previous steps, addition of galactose can be engineered to be highly efficient, yielding a strain that produces almost homogenous N-glycans and continuing a theme that has been a hallmark of the glycoengineered yeast system. The remarkable efficiency of this method in generating and extending complex N-glycans is an important finding for several reasons. Most important, this set of experiments demonstrates that an artificial N-glycosylation pathway can be generated that not only allows for the production of glycans via new pathways but to striking homogeneity. Second, our results demonstrate that the alg3 deletion offers an efficient alternative to functionally expressing, properly targeting, and maximizing mannosidase II activity. Moreover, the powerful combination of mutation and expression of heterologous enzymes afforded by fungal systems allows one to envision the attempt to assemble virtually any sugar structure *in vivo* for which a set of enzymes can be collected.

A significant unknown at the outset of this project was in the substrate specificity of the GnT I enzyme expressed in yeast. The endogenous substrate for GnT I in all known higher eukaryotes is Man9GlcNAc2 (Figure 1A), to which GnT I adds a single GlcNAc in β-1,2 linkage on the 1,3 arm. However, in the scenario described here, the two mannoses normally transferred onto the upper 1,6 arm are eliminated by deletion of *ALG3*, requiring GnT I to act on a Man9, GlcNAc2 structure following the action of α-1,2-mannosidase (Figure 1B). Analysis of several mammalian cell lines blocked in core assembly reveals that GnT I must act on a Man9GlcNAc2 structure to allow the generation of at least some complex N-glycans in these mutant cell lines (Kornfeld *et al.*, 1979; Lehman and Zeng, 1989; Stoll *et al.*, 1982). Based on *in vitro* studies, some GnT I enzymes are better equipped to handle this assignment than others (Altmann *et al.*, 1993; Strasser *et al.*, 1999; Vella *et al.*, 1984; Zhang *et al.*, 2003). Although all GnT I enzymes favor the Man9GlcNAc2 structure, the human GnT I used in this study acts with reasonable efficiency on a Man9GlcNAc2 substrate *in vitro*, whereas some (such as insect or plant GnT I) enzymes are substantially less effective when compared to their natural substrate, Man9GlcNAc2. The results presented here demonstrate that human GnT I, when properly targeted to the yeast secretory pathway, is able to act with high efficiency on a Man9GlcNAc2 substrate.

Given the importance of N-glycosylation in such essential cellular functions as cell wall formation, cell division, cell-cycle progression, protein folding, and even glycosylation itself, the effect of an extensively reengineered N-glycosylation pathway on growth has been an open question. However, the results presented here reveal that drastic changes can be made to the core glycosylation machinery as well as subsequent processing steps of a *P. pastoris* strain while maintaining viability. Previously, mutations in core glycosylation generated in *S. cerevisiae* resulted in a reduction in occupancy referred to as underglycosylation (Aebi *et al.*, 1996; Basco *et al.*, 1993; Cipollo and Trimble, 2000; Huffaker and Robbins, 1983; Cueva *et al.*, 1996; Munoz *et al.*, 1994; Verostek *et al.*, 1993a,b; Zufferey *et al.*, 1995). The underglycosylation phenotype observed here on K3 is mild. However, further studies will be required to analyze the effects of this core glycosylation mutant on other more complex glycoproteins in *P. pastoris* as well as possible.
development of artificial compensation mechanisms to improve occupancy. Moreover, although the cells do display a minor reduction in growth rate, we demonstrate that the performance of these strains in an industrial bioreactor and the levels of secreted protein compare favorably to that of the parental wild-type yeast.

Finally, this type of glycoengineering of yeast can lead the way toward highly sophisticated structure–function studies of glycosylated proteins that until now were not possible. Furthermore, in contrast to glycoproteins secreted by mammalian cells (which typically contain a mixed population of glycoforms), genetically engineered yeast cell lines are able to secrete glycoproteins of exceptional homogeneity. Thus, to understand the relationship between particular glycan structures and their biological function, these engineered cell lines may be used to generate libraries of homogenous glycoforms of one and the same protein. These glycoform libraries can then be used to dissect the specific structure–function relationship of different glycoforms and their effect on pharmacokinetic behavior and biological activity for any given glycoprotein and will prove to be of great interest in the elucidation of the function of specific glycoforms.

Materials and methods

Strains, culture conditions, and reagents

E. coli strains TOP10 or DH5a were used for recombinant DNA work. Protein expression was carried out at room temperature in a 96-well plate format 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 × 10⁻⁴% biotin, and 1% glycerol as a growth medium. The induction medium was buffered methanol-complex medium (BMMY) consisting of 1% methanol instead of glycerol in BMGY. Minimal medium is 1.4% yeast nitrogen base, 2% dextrose, 1.5% agar, and 4 × 10⁻⁵% biotin and amino acids supplemented as appropriate. Restriction and modification enzymes were from New England BioLabs (Beverly, MA). Oligonucleotides were obtained from the Dartmouth College Core facility (Hanover, NH) or Integrated DNA Technologies (Coralville, IA). DEAE ToyoPearl resin was from Tosohaas. Metal-chelating HisBind resin was from Novagen (Madison, WI). Lysate-clearing plates (96-well) were from Promega (Madison, WI). Protein-binding 96-well plates were from Millipore (Bedford, MA). Salts and buffering agents were from Sigma (St. Louis, MO). MALDI matrices were from Aldrich (Milwaukee, WI).

Construction of an alg3 deletion strain of P. pastoris

The detailed descriptions of P. pastoris strains BK64 and PB3 were published previously (Choi et al., 2003; Hamilton et al., 2003). The alg3::G418R allele used for deletion of the PpALG3 (AY653304) gene in strain PB3 was generated by the PCR overlap method (Davidson et al., 2000). Primers RCD142 (5’-CCACATCTCAGTCTA- CAGACATCTCAGA-3’) and RCD144 (5’-ACGAGG- CAAGCTAACCAGATCTCAGAAGCT-3’) paired with RCD146 (5’-AGCGTCTCAGGCAGTTGAG-3’) and RCD145 (5’-CCATCCAGTGTCAAAGAGGCTGAGGGTTA- TCCAG-3’) suited to amplify the 5’ and 3’ flanking regions of the ALG3 gene and the G418 resistance marker (G418R), respectively. Then primers RCD142 and RCD147 were used in a second reaction with all three first-round templates to generate an overlap product that contained all three fragments as a single linear alg3::G418R allele. This PCR product was then directly employed for transformation with selection on medium containing 200 µg/ml G418. Primers MG1 (5’-TCTGCTAGATCTCCATAGTGGTGCTG-3’) and PTEF (5’-AGCTGCGCCAGTGTCACCTGCAAGGGTT- CATCGGATGTTGCTGCTTGCCTCGT-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, a restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers.

Construction of a GnT II–expressing alg3 mutant yeast strain

The rat GnT II gene (GenBank accession number U21662) was PCR amplified using Takara EX Taq polymerase (Panvera, Madison, WI) from rat liver cDNA library (Clontech, Palo Alto, CA) with RAT1 (5’-TTTCTCCACTCTG- CAGTCTCCTAATAC-3’) and RAT2 (5’-TGGAGGACTCA- TGAGGTTCCGCACTGAATC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers. The PCR product was then cloned into pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced. Using this vector as a template, the restriction enzyme digestion with the restriction enzymes SacI and PstI gave the 5’ and 3’ termini for DNA fragments. The primers RCD178 (5’-TTATGCTGCTGAGGTCTGTTTC-3’) and RCD179 (5’-ATCAGATGCTGAGGTCTGTTTC-3’) primers.

Construction and expression of a UDP-galactose 4-epimerase/galactosyl transferase fusion construct

A gene encoding UDP-galactose 4-epimerase activity was previously amplified from S. pombe genomic DNA and
cloned (NC_003423) to generate plasmid pRCD406 (unpublished data). The SpGALE gene was subsequently amplified using primers RCD326 (5'-CTTGGATCTTGGTTTGGGCACCTAAGCCGCTGCC-3') and RCD329 (5'-CCTGGATCTTTATATGGGTCAGG-3') cloned into the pCR2.1 vector (Invitrogen) and sequenced. The hGalTIA43 truncated clone was amplified using primers RCD328 (5'-CTTGGATCTTGGTTTGGGCACCTAAGCCGCTGCC-3') and RCD199 (5'-CTTGGATCTTTATATGGGTCAGG-CGATGTCACC-3') cloned into the pCR2.1 vector (Invitrogen) and sequenced. The SpGALE gene was then digested with AscI/BamHI, and the hGalTI clone digested with BamHI/PacI, and both were inserted into pRCD452 digested with AscI/PacI. The plasmid pRCD452 contains a NotI/AscI/PacI polylinker flanked by a GAPDH/CYC1 cassette with a gene encoding the first 46 amino acids of the S. cerevisiae MNN2 gene (ScMNN2(s)) as a Golgi-targeting leader sequence as a NotI/AscI fragment. The AslI/BamHI SpGALE and BamHI/PacI hGalTIA43 fragments were ligated into the AslI/PacI in frame with the ScMNN2(s) leader sequence to create pRCD461. This new plasmid, pRCD461 contains a ScMNN2(s)/SpGALE/hGalTI fusion where the SpGALE and hGalTI proteins are encoded in a single polypeptide separated by a four-amino-acid (GSGG) linker containing the BamHI site, and all driven by the PpGAPDH promoter. The ScMNN2(s)/SpGALE/hGalTI fusion gene was subcloned as a NotI/PacI fragment into vector pJN701 to create plasmid pSH336. Plasmid pSH336 contains a NotI/PacI multiple cloning site flanked by a PpGAPDH/CYC1 cassette, the HYG<sup>R</sup> gene for selection, and flanking regions of the PpHIS1 gene for targeting. The fusion gene containing plasmid, pSH336, was digested with SfiI to release the expression/targeting fragment, transformed into strain PB6-5, and the transformants were selected on YPD medium containing 200 μg/ml hygromycin (Calbiochem, San Diego, CA).

**Reporter protein purification and release of N-linked glycans**

The K3 domain, under the control of the P. pastoris alcohol oxidase 1 (AOXI) promoter, was used as a model protein and was purified using the 6xHistidine tag as reported previously (Choi et al., 2003). The glycans were released and separated from the glycoproteins by a modification of a previously reported method (Papac et al., 1998). After the proteins were reduced and carboxymethylated and the membranes blocked, the walls were washed three times with water. The protein was deglycosylated by the addition of 30 μl 10 mM NH₂HCO₃, pH 8.3, containing 1 mM N-glycanase (Glyko, Novato, CA). After 16 h at 37°C, the solution containing the glycans was removed by centrifugation and evaporated to dryness.

**Analysis of K3 production in shake flasks and bioreactor**

The total protein in the supernatant (after 10 min centrifugation at 3000 × g) of flask and fermentation samples were analyzed using the Bio-Rad (Hercules, CA) Protein Assay using bovine serum albumin as a standard. Protein purity was assayed by quantitatively comparing the grayscale intensity of the bands on a SDS-PAGE (4–20% Tris–HCl, stained with Coomassie blue according to Laemmli, 1970) by using a freeware image analysis package (ImageJ 1.29x Wayne Rasband, National Institutes of Health, Java 1.3.1_03). Images of the SDS-PAGE gels were captured using a digital camera and stored in JPEG format.

The amount of K3 produced was estimated from the SDS-PAGE by comparing the grayscale intensity (using the image analysis software package) of the band on an SDS-PAGE corresponding to the molecular weight of K3 with the intensity of a band corresponding to a known amount of purified K3. These results were within 10% of the K3 obtained after purification using Ni-affinity chromatography column (based on the protocol provided by Novagen for HisBind resin) and assessing the amount of purified protein using both the Bio-Rad assay and absorbance at 420 nm after ensuring purity on a SDS-PAGE.

For the shake flask cultivations, a single colony was picked from a YPD plate (<2 weeks old) containing the strain of interest and inoculated into 10 ml BMGY medium in a 50-ml Falcon centrifuge tube. The cultivation was grown to saturation at 24°C (~8 h). The seed culture was transferred into a 500-ml baffled volumetric flask containing 150 ml BMGY medium and grown to OD<sub>600</sub> of 7 ±0.5 at 24°C (~24 h). The growth rate of the cells was determined as the slope of a plot of the natural log of OD<sub>600</sub> against time. The cells were harvested from the growth medium (BMGY) by centrifugation at 3000 × g for 10 min, washed with BMMY, and suspended in 15 ml BMMY in a 250-ml baffled volumetric flask. After 24 h, the expression medium flask was harvested by centrifugation (3000 × g for 10 min) and the supernatant analyzed for K3 production.

For bioreactor cultivations, a 500-ml baffled volumetric flask with 150 ml BMGY media was inoculated with 1 ml seed culture (see flask cultivations). The inoculum was grown to an OD<sub>600</sub> of 4–6 at 24°C (~18 h). The cells from the inoculum culture were then centrifuged and resuspended into 50 ml fermentation media (per L of media: CaSO₄·2H₂O 0.30 g, K₂SO₄ 6.00 g, MgSO₄·7H₂O 5.00 g, glycerol 40.0 g, PTM1 salts 2.0 ml, Biotin 0.20 g, NaI 0.8 g, MnSO₄·7H₂O 3.00 g, NaMoO₄·2H₂O 0.20 g, H₂BO₃ 0.02 g, CoCl₂·6H₂O 0.50 g, ZnCl₂ 20.0 g, FeSO₄·7H₂O 65.0 g, biotin 0.20 g, H₂SO₄ [98%] 5.00 ml). Fermentations were conducted in 3-L dished-bottom (1.5-L initial charge volume) Applikon bioreactors. The fermenters were run in a fed-batch mode at a temperature of 24°C, and the pH was controlled at 4.5 ±0.1 using 30% ammonium hydroxide. The dissolved oxygen was maintained above 40% relative to saturation with air at 1 atm by adjusting agitation rate (450–900 rpm) and pure oxygen supply. The air flow rate was maintained at 1 vvm. When the initial glycerol (40 g/L) in the batch phase is depleted, which is indicated by an increase of dissolved oxygen, a 50% glycerol solution containing 12 ml/L PTM₁ salts was fed at a feed rate of 12 ml/L/h until the desired biomass concentration was reached. After a starvation phase of 30 min, the methanol feed (100% methanol with 12 ml/L PTM₁) is initiated. The methanol feed rate is used to control the methanol concentration in the fermenter between 0.2% and 0.5%. The methanol concentration is measured online.
using a TGS gas sensor (TGS822 from Figaro Engineering, Glenview, IL) located in the offgass from the fermenter. The fermenters were sampled every 8 h and analyzed for bio-
mass (OD 600, wet cell weight and cell counts), residual carbon source level (glycerol and methanol by HPLC
using Aminex 87H) and extracellular protein content (by SDS-PAGE and Bio-Rad protein assay).

MALDI-TOF MS

Molecular weights of the glycans were determined using a Voyager DE PRO MALDI/TOF mass spectrometer
(Applied Biosciences, Foster City, CA) as described previously (Choi et al., 2003). Spectra were generated with
the instrument in the positive ion mode.

In vitro enzymatic assays

For in vitro galactose transfer, 200 μg of the purified K3
produced in the strain PB6P-5 was incubated with 4 mM
UDP-Galactose and 20 to 100 μU of β,1,4-galactosyl-
transferase (bovine milk, Calbiochem) in 50 mM
NH4HCO3, 1 mM MnCl2, pH 7.5 at 37°C for 20 h.
The glycans were then released and analyzed by
MALDI-TOF MS. The negative control without the enzyme
showed no galactose transfer under same conditions.

β-Galactosidase digest was carried out in 50 mM ammo-
nium acetate, pH 5.0, at 37°C for 16–20 h. N-glycans (2 μg)
from RDP93 released from K3 purified from supernatants
were incubated with 3 mU of β,1,4-galactosidase (QA
Bio, San Mateo, CA). Similar conditions were used for
Trichoderma reesei
α,1,2-mannosidase digests (gift from Dr.
Contreras R, Unit of Fundamental and Applied
Molecular Biology, Department of Molecular Biology,
Ghent University, Ghent, Belgium) and jack bean
exoglucanase leads to the
exoglucanase in
Saccharomyces cerevisiae
sequon of yeast exoglucanase in
Saccharomyces cerevisiae

Acknowledgments

We thank Byung-Kwon Choi for scientific discussions. We
also thank Tillman U. Gerngross and Phil Robbins for
guidance and support.

Abbreviations

GnT, N-acetylglucosaminyltransferase; GaT, β,1,4-galac-
tosyltransferase; MALDI-TOF matrix-assisted laser deso-
ration/ionization time-of-flight; MNS, α,1,2-mannosidase;
MS, mass spectrometry, PCR, polymerase chain reaction;
SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel
electrophoresis.

References

characterization of the ALG3 gene of Saccharomyces cerevisiae. Glycobiology, 6, 439–444.

Processing of asparagine-linked oligosaccharides in insect cells. N-
acetylglucosaminyltransferase I and II activities in cultured lepidop-

sequeom of Saccharomyces cerevisiae exoglucanase leads to the
secretion and secretion of a new glycoform of the molecule. Yeast, 9,
221–234.

Kinetic mechanism, substrate specificity, and inhibition characteristics
of UDP-N-acetylglucosamine:alpha-D-mannoside beta-1,2-N-acetyl-
glucosaminyltransferase II from rat liver. J. Biol. Chem., 262,
5784–5790.


Changing the donor cofactor of bovine alpha-1,3-galactosyltransfer-
ase by fusion with UDP-galactose 4-epimerase. More efficient
275, 31594–31600.

Cipollo, J.F. and Trimble, R.B. (2000) The accumulation of Man(6-
GlcNAc(2))-PP-dolichol in the Saccharomyces cerevisiae Deltaalg9
mutant reveals a regulatory role for the Alg3p alpha 1,3-mannide
arm addition in downstream oligosaccharide-lipid and glycoprotein

Cueva, R., Munoz, M.D., Andaluz, E., Basco, R.D., and Larriba G.
(1996) Preferential transfer to truncated oligosaccharides to the first

Davidson, R.C., Blankenship, J.R., Kraus, P.R., de Jesus Berrios, M.,
based strategy to generate integrative targeting alleles with large
regions of homology. Microbiology, 148, 2607–2615.

Use of combinatorial genetic libraries to humanize N-linked glycosyla-
tion in the yeast Pichia pastoris. Proc. Natl Acad. Sci. USA, 100, 5022–
5027.


oligosaccharide structures found in various yeast species. Biochim.

Wildt S. (2004) Functional analysis of the ALG3 gene encoding the
Dol-P-Man:ManGlcNAc2-PP-Dol mannosyltransferase enzyme of


Kornfeld, S., Gregory, W., and Chapman, A. (1979) Class E Thy-
1 negative mouse lymphoma cells utilize an alternate pathway of

Kukuruzinska, M.A. and Lennon, K. (1998) Protein N-glycosylation:


processing inhibitors in Chinese hamster ovary cells. The role of a
novel mutation in the asparagine-linked glycosylation pathway. J. Biol.
Chem. 264, 1584–1593.

of the asparagine-linked oligosaccharide processing pathway. Glyco-
biochemistry, 4, 113–125.

Munoz, M.D., Hernandez, L.M., Basco, R., Andaluz, E., and Larriba, G.
(1994) Glycosylation of yeast exoglucanase sequons in alg mutants
deficient in the glycosylation steps of the lipid-linked oligosaccharide.


