Glycosylation and protein transport

Peter Scheiffele*1 and Joachim Füllkrug†

*Department of Biochemistry and Biophysics, Howard Hughes Medical Institute, University of California, San Francisco, CA 94143-0452, U.S.A., †Max-Planck Institute for Molecular Cell Biology and Genetics, D-01307 Dresden, Germany

Introduction

Glycans have a multitude of functions in organisms. They can be found conjugated to proteins, most commonly via a glycosidic linkage between the carbohydrate and the side chains of asparagine (N-glycans) or serine/threonine residues (O-glycans). As large numbers of different monosaccharides are incorporated into linear or branched structures, an enormously complex repertoire of carbohydrates can be created with far more permutations as observed for nucleic acids or polypeptides.

Why does the cell create this overwhelming diversity of structures? Studies performed during the last 20 years have revealed two principally different roles for protein-bound glycans: specific carbohydrate epitopes can serve as ligands for receptors that mediate recognition events, as adhesion between select cells. On the other hand, rather general glycan structures can be employed to change such biophysical properties of a protein as its charge, solubility, folding or sensitivity towards proteases [1]. Inhibition or deletion of glycosyltransferases normally does not influence growth of isolated cells, but embryogenesis is severely impaired. This highlights the importance of glycan structures in the multicellular organism.

1To whom correspondence should be addressed (e-mail: Scheiffe@uclink4.berkeley.edu).
Intracellularly, a specific carbohydrate epitope on lysosomal enzymes, the mannose 6-phosphate modification, has been shown to act as a signal for their delivery to lysosomes [2]. Besides this well-characterized sorting determinant, carbohydrates were thought to be primarily of structural importance for the folding of glycoproteins. However, recent studies suggested that carbohydrates can also provide information for other sorting events in the secretory pathway and the discovery of a new class of intracellular lectins has made this hypothesis even more attractive.

In this chapter we will discuss the role of glycans as sorting signals in the secretory pathway, specifically for exit of cargo molecules from the endoplasmic reticulum (ER) and the Golgi complex.

**Transport to the cell surface by default?**

In the ER, proteins with consensus sequences for N-glycosylation all acquire an identical large carbohydrate structure that is transferred ‘en block’ to the amido group of an asparagine. While passing through the secretory pathway this structure is then trimmed extensively and modified by glycosidases and glycosyltransferases to yield the large variety of N-glycans found in different glycoproteins. The carbohydrate-modification enzymes display a polarized localization throughout the Golgi stack, with the early acting enzymes being active in the cis and medial cisternae, and the late-acting enzymes in the trans-Golgi and trans-Golgi network.

According to the ‘bulk-flow’ hypothesis, proteins are transported passively on their route to the plasma membrane of mammalian cells, and early experiments suggested that carbohydrate modifications are not required for transport along the secretory pathway [3]. However, more recently elegant work on the calnexin/calreticulin quality-control machinery has demonstrated that the sequential processing of carbohydrates in the ER co-ordinates folding and exit of newly synthesized glycoproteins from the compartment [4]. Furthermore, evidence has accumulated indicating that efficient transport along the secretory pathway does require specific signals. Instead of being transported passively to the cell surface, cargo molecules were shown to be concentrated actively in the transport carriers [5]. The identification of new lectins and the finding that carbohydrates can act as intracellular sorting signals suggests that carbohydrates might play an important role in cargo selection in the secretory pathway (see Figure 1).

**Lectins in the secretory pathway**

Lectins are proteins that selectively bind to specific carbohydrate structures. Recently, a novel class of animal lectins, homologous to leguminous-plant lectins, has been identified. Although the overall homology of this protein family to legume lectins is low, the amino acids constituting the carbohydrate-binding site are conserved [6,7].
ERGIC-53 (ER-Golgi intermediate-compartment protein of 53 kDa) is the most extensively characterized member of this new class of lectins. It binds specifically to mannose-containing carbohydrate structures in a calcium-dependent manner. This interaction is abolished by mutation of amino acid residues in the sugar-binding site, as predicted by sequence comparison with leguminous lectins [8].

Human individuals deficient in ERGIC-53 suffer from a bleeding disorder, a rare autosomal recessive disease referred to as combined deficiency of coagulation factors V and VIII [9]. Both coagulation factors are heavily glycosylated and are reduced to 5–30% of the normal levels in plasma. Other than that, these patients do not show any other obvious symptoms, implying that embryonic development and secretion of other glycoproteins are not seriously affected. This is somewhat surprising since ERGIC-53 is an abundant protein and widely expressed in different tissues, suggesting a housekeeping function. However, it seems likely that ERGIC-53 enhances the efficiency of secretion in general, and other functionally redundant factors might be able to compensate for the lack of ERGIC-53 during development.

ERGIC-53 has been localized to the ER, ER-Golgi intermediate compartment and cis-Golgi in mammalian tissue-culture cells. It is an integral membrane protein that spans the membrane once, with its N-terminus localized to the lumen of the membrane compartment (a type-I membrane protein). A short cytoplasmic tail contains several determinants for sorting in the early secretory pathway. A double phenylalanine motif at the C-terminus has been shown to interact with cytoplasmic coat proteins (COP II), which are involved in budding of transport vesicles from the ER. These coat proteins select cargo

---

**Figure 1. Involvement of lectins or carbohydrates in different steps of the secretory pathway**

See text for details. M6PR, mannose-6-phosphate receptor.
molecules (see Chapter 4 in this volume by Francis Barr) and an attractive hypothesis suggests that a certain class of membrane proteins would interact with COP II on the cytoplasmic side and with secretory cargo proteins in the lumen of the ER. Such putative coat-cargo receptors would thus provide a bridge through the membrane bilayer, extending the cargo-selection capabilities of the COP II coat. The cytoplasmic tail of ERGIC-53 also contains a dilyssine ER localization motif directly interacting with coatomer (COP I) [8]. COP I proteins are thought to bind to the cytoplasmic tail of proteins in the Golgi apparatus and guide their retrieval to the ER. The combination of anterograde (COP II) and retrograde (COP I) signals is assumed to confer cycling of ERGIC-53 between ER and cis-Golgi, leading to a steady-state localization in the ER-Golgi intermediate compartment.

In view of all these properties, ERGIC-53 has been suggested to constitute a cargo receptor which would concentrate glycoproteins at ER exit sites. Following transport to the Golgi apparatus, changes in calcium concentration and/or pH would release the glycoproteins, and enable ERGIC-53 to recycle back to the ER to collect a new set of glycosylated secretory proteins. Alternatively, ERGIC-53 could be involved in quality control in the early secretory pathway. Incorrectly glycosylated proteins would be either retained or retrieved until they were modified properly. Interestingly, extensive trimming of mannose residues occurs at the ER–Golgi interface, possibly modifying the mannose-containing epitope recognized by ERGIC-53.

Expression of a mutant of ERGIC-53 that localizes exclusively to the ER (but retains lectin activity) results in impaired secretion of at least one lysosomal enzyme. However, other glycoproteins tested were not affected. Therefore, not all secretory proteins depend on the recycling of ERGIC-53 for efficient secretion. Still, these experiments prove the point that cycling of this protein is functionally relevant [8].

In mammalian cells, another abundant type-I membrane protein related to ERGIC-53 has been described. First identified in epithelial cells, VIP36 (vesicular integral membrane protein of 36 kDa) also displays significant sequence homology to leguminous-plant lectins. Similar to ERGIC-53, VIP36 recognizes a high-mannose sugar structure on glycoproteins and localizes to the early secretory pathway [10,11]. Future experiments will be required to clarify whether VIP36 and ERGIC-53 have common ligands and whether they are functionally redundant or complementary.

**Glycosylation-dependent sorting of proteins in the secretory pathway**

Studies performed in 1985 suggested that N-glycosylation might be required for transport of transmembrane proteins to the cell surface. Rose and co-workers [12] observed that a non-glycosylated membrane-anchored secretory protein (rat growth hormone) was efficiently transported to the Golgi
complex, but did not reach the cell surface. Introduction of N-glycosylation sites by site-directed mutagenesis allowed exit from the Golgi and subsequent transport to the cell surface. Similarly, surface transport of the vesicular stomatitis virus glycoprotein (VSVG) required the presence of at least one N-linked carbohydrate chain. However, N-glycosylation is not required for exocytosis of all proteins.

Some new insight was gained from the analysis of protein sorting in polarized epithelial cells (see Figure 2). Apical sorting of the major secretory protein in Madin–Darby canine kidney (MDCK) cells, gp80, was lost when cells were treated with tunicamycin, an inhibitor of N-glycosylation [13]. Furthermore, addition of N-glycosylation sites into the normally non-glycosylated and non-polarized secreted rat growth hormone was sufficient to target the protein apically. Interestingly, the number of N-glycosylation sites introduced correlated with the efficiency of apical sorting [14].

Figure 2. Sorting in the trans-Golgi network of polarized epithelial cells
In MDCK cells N-glycans can act as apical sorting signals for secretory and transmembrane proteins. For the p75NTR the O-glycosylated stalk is required for apical sorting and analysis of secreted proteoglycans revealed that chondroitin sulphate is apically secreted. Basolateral signals are found in the cytoplasmic tail of transmembrane proteins and can override an apical signal. Secretory proteins without signals are secreted in a non-polarized fashion and transmembrane proteins without signals are retained in the Golgi complex. TGN, trans-Golgi network; PG, proteoglycan.
blasts. In all cases the unglycosylated forms of the analysed proteins accumulated in the Golgi complex, whereas N-glycosylation allowed surface delivery. In MDCK cells the same glycosylated proteins were sorted to the apical membrane, suggesting that also for membrane proteins N-glycans can act as an apical signal [16]. Basolateral sorting signals are generally found in the cytoplasmic tail of membrane proteins. When such a cytoplasmic signal was added to any of the three model proteins they were targeted basolaterally, independently of N-glycosylation. Importantly, also in fibroblasts addition of a basolateral signal allowed carbohydrate-independent surface delivery. The simplest explanation for these results is that exit from the Golgi complex is signal-mediated. According to this hypothesis, transport of membrane proteins to the cell surface requires sorting determinants to direct cargo into either the apical or the basolateral (cognate) carriers. As suggested by previous studies the cytoplasmic basolateral signals are generally dominant over the apical information [17]. However, in case a protein carries no basolateral signal then positive apical information (as N-glycosylation) is required for exit from the Golgi complex.

Clearly, N-glycans can act as determinants of apical sorting; however, the presence of N-glycans in a molecule does not guarantee that the protein will be transported apically. Deletion of the basolateral signal in the cytoplasmic tail of the transferrin receptor results in non-polarized delivery, although the protein contains two N-linked glycans. Most likely, the positioning of carbohydrates on the surface of the molecule is critical, but also additional sorting information in the transmembrane domain can be required for apical delivery [18]. In the case of the neurotrophin receptor (p75NTR) deletion of the single N-glycosylation site did not abolish apical delivery, instead the juxtamembrane domain, which is modified by O-glycosylation, was required for sorting [19]. Direct evidence for a role of O-glycosylation in apical targeting is still lacking, although there are other examples of O-glycosylated proteins that are apically sorted in the absence of N-glycans.

Besides N- and O-glycans, secretory proteins can also be modified by addition of proteoglycan chains. Interestingly, a differential sorting of secreted proteoglycans in polarized cells was observed recently. Whereas only 20% of heparan sulphate proteoglycans are secreted apically, most of the chondroitin sulphate proteoglycans (75%) are secreted into the apical medium [20]. Heparan sulphate and chondroitin sulphate glycosaminoglycan chains differ mainly in their repeating disaccharides. Since protein-free chondroitin sulphate chains are also secreted predominantly to the apical medium, the sorting information must be localized in the sugar chains. However, the precise nature of this sorting signal remains to be elucidated.

How is sorting information in glycans decoded?

While there is now strong evidence for carbohydrate signals in the secretory pathway, it is far less clear how all these different sorting determinants are
interpreted. The mannose 6-phosphate epitope in the lysosomal enzymes is recognized by its receptor which has lectin activity. Are there more intracellular lectins recognizing sugar structures in N-glycans, O-glycans and proteoglycans in the apical post-Golgi route? Another possibility would be that instead of acting as specific epitopes the glycans could change general properties of proteins and for example drive self-association. Clustering of cargo molecules into targeting patches with specific lipid composition, so-called rafts, has been proposed to mediate apical sorting [21]. This clustering could be driven by lectins associated with raft domains, which display low-affinity interactions with carbohydrates of the cargo proteins. However, another possibility is that the glycan structures on the cargo molecules themselves mediate co-aggregation under the chemical conditions of the trans-Golgi network, similar to that which has been proposed for sorting of proteins into secretory granules for regulated secretion [22].

A requirement for specific terminal carbohydrate modifications in apical sorting has been tested by treatment of tissue-culture cells with pharmacological inhibitors of terminal glycosylation. In MDCK cells deoxymannojirimycin or swainsonine, which both inhibit the formation of complex-type N-glycans, did not disturb apical sorting of gp80 [23]. In HT-29 cells, however, inhibition of Galβ1-3GalNAc:2,3-sialyltransferase (and thereby terminal sialylation) resulted in an intracellular accumulation of several apical proteins whereas basolateral delivery was not affected [24]. It is possible that in the mucus-producing HT-29 cells signals additional to those found in MDCK cells operate. However, clearly further studies and the identification of apical sorting receptors will be required to define precisely the mechanism of apical sorting directed by carbohydrate signals.

Perspectives

Only a few years ago, the role of glycosylation in the secretory pathway was thought to be limited to sorting of lysosomal proteins by mannose-6-phosphate receptors. Since then, surprising and exciting progress has unveiled a role for glycosylation in ER quality control and generation of epithelial polarity at the level of the trans-Golgi network. In addition, a new class of lectins localized to the Golgi apparatus has been discovered, which still waits for a firm functional assignment.

Nevertheless, these intracellular leguminous lectin homologues have been one of the pieces of evidence which led to the general idea that protein transport in the secretory pathway does not occur by default, but is signal-mediated. The human proteins ERGIC-53 and VIP36 would be engaged in lectin–glycoprotein interactions, concentrating and chaperoning glycoproteins on their route along the secretory pathway, or proof-reading their carbohydrate modifications. Sequence database searches reveal ERGIC-53 and VIP36
homologues in eukaryotic organisms as diverse as yeast, worms and fruitflies, suggesting evolutionarily conserved roles in the exocytic pathway.

The multitude of different carbohydrate signals operating in post-Golgi trafficking makes it entirely possible that whole families of specific intracellular lectins have not yet been identified. However, an alternative explanation is that those signals act rather by driving co-aggregation of cargo molecules or association with few lectin-like receptors of very low carbohydrate specificity. Although there are still many open questions it now seems likely that in every eukaryotic cell exocytosis requires signals that target cargo on specific post-Golgi routes to the cell surface.

Summary

- Transport along the secretory pathway is largely signal-mediated.
- Proteins in the secretory pathway can be covalently modified with various carbohydrate structures, most commonly O-glycans, N-glycans and/or proteoglycans.
- Carbohydrate modifications can change the physical properties of proteins or can function as specific recognition epitopes.
- Glycosylation can act as an apical sorting signal in polarized epithelial cells and provide a signal for surface transport in non-polarized fibroblasts.
- Homologues of leguminous plant lectins have been identified in yeast, fruitflies, worms and humans.
- Intracellular lectins are candidate receptors in the secretory pathway to mediate concentration of cargo in carrier vesicles.

We are grateful for financial support from the German research foundation (Deutsche Forschungsgemeinschaft), the Max-Planck Society and the European Molecular Biology Organization.

References