

THE HUMANIZATION OF N-GLYCOSYLATION PATHWAYS IN YEAST

Stefan Wildt* and Tillman U. Gerngross[‡]

Abstract | Yeast and other fungal protein-expression hosts have been extensively used to produce industrial enzymes, and are often the expression system of choice when manufacturing costs are of primary concern. However, for the production of therapeutic glycoproteins intended for use in humans, yeast have been less useful owing to their inability to modify proteins with human glycosylation structures. Yeast *N*-glycosylation is of the high-mannose type, which confers a short half-life *in vivo* and thereby compromises the efficacy of most therapeutic glycoproteins. Several approaches to humanizing yeast *N*-glycosylation pathways have been attempted over the past decade with limited success. Recently however, advances in the glycoengineering of yeast and the expression of therapeutic glycoproteins with humanized *N*-glycosylation structures have shown significant promise — this review summarizes the most important developments in the field.

Recombinant DNA technology, and the ability to produce recombinant proteins, has made important contributions to human health and has profoundly impacted on scientific discovery itself. Beginning with the introduction of recombinant forms of human insulin (formerly extracted from animal sources), many therapeutic proteins are now produced in recombinant form, including erythropoietins, interferons, blood-clotting factors and monoclonal antibodies. However, many of these proteins require post-translational processing — mostly glycosylation — to attain full biological and therapeutic function. Yeasts, although generally viewed as attractive fermentation organisms and hosts for recombinant protein expression, are unable to perform human glycosylation reactions (FIG. 1).

Obtaining a properly folded and active protein is usually the first requirement of a recombinant protein-expression host, although some proteins can be obtained as insoluble inclusion bodies and reconstituted by denaturation and refolding. The ease of handling and rapid read-out often make *Escherichia coli* the organism of choice; however, many proteins cannot be obtained

in an active and soluble form in this host. The folding environment in the cytosol and the requirement for glycosylation are often responsible for this. Yeast can be used to overcome some of these shortcomings, and the expression of recombinant proteins in yeast goes back to the early days of recombinant DNA technology¹.

Initial attempts to express recombinant proteins in yeast focused on the 'baker's yeast' *Saccharomyces cerevisiae*, which has become a well-understood model organism with an extensive body of literature describing its cellular and molecular biology. An abundance of molecular biological tools are available to the yeast biologist, including genome sequences, the ability to perform direct gene knockouts, gene replacements and site-specific chromosomal integration, and the availability of recyclable genetic markers allowing for the repeated introduction of new genetic constructs. In addition, a variety of strong constitutive promoters (for example, the glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) promoter) and inducible promoters (such as the *GALI* promoter) have been developed. Over the past decade additional yeast expression systems

*GlycoFi Inc., 21 Lafayette Street, Lebanon, New Hampshire 03766, USA.
[‡]Thayer School of Engineering, Department of Biological Sciences, and Department of Chemistry, Dartmouth College, Hanover, New Hampshire 03755, USA.
 Correspondence to T.U.G.
 e-mail: tillman.gerngross@dartmouth.edu
 doi:10.1038/nrmicro1087

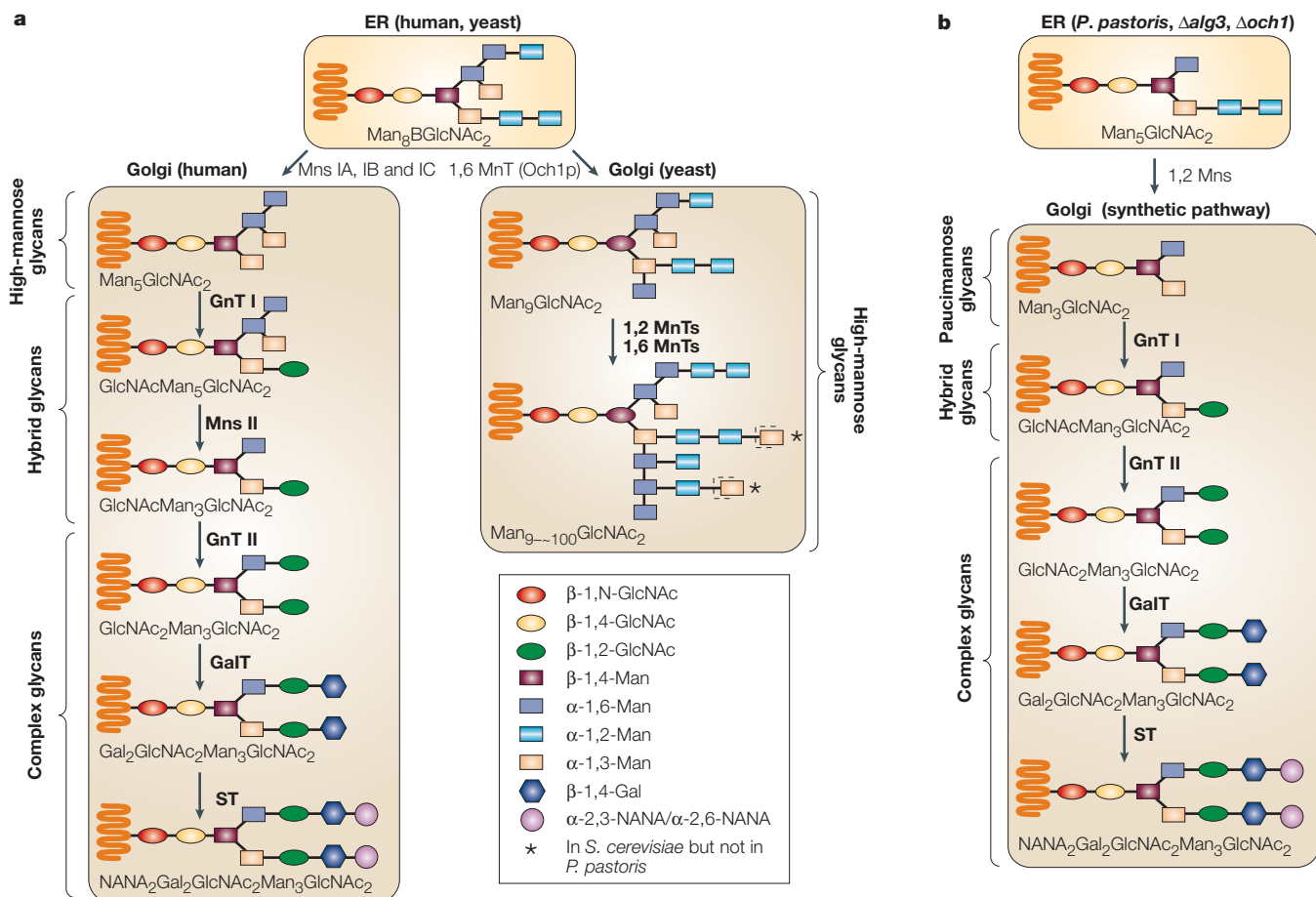


Figure 1 | **Major N-glycosylation pathways in humans and yeast.** **a** | Representative pathway of N-glycosylation pathways in humans (left) provides a template for humanizing N-glycosylation pathways in yeast (right). **b** | Early oligosaccharide assembly mutants can be used to recreate synthetic glycosylation pathways that lead to complex N-glycosylation in yeast (see main text). ER, endoplasmic reticulum; GalT, galactosyltransferase; GlcNAc, N-acetylglucosamine; GnT I, N-acetylglucosaminyl transferase I; GnT II, N-acetylglucosaminyl transferase II; Man, mannose; Mns II, mannosidase II; MnTs, mannosyltransferase; NANA, N-acetylneuraminic acid; ST, sialyltransferase.

have evolved, driven by the need for stronger inducible promoters and the desire to secrete the recombinant product into the culture medium to facilitate downstream recovery^{2,3}. Three yeast expression hosts are of note in this context — *Kluyveromyces lactis* and the methylotrophic yeasts *Pichia pastoris* and *Hansenula polymorpha*. Each of these systems has been used to produce recombinant enzymes on a commercial scale and there is an extensive body of literature describing their use as hosts for recombinant protein expression (reviewed in REFS 3–6).

Industrial yeast fermentations are characterized by high cell densities, short fermentation times (several days) and the use of chemically defined media that do not require components derived from animals (for example, calf serum), which can be of concern in the production of therapeutic products for human use due to the possibility of contamination with viruses or prions. In these systems, recombinant proteins, including mammalian proteins, can often be expressed at high protein titres, as shown by the production of 14.8 g l⁻¹ of gelatin in *P. pastoris*⁷.

Therapeutic proteins can typically be divided into two major categories — those with N-linked glycosylation (that is, glycoproteins) and those without. N-linked glycosylation often constitutes a small fraction of the molecular mass of a given protein, yet it contributes significantly to its hydrodynamic volume and therefore to its pharmacodynamic behaviour (FIG. 2). Until recently, it was thought that N-linked glycosylation occurs exclusively in eukaryotes, in which a pre-assembled oligosaccharide is transferred in a site-specific manner to an asparagine residue of the nascent protein in the endoplasmic reticulum (ER). Prokaryotes such as *E. coli* do not have an ER, and typically do not perform N-linked glycosylation reactions. Recently however, this notion was challenged by the finding that the prokaryotic organism *Campylobacter jejuni* contains N-glycosylated proteins, albeit with N-glycan structures that differ markedly from those that are found in eukaryotes^{8–10}, as well as O-glycosylated proteins¹¹.

Currently, most non-glycosylated therapeutic proteins are expressed in either *E. coli* or yeast. For example, about half of the world's supply of insulin is

made in yeast by Novo Nordisk, whereas Eli Lilly & Co. produces insulin in *E. coli*. Although almost one-sixth of the currently approved therapeutic proteins are produced in yeast, including recombinant granulocyte macrophage colony-stimulating factor (GM-CSF), insulin and several vaccines (for example, the first therapeutic protein to be expressed in yeast was the hepatitis B vaccine¹²), yeast have typically not been used to make recombinant glycoproteins owing to the non-human *N*-glycosylation reactions that are associated with yeast protein expression (FIG. 1). Yeast *N*-glycosylation is of the high-mannose type, which in humans mediates binding to mannose receptors and results in poor pharmacokinetic behaviour¹³.

Recombinant glycoproteins constitute about 70% of all approved therapeutic proteins and, given their requirement for human-like *N*-glycosylation, have typically required the use of mammalian cell lines for their production (for example, CHO or BHK cell lines), which are able to perform human-like *N*-glycosylation reactions, although the end product invariably differs in composition to its human counterpart. Examples of therapeutic glycoproteins that have been approved and are produced in mammalian cell lines include recombinant interferon- β , which is used for the treatment of multiple sclerosis; erythropoietin, which is used for the treatment of anaemia; and recombinant antibodies, such as Herceptin and Rituxan, which are used to treat breast cancer and non-Hodgkin's lymphoma, respectively. Mammalian cell culture is an expensive and lengthy process that affords little control over the final product. The generation of mammalian production cell lines is a cumbersome task that typically takes several months. In addition, mammalian cells are very sensitive to environmental changes during fermentation. An increase in the heterogeneity of *N*-glycosylation, and therefore a decrease in product quality, has been correlated with several factors including pH, process time, glutamine levels and the availability of nucleotide sugars.

Notwithstanding these difficulties, in the absence of better alternatives the use of mammalian cell culture has emerged as the dominant platform for the production of recombinant human glycoproteins for therapeutic purposes. The market for therapeutic glycoproteins has experienced robust growth (25.8% annual growth over the past five years¹⁴), and recent interest has begun to focus on protein expression platforms that are able to reduce costs and offer greater control over the post-translational processing of the final product. The latter has become increasingly important as the extent and type of *N*-glycosylation have been found to have a marked impact on the therapeutic properties of many commercially relevant therapeutic proteins^{15–17}.

Engineering *N*-glycosylation in yeast

The production of recombinant proteins involves three main steps that contribute about equally to the cost of the final product: protein production (that is, fermentation); protein purification; and 'fill and finish', which is, the filling and packaging of a formulated and stabilized product. As the second and third steps are essentially fixed

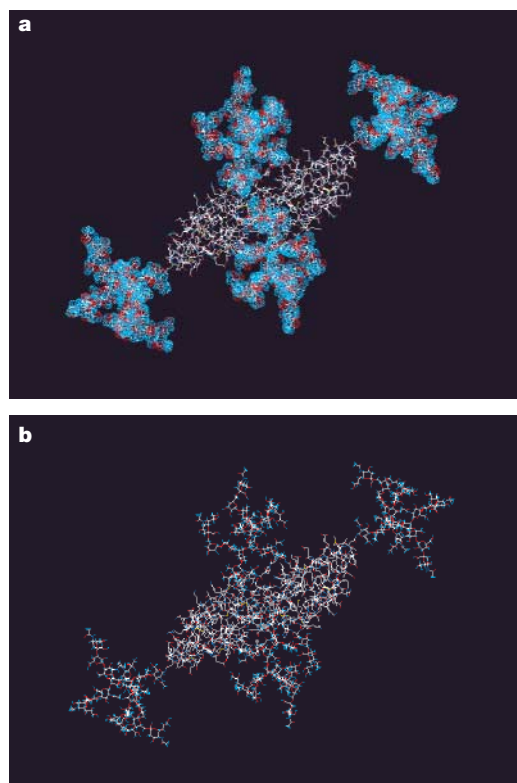


Figure 2 | Structure of glycosylated glucocerebrosidase. Panels show glycosylated glucocerebrosidase with (a) and without (b) glycans highlighted to illustrate the extent to which glycans influence the hydrodynamic volume and surface properties of glycoproteins.

expenses that are required irrespective of the production process, most efforts to reduce costs have focused on improving the fermentation process by increasing productivity, and hence the titre and quality of the final product. This can be achieved by improving the fermentation process itself or by considering alternative production systems. Although yeast have proven to be robust protein-expression platforms for many industrial enzymes and some non-glycosylated therapeutic proteins, as already discussed, the inability to perform human-like *N*-glycosylation reactions has precluded yeast from being used to produce most recombinant human glycoproteins intended for therapeutic use in humans. To overcome this shortcoming, several laboratories have been investigating the possibility of altering endogenous *N*-glycosylation reactions and genetically engineering human *N*-glycosylation pathways into yeast and other fungi.

N-glycosylation pathways: human versus yeast

Two types of protein glycosylation can be distinguished — *O*-linked and *N*-linked. *O*-linked glycosylation is the attachment of carbohydrates to the hydroxyl groups of serine or threonine residues of proteins. Unlike *N*-glycosylation, which has been shown to be crucial for protein function, relatively little is known about *O*-glycosylation and its biological role (reviewed in REFS 18–20). Given the importance of *N*-glycosylation for therapeutic functions, this review focuses on

Table 1 | Yeast *N*-glycosylation

Reaction	<i>S. cerevisiae</i>	<i>K. lactis</i>	<i>P. pastoris</i>	<i>H. polymorpha</i>
ER α -1,2-mannosidase	+	–	+	?
Extensions to Man₈/Man₉ core				
α -1,3-mannose	+	+	–	–
α -1,2-mannose	+	+	+	+
Mannosylphosphate	+	–	+	?
α -1,6-mannose	+	+	+	+

efforts to modify *N*-glycosylation in yeast and other fungal protein-expression hosts; however, similar efforts are underway to modify and humanize *O*-glycosylation in yeast.

Re-engineering a protein-expression host to replicate human *N*-glycosylation reactions requires an inventory of all the *N*-glycosylation reactions that are known to occur in that host (that is, the glycome). By the early 1990s, much was known about the main *N*-glycosylation pathways in several yeasts and in humans, and a significant amount of conservation was found in the early steps of *N*-glycan assembly and processing^{21–23} (FIG. 1).

As in humans and other higher mammals, yeast *N*-glycosylation is initiated in the ER, where a core oligosaccharide (Glc₃Man₉GlcNAc₂) is transferred onto the nascent polypeptide. This site-specific transfer seems to be conserved in higher and lower eukaryotes and involves the recognition of the sequence Asn-X-Ser/Thr (where X is any amino acid other than proline). Following transfer of the core oligosaccharide to the asparagine residue within the Asn-X-Ser/Thr motif, three glucose moieties and one terminal α -1,2-mannose moiety are removed by glucosidase I and glucosidase II, and an ER-residing α -1,2-mannosidase, respectively. The resulting Man₈GlcNAc₂-containing glycoprotein is then transported to the Golgi apparatus where *N*-glycan processing differs markedly between yeast and mammals (FIG. 1). It is of note that some yeasts such as *Schizosaccharomyces pombe* and *K. lactis* seem to lack ER-specific α -1,2-mannosidase activity, resulting in the transport of Man₉GlcNAc₂-containing glycoproteins to the Golgi (REF. 24; S.W. and T.U.G., unpublished observations).

In humans and other mammals, early Golgi *N*-glycan processing involves the trimming of Man₈GlcNAc₂ to Man₅GlcNAc₂ by α -1,2-mannosidase(s), a process that generates the substrate for *N*-acetylglucosaminyl transferase I (GlcNAcT I), which transfers a single *N*-acetylglucosamine (GlcNAc) sugar onto the terminal 1,3-mannose of the tri-mannose core. Following this transfer, mannosidase II removes the two remaining α -1,3- and α -1,6 terminal mannose sugars to produce GlcNAcMan₃GlcNAc₂. This is the substrate for *N*-acetylglucosaminyl transferase II (GlcNAcT II), which adds one GlcNAc sugar to the terminal α -1,6-mannose arm of the tri-mannose core^{25–29} (FIG. 1). Further processing typically involves the attachment of additional GlcNAc³⁰, galactose and *N*-acetylneuraminic acid (NANA) moieties. However, a host of additional glycosyltransferases, including fucosyltransferases, GalNAc transferases and

GlcNAc phosphotransferases, are known to exist, which further broadens the range of *N*-glycans found on proteins isolated from human sources. It is of note that many glycosylation reactions are not complete and analysis of proteins purified from human serum typically reveals several intermediates ranging from high-mannose glycans to complex, terminally sialylated glycans. Many of these glycan structures are thought to confer specific functions, yet the elucidation of glycosylation-dependent structure–activity relationships is far from complete and remains an active area of research.

In contrast to human *N*-glycan processing, which involves the removal of mannose followed by the addition of GlcNAc, galactose, fucose and NANA, early *N*-glycan processing in yeast is limited to the addition of mannose and mannosylphosphate sugars. The Golgi apparatus of *S. cerevisiae* contains α -1,2-, α -1,3- and α -1,6-mannosyltransferases as well as mannosylphosphate transferases, which produce *N*-glycan structures that are mannosylated and hypermannosylated to varying extents. In *K. lactis*, *H. polymorpha* and *P. pastoris*, a similar set of mannosyltransferases exists, resulting in the production of mostly high-mannose structures that resemble those produced in *S. cerevisiae* but which are typically smaller in size (TABLE 1; FIG. 1). It is of note that *P. pastoris* lacks a Golgi-resident α -1,3-mannosyltransferase yet seems to contain at least one β -mannosyltransferase; it has also been shown that *P. pastoris* has a secretory pathway with distinct Golgi stacks similar to those found in mammalian cells (FIG. 3).

Early approaches

Once the main *N*-glycosylation reactions for a given organism have been elucidated, a rational strategy for re-engineering human *N*-glycan processing can be devised. This typically involves the elimination of non-human *N*-glycosylation reactions followed by the introduction of human *N*-glycosylation reactions. Early efforts focused on the yeast *S. cerevisiae*, in part because many of the *N*-glycosylation pathways had been characterized and some of the corresponding genes had been cloned³¹. In particular, the α -1,6-mannosyltransferase **Och1p** of *S. cerevisiae*, which initiates the synthesis of the outer chain, became the focus of initial attempts to eliminate yeast-specific hyper-mannosylation. The first such report by Jigami and colleagues showed that a double knockout of *S. cerevisiae* lacking *och1* (which encodes an α -1,6-mannosyltransferase) and *mmn1* (encoding an α -1,3-mannosyltransferase) was able to produce mostly high-mannose glycans of the Man₈GlcNAc₂ type³². As this structure is an early intermediate of human *N*-glycosylation, it enabled further humanization of *N*-glycosylation pathways in this yeast. Processing of Man₈GlcNAc₂ in the human Golgi involves the removal of three α -1,2-mannose sugars by Golgi-resident α -1,2-mannosidases to generate Man₅GlcNAc₂²⁶. Replicating this reaction in yeast remained problematic until 1998, when Chiba and co-workers first demonstrated that some Man₅GlcNAc₂ could be produced by localizing a fungal α -1,2-mannosidase from *Aspergillus*

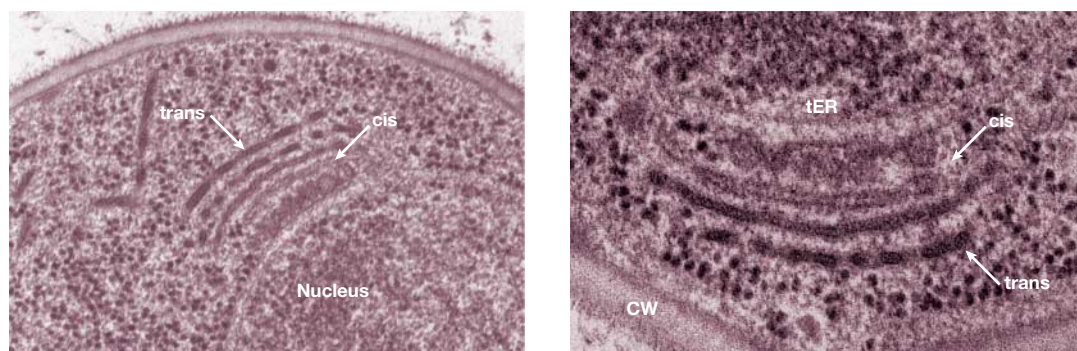


Figure 3 | **Electron micrographs of *Pichia pastoris*.** The cellular organization of the various organelles (endoplasmic reticulum (ER) and Golgi apparatus) that constitute the secretory pathway can be seen. It is of note that this yeast has a secretory pathway that more closely resembles the mammalian counterpart in its organization. Cis and trans indicate the cis and trans Golgi apparatus. Images provided by Soren Mogelsvang and Nathalia Gomez-Ospina, University of Colorado School of Medicine. CW, cell wall; tER, transitional ER.

saitoi in the ER of *S. cerevisiae*. However, despite the use of a strong constitutive promoter (*GAPDH*) and a multicopy 2 μ plasmid, only ~27% of the glycans of a reporter protein (carboxypeptidase Y) contained the desired Man₅GlcNAc₂ structure, with the remainder being Man₆GlcNAc₂–Man₈GlcNAc₂³³.

Other groups attempted to re-engineer glycosylation pathways by expressing various enzymes that are involved in early *N*-glycan processing in humans, yet failed to eliminate the competing endogenous glycosylation pathways or chose hosts that did not provide an appropriate substrate. Two such early studies are of note. First, Harry Schachter and co-workers showed in 1995 that active GlcNAcT I could be expressed in the filamentous fungus *Aspergillus nidulans*, and that it could be assayed in cell-free extracts, yet the proteins produced by the engineered strain showed no transfer of GlcNAc to *N*-glycans³⁴. Several explanations were proposed including improper localization of GlcNAcT I, the absence of a Man₅GlcNAc₂ acceptor substrate at the site of GlcNAcT I localization, conditions at the site of localization that were not conducive to GlcNAcT I activity (for example, incorrect pH), lack of UDP-GlcNAc at the site of GlcNAcT I localization or any combination of these factors. Another important study by Roland Contreras and colleagues showed that the expression of an α -1,2- mannosidase from *Trichoderma reesei* in the secretory pathway of *P. pastoris* resulted in the production of higher molecular mass glycans – which is in contrast to the smaller molecular mass glycans one would expect if the enzyme is active in the secretory pathway³⁵. These studies highlight the importance of ensuring the absence of competing endogenous reactions and ensuring that an appropriate substrate is available for the corresponding glycosylation enzyme. As such, the work by Jigami and co-workers is the first report in which both elements — elimination of endogenous pathways and the introduction of an exogenous α -1,2-mannosidase — are combined to show the conversion of Man₈GlcNAc₂ to the desired Man₅GlcNAc₂ human intermediate, albeit at suboptimal efficiency³³.

Efficient replication of human glycosylation

In yeast and humans, host-specific glycosyltransferases and glycosidases line the luminal surface of the ER and Golgi apparatus, and thereby provide catalytic surfaces that allow the sequential processing of glycoproteins as they proceed from the ER through the Golgi network into the medium (FIG. 4). As a glycoprotein proceeds from the ER through the secretory pathway, it is sequentially exposed to different mannosidases and glycosyltransferases. Several processing steps rely on previous reactions because some *N*-glycosylation enzymes depend on a particular substrate that is created by the previous enzyme — for example, mannosidase II can only trim terminal α -1,3- and α -1,6-mannoses from the pentamannose core if a terminal GlcNAc is present on the α -1,3 arm (FIG. 1). *N*-glycosylation enzymes must therefore be arranged in a predetermined sequence to allow for the synthesis of specific *N*-glycan structures.

Targeting to compartments of the secretory pathway.

Establishing the sequential processing environments of the secretory pathway requires the proper localization of *N*-glycosylation enzymes. The mechanisms by which secreted proteins can be transported through the secretory pathway (from the ER to the cis-, medial- and trans-Golgi compartments and into the medium), while each compartment maintains a specific set of resident (for example, *N*-glycosylation) enzymes, has been the subject of extensive study and remains an active area of basic cell biological research^{36–38}.

Two well-established mechanisms that localize proteins to the various compartments of the secretory pathway are retrieval and retention (reviewed in REFS 36,39). Retrieval is a process by which proteins are localized to certain organelles through interaction with other proteins. Several ER-residing proteins contain a carboxy-terminal tetrapeptide with the consensus sequence KDEL (or HDEL in yeast), which has been shown to be required for efficient localization to the ER^{40,41}. This tetrapeptide is recognized by the H/KDEL receptor *Erd2p*, which binds to the H/KDEL sequence in the Golgi and releases the protein on retrograde

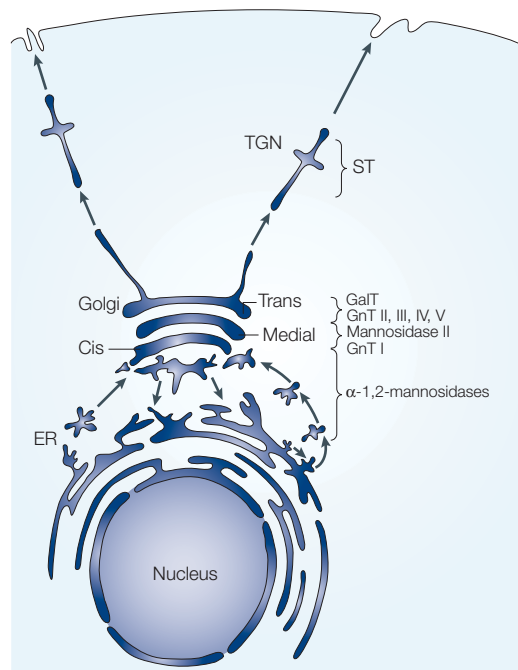


Figure 4 | A working model for the cellular distribution of glycosyltransferases throughout the secretory pathway.

Specific glycosyltransferases and glycosidases line the luminal surface of the endoplasmic reticulum (ER) and Golgi, allowing the sequential processing of glycoproteins as they are shuttled through the secretory pathway. GalT, galactosyltransferase; GnT, *N*-acetylglucosaminyltransferase; ST, sialyltransferase; TGN, trans-Golgi network.

transport back to the less-acidic ER^{42–44}. At steady state, this results in the predominant localization of HDEL-tagged proteins to the ER. It has also been shown that the HDEL retention mechanism can be saturated by overexpressing proteins containing the HDEL tetrapeptide; presumably Erd2p can only bind a finite amount of HDEL and, once saturated, is no longer able to efficiently retrieve HDEL-containing proteins, resulting in their secretion into the medium⁴².

Several ER- and Golgi-residing enzymes, such as α -1,2-mannosidase, mannosidase II and GlcNAcT I, are type II membrane proteins (FIG. 5). These proteins have a common domain structure comprising a short cytoplasmic tail at the amino terminus, a hydrophobic transmembrane domain, a luminal stem and a C-terminal catalytic domain⁴⁵. Deletion studies as well as fusions to non-Golgi-residing proteins have identified the N-terminus, and in particular the transmembrane region, as containing the targeting information of many type II membrane proteins. Although it is clear that N-terminal domains are involved in targeting, the extent to which their targeting ability is transferable between different species is unclear. For example, it has been shown that, when expressed in *S. cerevisiae*, full-length human 2,6-sialyltransferase and galactosyltransferase (GalT) localized to the ER, whereas in mammalian cells the two enzymes are found in the trans-Golgi network and the late Golgi, respectively⁴⁶. By contrast, the N-terminal transmembrane domain of rat 2,6-sialyltransferase is able

to localize yeast invertase to the Golgi of *S. cerevisiae*⁴⁷. Yoshida and colleagues showed that rat GlcNAcT I, typically an early- to mid-Golgi enzyme, localizes to the ER, Golgi and vacuole when expressed in *S. cerevisiae*⁴⁸. Schwientek and co-workers demonstrated that full-length human GalT cannot be expressed in an active form despite considerable transcript levels, yet an active enzyme can be obtained by fusing the catalytic domain of the enzyme to the leader sequence from a type II membrane protein from yeast (MntI). A third localization mechanism involves the association of different Golgi enzymes with one another, a process known as 'kin recognition'. For example, adding an ER-retention motif to GlcNAcT I results in the localization of this Golgi enzyme to the ER, yet mannosidase II, a medial Golgi enzyme, is now also partially localized to the ER⁴⁹. Although these studies have established a correlation between certain domains and targeting events, there is currently no tool that allows us to predict the behaviour of a given targeting peptide across different hosts and what role the corresponding catalytic domain might have in this process.

Notwithstanding this uncertainty, several researchers have tried to exploit targeting mechanisms empirically to establish human-like *N*-glycosylation pathways in heterologous organisms. Early efforts to localize mannosidases in the secretory pathway of various yeasts have shown some promise. Martinet and co-workers demonstrated that an α -1,2-mannosidase from *T. reesei* containing a putative N-terminal transmembrane domain is not efficiently retained when expressed in *P. pastoris*³⁵. Fusing the catalytic domain of the same enzyme to the leader sequence of *S. cerevisiae* Mns1p prevented secretion but did not result in mannose trimming³⁵. Chiba and co-workers reported similar findings when using the leader sequence of *S. cerevisiae* Och1p to localize the α -1,2-mannosidase from *A. saitoi* to the Golgi of *S. cerevisiae*³³. Although this fusion localized to the Golgi and the cytosol, Man₈GlcNAc₂ trimming was not observed. It was only when an HDEL retrieval mechanism was used to localize the catalytic domain of α -1,2-mannosidase to the ER that some mannosidase activity could be detected on an intracellular reporter protein. However, the amount of mannose trimming was small — only 27% of glycans had the desired Man₅GlcNAc₂ structure. As it was shown that at least some mannosidase was localized within the cell (mostly in the ER) and an intracellular reporter was used, there are only two explanations for these findings — either the mannosidase is not fully active at the site of localization or endogenous mannosyltransferases of the Golgi apparatus are able to outcompete the exogenous mannosidase, thereby preventing the formation of appreciable amounts of Man₅GlcNAc₂. A similar approach in *P. pastoris* demonstrated that the HDEL sequence was able to localize an α -1,2-mannosidase–HDEL fusion to the ER⁵⁰; however, when the construct was expressed under the control of the strong *AOX1* promoter, the α -1,2-mannosidase–HDEL fusion protein was also secreted into the medium, consistent with previous findings⁴². Although this work demonstrated

efficient mannosidase trimming, the problem with studies involving HDEL-targeted α -1,2-mannosidases is the ambiguity of where the mannose trimming actually occurs when secreted reporter proteins are used⁵¹. As HDEL-tagged mannosidases always have the potential to secrete into the medium, detecting $\text{Man}_3\text{GlcNAc}_2$ on a secreted glycoprotein can have two possible causes: the mannosidase was active *in vivo* and trimming occurred inside the cell, or trimming occurred post-secretion in the medium. True intracellular trimming can be demonstrated by the absence of mannosidase activity in the medium or by engineering a second processing step that can occur only *in vivo*, such as the addition of GlcNAc, which requires an intracellular pool of UDP-GlcNAc and intracellular GlcNAcT I activity.

The evolving picture from these studies is complex and it is not possible to reliably predict which leader sequences can be used to localize catalytic domains to specific regions of the secretory pathway, and which catalytic domains might in fact be active when properly localized. To overcome these shortcomings, our laboratory designed a genetic library of known type II membrane protein domains that encode peptides that are associated with proteins that naturally localize to the ER and Golgi of *S. cerevisiae* or *P. pastoris*⁵². The library includes N-terminal fragments of **Gls1p**, **Mns1p**, **Sec12p**, **Mnn9p**, **Van1p**, **Anp1p**, **Hoc1p**, **Mnn10p** and **Mnn11p** from *S. cerevisiae* and **Och1p** and **Sec12p** from *P. pastoris*. A second library contained catalytic domains of several α -1,2-mannosidases from *Homo sapiens*, *Mus musculus*, *Aspergillus nidulans*, *Caenorhabditis elegans*, *Drosophila melanogaster* and *Penicillium citrinium*, all of which lack their endogenous N-terminal leader sequence. The libraries were designed in such a way that any combination of a leader construct and a catalytic domain resulted in a chimeric fusion protein between a yeast leader sequence and a catalytic domain. A *P. pastoris* strain that is deficient in initiating 1,6-mannosyltransferase activity (Δoch1) but able to secrete a hexahistidine-tagged fragment of human plasminogen as a reporter protein was used to receive the combinatorial library of 608 chimeric fusion constructs. On screening of all fusion constructs, only a few efficiently trimmed $\text{Man}_8\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$, while concomitantly lacking measurable mannosidase activity in the medium. This work further demonstrated the importance of proper cellular targeting and the fact that specific interactions between yeast leader sequences and catalytic domains are necessary to achieve proper localization and efficient intracellular mannosidase activity.

In mammals, the next step towards the formation of complex human glycoproteins involves the addition of *N*-acetylglucosamine to the terminal α -1,3-mannose moiety of $\text{Man}_3\text{GlcNAc}_2$. This reaction occurs early in the Golgi and is catalysed by GlcNAcT I, an enzyme that requires UDP-GlcNAc as an energy-rich sugar donor and a glycoprotein containing an appropriate *N*-glycan substrate such as $\text{Man}_3\text{GlcNAc}_2$. As mentioned above, rabbit GlcNAcT I has been introduced into the filamentous fungus *A. nidulans*

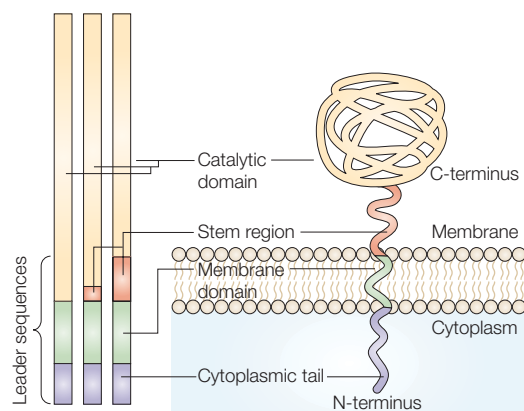


Figure 5 | Type II membrane proteins. The anchoring mechanism for most glycosyltransferases is through a type II transmembrane anchor, by which the C-terminal catalytic domain is anchored to the membrane through an N-terminal transmembrane domain.

without finding any addition of terminal GlcNAc to glycoproteins.

Given the unsuccessful attempts to produce the hybrid glycan intermediate $\text{GlcNAcMan}_3\text{GlcNAc}_2$ in yeast, we concluded that several issues had to be addressed. First, one had to ensure the availability of appropriate substrate ($\text{Man}_3\text{GlcNAc}_2$) within the yeast (see above); second, one had to express and localize an active GlcNAcT I enzyme; and third, one had to ensure the availability of UDP-GlcNAc in the Golgi to enable the interaction of substrate, sugar nucleotide donor and the catalysing enzyme GlcNAcT I. Given the assembly of GlcNAc-containing dolichol-linked precursors on the cytosolic side of the ER, it was assumed that yeast have an endogenous pool of UDP-GlcNAc in the cytosol. After showing that $\text{Man}_5\text{GlcNAc}_2$ can be efficiently produced in *P. pastoris*, we proceeded to screen a library of GlcNAcT I fusions in a strain that also expressed the UDP-GlcNAc transporter from *K. lactis*. This screen yielded strains of *Pichia* that had the ability to glycosylate a secreted reporter protein (kringle 3 of human plasminogen) with an essentially uniform hybrid glycan $\text{GlcNAcMan}_3\text{GlcNAc}_2$ ⁵². The reporter protein used in these studies is expressed at about 1 g l^{-1} , and we have found that the presence of a UDP-GlcNAc transporter is necessary for complete conversion of $\text{Man}_3\text{GlcNAc}_2$ to $\text{GlcNAcMan}_3\text{GlcNAc}_2$ on highly expressed proteins. Unlike glycosyltransferases, which can be linked to any transmembrane leader sequence to influence their targeting, sugar nucleotide transporters typically contain multiple transmembrane domains, which makes domain swapping a daunting proposition.

Removing more mannose

In mammals, $\text{GlcNAcMan}_3\text{GlcNAc}_2$ is the substrate for mannosidase II, which is a 131-kDa Golgi-resident protein that removes one terminal α -1,3- and one terminal α -1,6-mannose to yield $\text{GlcNAcMan}_3\text{GlcNAc}_2$. This reaction is dependent on a GlcNAc residue being

present on the α -1,3 arm of the tri-mannose core, and therefore requires either co-localization with GlcNAcT I or location downstream of GlcNAcT I in the secretory pathway. Given the size of mannosidase II and the absence of reports demonstrating successful expression of this enzyme in yeast, we were concerned about the ability to localize an active version of mannosidase II in the secretory pathway of *P. pastoris*. We therefore chose to pursue two alternative approaches: one in which we recreated the human pathway based on mannosidase II and a second in which we recreated a synthetic pathway based on mutants in which core oligosaccharide assembly is prematurely terminated.

The second approach was based on earlier reports that some *N*-glycosylation mutants in *S. cerevisiae* were able to transfer truncated lipid-linked structures onto glycoproteins. As the formation of complex glycans in humans involves a series of lipid-linked assembly steps, followed by a transfer step to the protein and a series of post-transfer mannosidase-removal steps, we reasoned that blocking the addition of certain mannose (terminal α -1,3- and terminal α -1,6-mannose of the pentamannose core) moieties before the transfer of the core oligosaccharide onto the protein could obviate the need to introduce an active mannosidase II (FIG. 1). In *S. cerevisiae*, such mutants have been described, and **Alg3p**, which adds the first α -1,3-mannose to the α -1,6 arm of the tri-mannose core, is of particular interest, as a lack of Alg3p activity was shown to prevent the subsequent addition of α -1,6-mannose to the α -1,6 arm of the tri-mannose core, resulting in unnatural Man₃GlcNAc₂ structures that only contain α -1,2-mannoses attached to the tri-mannose core^{53–55} (FIG. 1b). This structure can be converted to a paucimannose structure (tri-mannose core) by the action of a single enzyme (α -1,2-mannosidase), and since earlier work by Harry Schachter and colleagues had shown that GlcNAcT I is able to transfer GlcNAc onto a tri-mannose core *in vitro*, we assumed that a synthetic pathway leading to complex glycans in the absence of mannosidase II can be established *in vivo*. However, the extent to which an *alg3*, *och1* double mutant would be viable, whether the *alg3* deletion is sufficiently tight to prevent subsequent addition of α -1,6-mannose by Alg12p and whether occupancy could be adversely impacted were not known.

We have recently shown that a viable *alg3*, *och1* double mutant of *P. pastoris* can be obtained and that the introduction of active α -1,2-mannosidase, GlcNAcT I, GlcNAcT II and a UDP-GlcNAc transporter allows for the generation of uniform complex glycoproteins with a terminal GlcNAc⁵⁶. Here again, combinatorial leader and catalytic domain fusions were used to find enzymes with optimal activity in the secretory pathway. In parallel, we also tried to recreate the mammalian *N*-glycosylation pathway based on the expression of mannosidase II. In this work, we were able to show that catalytic domains of mannosidase II and GlcNAcT II can be introduced into *P. pastoris* to generate strains that have the ability to secrete uniform complex glycans with terminal GlcNAc moieties⁵⁷.

Further maturation – galactosylation

The penultimate step in obtaining terminally sialylated glycoproteins involves the addition of galactose to terminal GlcNAc sugars. These GlcNAc sugars are attached to the paucimannose core, which can contain between two and four terminal GlcNAc sugars, resulting in two, three or four ‘antennae’ that can be extended by the action of β -1,4-GalT. The transfer of galactose requires a productive interaction between a terminal GlcNAc-containing glycoprotein, UDP-galactose and active β -1,4-GalT. Schwientek and co-workers demonstrated the ability to localize active human β -1,4-GalT in the Golgi of *S. cerevisiae* using an elegant approach that mimics the natural substrate. By using an *alg1* mutant of *S. cerevisiae*, which, at the non-permissive temperature, transfers short GlcNAc₂ ‘stumps’ onto glycoproteins, the authors were able to produce an artificial substrate for GalT and demonstrate its modification by GalT *in vivo*. Lectin-binding assays indicated that only a small amount of terminal GlcNAc received a galactose, suggesting that transfer is not efficient⁵⁸. In another approach, the group led by Roland Contreras introduced a Golgi-targeted β -1,4-GalT into a *P. pastoris* strain that was able to produce hybrid GlcNAcMan₃GlcNAc₂ *N*-linked glycans, and here also only incomplete transfer onto the hybrid acceptor substrate was observed⁵⁹.

Our group recently reported two approaches to obtain essentially quantitative transfer of galactose onto complex glycoproteins expressed in the yeast *P. pastoris*. Both approaches are based on host cells that produce bi-antennary structures that terminate in GlcNAc^{56,57}. In one approach, we screened a library of GalT–leader fusions and did not find complete transfer of galactose in any of the tested constructs. We reasoned that transport of UDP-galactose into the Golgi and/or the intracellular pool of UDP-galactose in the cytosol were suboptimal, although the existence of a UDP-galactose transporter in *S. cerevisiae* had been described by Roy and co-workers⁶⁰. To overcome these limitations, we isolated a previously uncharacterized gene from *S. pombe* that had high amino acid sequence homology to known UDP-galactose-4-epimerases, enzymes that are conserved among galactose-assimilating organisms, including bacteria and mammals, and are responsible for the reversible conversion of UDP-glucose and UDP-galactose. By co-expressing this gene as a cytosolic enzyme, we found a marked improvement in the amount of galactose transfer, yet failed to demonstrate complete galactosylation. Finally, by co-expressing GalT, UDP-galactose-4-epimerase from *S. pombe* and a UDP-galactose transporter from *D. melanogaster*, we were able to demonstrate that terminally galactosylated *N*-glycans can be produced in yeast at high homogeneity (R. C. Davidson *et al.*, manuscript submitted). These results indicate that highly efficient galactosylation of complex glycoproteins in *P. pastoris* can be achieved through the metabolic engineering of a sufficient intracellular UDP-galactose pool, the expression of an active and properly localized GalT, and the translocation of UDP-galactose into the Golgi by an active UDP-galactose transporter.

In a second approach we tried to mitigate the risk associated with introducing a mammalian UDP-galactose transporter into the secretory pathway of yeast. As mentioned above, sugar nucleotide transporters have multiple transmembrane domains and so are not amenable to domain swapping to redirect localization to a different organelle. Accordingly, we created a fusion protein containing UDP-galactose-4-epimerase, the catalytic domain of GalT and a yeast leader sequence to anchor the fusion protein in the Golgi compartment of the yeast's secretory pathway. This synthetic fusion protein was able to mediate the conversion of UDP-glucose to UDP-galactose in the Golgi and catalyse the transfer of galactose to terminal GlcNAc to obtain complex bi-antennary glycans terminating in galactose⁵⁶.

Sialic acid transfer – the final step

After the transfer of galactose, human *N*-glycosylation continues with the terminal capping of *N*-glycans with sialic acid. Three criteria have to be met to allow for the production of sialylated glycoproteins in yeast: the availability of a terminal-galactose-containing *N*-glycan (see above), an active sialyltransferase and a pool of CMP-sialic acid. Several reports describe the recombinant production of secreted active sialyltransferases in yeast^{61,62}. We have recently shown that an active sialyltransferase can be localized to the late Golgi, where the enzyme is able to catalyse sialic acid transfer *in vivo* in the presence of an appropriate CMP-NANA transporter and a pathway to produce CMP-NANA in the cytosol (S.W. and T.U.G., unpublished observations).

However, these results are preliminary and need to be further improved before quantitative *in vivo* sialylation can be demonstrated.

Summary

Over the past decade our ability to manipulate *N*-glycosylation pathways in yeast and other protein-expression hosts has developed rapidly. In the past two years alone, we have seen the development of yeast strains that perform hybrid *N*-glycosylation at high uniformity⁵², the production of complex glycans⁵⁷ and, more recently, the production of terminally galactosylated complex glycans. Given this progress, it is expected that yeast strains with the ability to generate fully sialylated proteins will be attainable in the near future. This may obviate the need for mammalian cell culture for a number of protein-expression needs. Unlike mammalian cell culture, which typically results in recombinant products that have a heterogeneous distribution of glycans, engineered yeast cells have been evolved to perform essentially uniform *N*-glycosylation. As it has been shown that *N*-glycosylation can markedly impact on the pharmacokinetic and pharmacodynamic properties of glycoproteins, as well as their tissue distribution and receptor interactions, it is evident that individual glycoforms have different properties. Yeast strains that are able to perform homogeneous *N*-glycosylation are expected to allow the exploration of *N*-glycosylation-related structure-activity relationships, and ultimately might lead to the generation of better protein-based therapeutics.

- Nasmyth, K. Eukaryotic gene cloning and expression in yeast. *Nature* **274**, 741–743 (1978).
- Sudbery, P. E. The expression of recombinant proteins in yeasts. *Curr. Opin. Biotechnol.* **7**, 517–524 (1996).
- Gellissen, G. & Hollenberg, C. P. Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* – a review. *Gene* **190**, 87–97 (1997).
- Cregg, J. M., Vedvick, T. S. & Raschke, W. C. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology* **11**, 905–910 (1993).
- Cregg, J. M. & Higgins, D. R. Production of foreign proteins in the yeast *Pichia pastoris*. *Can. J. Bot.* **73**, S891–S897 (1995).
- Gellissen, G. & Melber, K. Methylotrophic yeast *Hansenula polymorpha* as production organism for recombinant pharmaceuticals. *Arzneimittelforschung* **46**, 943–948 (1996).
- Werten, M. W. T., Van den Bosch, T. J., Wind, R. D., Mooibroek, H. & De Wolf, F. A. High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **15**, 1087–1096 (1999).
- Young, N. M. *et al.* Structure of the *N*-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* **277**, 42530–42539 (2002).
- Szymanski, C. M., Yao, R. J., Ewing, C. P., Trust, T. J. & Guerny, P. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **32**, 1022–1030 (1999).
- Wacker, M. *et al.* *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298**, 1790–1793 (2002).
This is the first paper to describe transfer of a prokaryotic N-glycosylation machinery from C. jejuni to E. coli.
- Thibault, P. *et al.* Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **276**, 34862–34870 (2001).
- McAleer, W. J. *et al.* Human hepatitis B vaccine from recombinant yeast. *Nature* **307**, 178–180 (1984).
- Walsh, G. Biopharmaceutical benchmarks — 2003. *Nature Biotechnol.* **21**, 1396 (2003).
- Humphreys, A. & Boersig, C. Cholesterol drugs dominate. *MedAdNews* **22**, 42–57 (2003).
- Yuen, C. T., Carr, S. A. & Feizi, T. The spectrum of *N*-linked oligosaccharide structures detected by enzymatic microsequencing on a recombinant soluble CD4 glycoprotein from Chinese hamster ovary cells. *Eur. J. Biochem.* **192**, 523–528 (1990).
- Ulloa-Aguirre, A., Timossi, C., Barrios-de-Tomasi, J., Maldonado, A. & Nayudu, P. Impact of carbohydrate heterogeneity in function of follicle-stimulating hormone: studies derived from *in vitro* and *in vivo* models. *Biol. Reprod.* **69**, 379–389 (2003).
- Shields, R. L. *et al.* Lack of fucose on human IgG1 *N*-linked oligosaccharide improves binding to human FcγRIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277**, 26733–26740 (2002).
This publication highlights the importance of glycosylation for antibody effector function.
- Gentzsch, M. & Tanner, W. Protein *O*-glycosylation in yeast: protein-specific mannosyltransferases. *Glycobiology* **7**, 481–486 (1997).
- Lehle, L. & Bause, E. Primary structural requirements for *N*-glycosylation and *O*-glycosylation of yeast mannoproteins. *Biochim. Biophys. Acta* **799**, 246–251 (1984).
- Strahl-Bolsinger, S., Gentzsch, M. & Tanner, W. Protein *O*-mannosylation. *Biochim. Biophys. Acta* **1426**, 297–307 (1999).
- Ziegler, F. D., Gemmill, T. R. & Trimble, R. B. Glycoprotein synthesis in yeast early events in *N*-linked oligosaccharide processing in *Schizosaccharomyces pombe*. *J. Biol. Chem.* **269**, 12527–12535 (1994).
- Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–664 (1985).
- Burda, P. & Aebi, M. The dolichol pathway of *N*-linked glycosylation. *Biochim. Biophys. Acta* **1426**, 239–257 (1999).
- Gemmill, T. R. & Trimble, R. B. Overview of *N*- and *O*-linked oligosaccharide structures found in various yeast species. *Biochim. Biophys. Acta* **1426**, 227–237 (1999).
- Tabas, I. & Kornfeld, S. Purification and characterization of a rat-liver golgi α -mannosidase capable of processing asparagine-linked oligosaccharides. *J. Biol. Chem.* **254**, 1655–1663 (1979).
- Herscovics, A. Structure and function of class I α 1,2-mannosidases involved in glycoprotein synthesis and endoplasmic reticulum quality control. *Biochimie* **83**, 757–762 (2001).
- Herscovics, A. Importance of glycosidases in mammalian glycoprotein biosynthesis. *Biochim. Biophys. Acta* **1473**, 96–107 (1999).
- Tulsiani, D. R. P., Hubbard, S. C., Robbins, P. W. & Touster, O. α -*D*-mannosidases of rat-liver golgi membranes. Mannosidase-II is the GlcNAcMan₅-cleaving enzyme in glycoprotein-biosynthesis and mannosidase-Ia and Mannosidase-Ib are the enzymes converting Man₅ precursors to Man₆ intermediates. *J. Biol. Chem.* **257**, 3660–3668 (1982).
- Moremen, K. W. & Robbins, P. W. Isolation, characterization, and expression of cDNAs encoding murine α -mannosidase-II, a golgi enzyme that controls conversion of high mannose to complex *N*-glycans. *J. Cell Biol.* **115**, 1521–1534 (1991).
- Bendiak, B. & Schachter, H. Control of glycoprotein-synthesis. Kinetic mechanism, substrate-specificity, and inhibition characteristics of UDP-*N*-acetylglucosamine- α -*D*-mannoside β -1-2-*N*-acetylglucosaminyltransferase-II from rat liver. *J. Biol. Chem.* **262**, 5784–5790 (1987).
- Schachter, H. The joys of HexNAc. The synthesis and function of *N*- and *O*-glycan branches. *Glycoconjugate J.* **17**, 465–483 (2000).
- Nakanishi-Shindo, Y., Nakayama, K., Tanaka, A., Toda, Y. & Jigami, Y. Structure of the *N*-linked oligosaccharides that show the complete loss of α -1,6-polymannose outer chain from Och1, Och1 Mnn1, and Och1 Mnn1 Alg3 mutants of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **268**, 26338–26345 (1993).
This paper is the first successful attempt to eliminate yeast-specific N-glycan processing and to provide a human intermediate of N-glycan biosynthesis.

33. Chiba, Y. *et al.* Production of human compatible high mannose-type (Man₅GlcNAc₂) sugar chains in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 26298–26304 (1998).
This paper describes the first successful attempt to target mannosidases to the secretory pathway of yeast and demonstrate some effect on N-linked glycan processing.
34. Kalsner, I., Hintz, W., Reid, L. S. & Schachter, H. Insertion into *Aspergillus nidulans* of functional UDP-GlcNAc- α -3- α -mannoside β -1,2-*N*-acetylglucosaminyltransferase-I, the enzyme catalyzing the first committed step from oligomannose to hybrid and complex *N*-glycans. *Glycoconjugate J.* **12**, 360–370 (1995).
35. Martinet, W., Maras, M., Saelens, X., Jou, W. M. & Contreras, R. Modification of the protein glycosylation pathway in the methylotrophic yeast *Pichia pastoris*. *Biotechnol. Lett.* **20**, 1171–1177 (1998).
36. van Vliet, C., Thomas, E. C., Merino-Trigo, A., Teasdale, R. D. & Gleeson, P. A. Intracellular sorting and transport of proteins. *Prog. Biophys. Mol. Biol.* **83**, 1–45 (2003).
37. Gleeson, P. A. Targeting of proteins to the Golgi apparatus. *Histochem. Cell Biol.* **109**, 517–532 (1998).
38. Harter, C. & Wieland, F. The secretory pathway: mechanisms of protein sorting and transport. *Biochim. Biophys. Acta* **1286**, 75–93 (1996).
39. Teasdale, R. D. & Jackson, M. R. Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annu. Rev. Cell Dev. Biol.* **12**, 27–54 (1996).
40. Pelham, H. R. B., Hardwick, K. G. & Lewis, M. J. Sorting of soluble ER proteins in yeast. *EMBO J.* **7**, 1757–1762 (1988).
Excellent early paper describing the sorting of proteins in the secretory pathway in the yeast *Saccharomyces cerevisiae*.
41. Munro, S. & Pelham, H. R. B. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**, 899–907 (1987).
42. Dean, N. & Pelham, H. R. B. Recycling of proteins from the Golgi compartment to the ER in yeast. *J. Cell Biol.* **111**, 369–377 (1990).
43. Lewis, M. J., Sweet, D. J. & Pelham, H. R. B. The *ERD2* gene determines the specificity of the luminal ER protein retention system. *Cell* **61**, 1359–1363 (1990).
44. Semenza, J. C., Hardwick, K. G., Dean, N. & Pelham, H. R. B. *ERD2*, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* **61**, 1349–1357 (1990).
45. Paulson, J. C. & Colley, K. J. Glycosyltransferases — structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.* **264**, 17615–17618 (1989).
46. Krezdorn, C. H. *et al.* Human β -1,4 galactosyltransferase and α -2,6 sialyltransferase expressed in *Saccharomyces cerevisiae* are retained as active enzymes in the endoplasmic reticulum. *Eur. J. Biochem.* **220**, 809–817 (1994).
47. Schwientek, T., Lorenz, C. & Ernst, J. F. Golgi localization in yeast is mediated by the membrane anchor region of rat-liver sialyltransferase. *J. Biol. Chem.* **270**, 5483–5489 (1995).
48. Yoshida, S. *et al.* Expression and characterization of rat UDP-*N*-acetylglucosamine: α -3- α -mannoside β -1,2-*N*-acetylglucosaminyltransferase I in *Saccharomyces cerevisiae*. *Glycobiology* **9**, 53–58 (1999).
49. Nilsson, T. *et al.* Kin recognition between medial Golgi enzymes in HeLa cells. *EMBO J.* **13**, 562–574 (1994).
50. Callewaert, N. *et al.* Use of HDEL-tagged *Trichoderma reesei* mannosyl oligosaccharide 1,2- α - α -mannosidase for *N*-glycan engineering in *Pichia pastoris*. *FEBS Lett.* **503**, 173–178 (2001).
51. Bretthauer, R. K. Genetic engineering of *Pichia pastoris* to humanize *N*-glycosylation of proteins. *Trends Biotechnol.* **21**, 459–462 (2003).
52. Choi, B. K. *et al.* Use of combinatorial genetic libraries to humanize *N*-linked glycosylation in the yeast *Pichia pastoris*. *Proc. Natl Acad Sci. USA* **100**, 5022–5027 (2003).
53. Sharma, C. B., Knauer, R. & Lehle, L. Biosynthesis of lipid-linked oligosaccharides in yeast: the *ALG3* gene encodes the Dol-P-Man:Man₅GlcNAc₂-PP-Dol mannosyltransferase. *Biol. Chem.* **382**, 321–328 (2001).
54. Aebi, M., Gassenhuber, J., Domdey, H. & Heesen, S. T. Cloning and characterization of the *ALG3* gene of *Saccharomyces cerevisiae*. *Glycobiology* **6**, 439–444 (1996).
55. Davidson, R. C. *et al.* Functional analysis of the *ALG3* gene encoding the Dol-P-Man:Man₅GlcNAc₂-PP-Dol mannosyltransferase enzyme of *P. pastoris*. *Glycobiology* **14**, 399–407 (2004).
56. Bobrowicz, P. *et al.* Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris* — production of complex humanized glycoproteins with terminal galactose. *Glycobiology* **14**, 757–766 (2004).
57. Hamilton, S. R. *et al.* Production of complex human glycoproteins in yeast. *Science* **301**, 1244–1246 (2003).
References 52, 56 and 57 describe the use of combinatorial genetic libraries to obtain uniform human-like N-linked glycosylation in the yeast *P. pastoris*.
58. Schwientek, T., Narimatsu, H. & Ernst, J. F. Golgi localization and *in vivo* activity of a mammalian glycosyltransferase (human β -1,4-galactosyltransferase) in yeast. *J. Biol. Chem.* **271**, 3398–3405 (1996).
- This is the first paper describing the targeting of a mammalian glycosylation enzyme to the secretory pathway of a yeast.**
59. Vervecken, W. *et al.* *In vivo* synthesis of mammalian-like, hybrid-type *N*-glycans in *Pichia pastoris*. *Appl. Environ. Microbiol.* **70**, 2639–2646 (2004).
60. Roy, S. K., Yoko-o, T., Ikenaga, H. & Jigami, Y. 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 (1998).
61. Malissard, M., Zeng, S. & Berger, E. G. Expression of functional soluble forms of human β -1,4-galactosyltransferase I, α -2,6-sialyltransferase, and α -1,3-fucosyltransferase VI in the methylotrophic yeast *Pichia pastoris*. *Biochem. Biophys. Res. Commun.* **267**, 169–173 (2000).
62. Malissard, M., Zeng, S. & Berger, E. G. The yeast expression system for recombinant glycosyltransferases. *Glycoconjugate J.* **16**, 125–139 (1999).

Acknowledgements

The authors are grateful to the outstanding scientists that provided the basic knowledge upon which the engineering of humanized yeast strains was made possible. In particular, we wish to express our gratitude to P. Robbins, R. Bretthauer, K. Moremen, C. Barlowe, M. Aebi and R. Hitzman for helpful discussions and support. Most importantly, we are grateful to the exceptional group of scientists we have been able to work with over the past four years. It is their work that has advanced the field in such a significant way and may lead to a paradigmatic change in the production of complex protein-based therapeutics.

Competing interests statement

The authors declare no competing financial interests.

 **Online links**

DATABASES

The following terms in this article are linked online to:

Entrez: <http://www.ncbi.nlm.nih.gov/Entrez/ACX1> | *Campylobacter jejuni* | Erd2p | *Escherichia coli* | *Kluyveromyces lactis* | *mnn1* | *Saccharomyces cerevisiae*
SwissProt: <http://www.expasy.org/sprot/>
 Alg3p | Anp1p | Gls1p | Hoc1p | Mnn9p | Mnn10p | Mnn11p | Mns1p | Och1p | Sec12p | Van1p

FURTHER INFORMATION

Tillman Gerngross's laboratory:
<http://engineering.dartmouth.edu/thayer/faculty/tillmangerngross.html>

Access to this links box is available online.