We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,600
Open access books available

137,000
International authors and editors

170M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 14

The Role of Glycans in Apical Sorting of Proteins in Polarized Epithelial Cells

Kristian Prydz, Gro Live Fagereng and Heidi Tveit

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/48214

1. Introduction

Epithelia consist of highly differentiated cells that form tight layers lining the inner cavities and the outer surface of the body. In classical cell biology, carbohydrate groups are attached to proteins (called glycoproteins or proteoglycans, depending on the carbohydrate structure) and lipids (glycolipids) - facing towards the exterior from the outer lipid leaflet of the cell membrane. These carbohydrate structures were previously thought to constitute a protective water-binding glycocalix, but were also known early on to be attachment sites for viruses, parasites, bacteria and bacterial toxins. Today, the glycan moieties of glycoproteins and proteoglycans are known to participate in protein sorting, transport and signaling processes in eukaryotic cells. Such processes are more complex in the epithelial cell, which possesses two distinct plasma membrane areas; the apical membrane domain, directed towards the inner cavities of the body, and the basolateral membrane domain facing the bloodstream. These two opposite membrane domains are segregated and differ in protein and lipid composition, and to maintain this polarity, vectorial transport of newly synthesized protein and lipid molecules is required. Proteins destined for the plasma membrane must traverse the secretory pathway. Such proteins emerge from ribosomes that dock onto endoplasmic reticulum (ER) membranes, due to recognition of a signal sequence in the protein itself. During co-translational import into the ER lumen, the starting point of the secretory pathway, the majority of translocated proteins receive carbohydrate modifications of the N-glycan type on asparagine residues, when a mannose-rich branched glycan structure with three terminal glucose residues is transferred en bloc from the lipid carrier dolichol to a proper modification site by an oligosaccharyl transferase. Such N-glycan structures are utilized as recognition units on glycoproteins by the protein folding quality control system which either allows further transport along the secretory pathway or diverts misfolded proteins to proteasomal degradation in the cytoplasm (1).
Properly folded proteins exit the ER in membrane carriers that are transported to the intermediate compartment (IC), and from there further to the Golgi apparatus. While N-glycans are modified to become more complex structures by enzymes in the Golgi apparatus, are other classes of glycans attached to alternative glycosylation sites in available protein cores. O-glycosylation on serine and threonine residues, starting with N-acetyl-galactosamine often results in mucin-type branched glycans, while glycosaminoglycan modification of serines with a neighboring glycine results in proteoglycans. Proteins and proteoglycans exiting from the trans side of the Golgi apparatus – the trans-Golgi network - are either transported in vesicles moving directly to the cell surface or to an endosomal compartment, from where the cell surface can be reached via indirect routes. Besides being the main organelle for carbohydrate synthesis and modification, the Golgi apparatus has been regarded as the main sorting center for polarized delivery of proteins in epithelial cells. In the trans-Golgi network, molecules are sorted and packaged into vesicles destined either for the apical or the basolateral cell surface domain. Although the trans-Golgi network has been established as a main protein sorting center (2), both pre- and post-Golgi polarized sorting has been reported for epithelial cells (3).

Molecular mechanisms governing polarized protein sorting in epithelial cells have been studied for to decades. Adapter proteins, tethering factors and fusion mediators have been assigned to different organelles and transport steps. Some proteins are shown to be required for particular routes although the molecular role played, is uncertain. Sorting signals have been localized to cytoplasmatic protein tails, transmembrane regions, specific globular protein domains, as well as post-translational modifications like lipid anchors and carbohydrate structures. The definition of a protein sorting signal, is that this molecular feature works after transfer to a previously non-sorted or oppositely sorted protein. In studies of glycans as sorting mediators, this requires manipulation and transfer of glycan modification sites. Such studies have indicated that both N- and O-glycans can mediate apical sorting, as well as long unbranched glycosaminoglycan chains of the chondroitin sulfate type. In this chapter we emphasize how different post-translationally added carbohydrate structures participate in polarized sorting of proteins. Although there are many examples that glycan structures and glycosylation sites in proteins are central to protein sorting, the underlying molecular mechanisms and their intracellular sites of action have not been thoroughly described.

2. Polarized epithelial cells

Single layers of epithelial cells coat the inner cavities of the body, like the intestinal tract, lungs, and kidney tubules. Attachment to the underlying tissues is mediated by proteins in the basal membrane domain that bind firmly to extracellular matrix components, while an apical region of the plasma membrane, with microvillus protrusions faces the lumenal environment. The apical membrane domain of the cell surface is segregated from the lateral and basal regions, together termed the basolateral domain, by tight junctions, protein complexes consisting of transmembrane anchored adhesion proteins, like claudins and occludins. The tight junctions fulfill several functions. Together with other junctional
complexes, they connect the individual cells in an epithelial layer by closing the gap between all neighboring cells in the monolayer. At the same time, the tight junctions introduce a restriction to lateral movement of proteins and lipids from the apical domain of the plasma membrane to the basolateral domain, and *vice versa* (4). In addition, tight junctions contribute to the fence function of epithelia, by restricting the intercellular route for passage of macromolecules. The tightness of epithelia is variable with the physiological functions of the tissue in question, some epithelia allow passage of small proteins, while some even restrict the passage of water molecules, like the epithelium of the distal kidney tubules where reabsorption of water molecules takes place (Figure 1).

**Figure 1. Polarized kidney epithelial cells.** A single kidney epithelium monolayer consists of polarized cells with two surface domains segregated by tight junctions. The apical domain facing the kidney tubule lumen, and the basolateral domain attached to the extracellular matrix (ECM), facing the bloodstream.

The apical membrane domain takes on specialized tissue functions, while the basolateral domain fulfills more general functions, like uptake of nutrients from the blood, building blocks like amino acids, lipids, and monosaccharides for synthesis of macromolecules. The functional differences of the two cell surface domains requires some differences in protein and lipid composition of these two opposite membrane areas, which again requires sorting mechanisms that direct newly synthesized proteins and lipids to the proper plasma membrane region. Receptors and transporters for binding of ligands like lipoproteins and growth factors and sugars, amino acids, sulfate and phosphate must be sorted basolaterally, together with attachment proteins, while several proteins involved in ion secretion to the body cavities must be sorted apically. A number of epithelia are immunologically active...
tissues that engage in uptake and transcytosis of immunologically active molecules, like immunoglobulins. Thus, it is of the utmost importance that the corresponding receptors are correctly sorted and transported.

2.1. General aspects of cell polarity

Most proteins destined for the cell surface in animal cells are modified with carbohydrate structures during their passage through the secretory pathway. Such carbohydrate structures are functionally important to the individual protein at the final cellular destination, but may also play a role in guiding the protein to its correct site of action. Thus, glycan modifications can contribute to the generation and maintenance of proper cellular organization in several ways, by positioning the protein correctly, and by fine-tuning the activity. Cell polarity is not restricted only to epithelial cells. Endothelial cells also possess tight junctions and apical and basolateral membrane domains similar to those observed in epithelia, and even in neurons there are restrictions to lateral protein mobility in the plasma membrane (5). In the hepatocyte, the apical membrane is divided into several separate domains that form bile canaliculi together with corresponding domains of neighboring hepatocytes. In these cells, the basolateral surface domain is continuous and constitutes the non-canonical (sinusoidal) regions of the plasma membrane, covering approximately 85% of the total cell surface area. Studies of protein sorting in hippocampal neurons have indicated that the apical membrane of epithelial cells and the hippocampal axons share some characteristics with respect to protein sorting (6, 7), suggesting also that dendritic and basolateral transport mechanisms share common features.

Much of our knowledge on polarized sorting and transport of proteins has been derived from studies with epithelial cell lines, such as Madin-Darby canine kidney (MDCK) cells and the colon cancer cell line CaCo². These, and other epithelial cell lines, have been cultured on filters with an appropriate pore size and allowed to form confluent monolayers with a measurable electrical resistance. MDCK cells can grow on almost any type of support, since these cells produce their own matrix components needed for subsequent attachment and differentiation. The relative tightness of epithelial monolayers may in many instances be determined from trans-epithelial resistance values that are easily measured when the epithelium has achieved a sufficient tightness. The ease by which kidney epithelial cells are grown on filters, and the tightness of the monolayers formed, grants differential experimental access to the apical and basolateral membrane domains and their corresponding medium reservoirs. This allows for both biochemical studies and various modes of microscopy. Apical and basolateral medium samples can be collected without mixing, and the apical and basolateral membrane domains are differentially accessible to biotinylation reagents, antibodies and reagents for chemical modification. Biochemical studies of polarity are not carried out with similar ease for neuronal cells and hepatocytes. Although hepatocytes sometimes are classified as epithelial cells, the apical membrane normally forms perpendicularly to the growth substratum in culture. Thus, polarity studies with hepatocytes are more rare than for other epithelial cell types. While some conclusions from studies carried out with the MDCK cell line might have transition value to hepatocytes
and neurons, there are also evident differences. In MDCK cells, proteins destined for the apical membrane are mostly delivered directly to this membrane domain, without transient basolateral appearance. In hepatocytes, however, most newly synthesized apical proteins are first transported to the basolateral membrane domain, followed by endocytosis and transcytosis across to the apical, bile canalicular membrane (8). Many polarized cell types possess both direct and indirect transport routes for newly synthesized proteins to the cell surface, observed both in epithelial cell lines like CaCo-2 (9) and in neurons (10).

3. The secretory pathway

Access to the secretory pathway is normally granted to polypeptides that emerge from ribosomes with an N-terminal signal sequence that is recognized by the signal recognition particle (SRP) in the cytoplasm followed by docking onto the ER membrane. However, signals elsewhere in a protein may also mediate uptake into the ER lumen, sometimes giving rise to a different protein topology. After ER entry, and during further transport through the secretory pathway, a large number of different types of post-translational modifications may occur. Proteolytic removal of the N-terminal signal sequence is a frequent event, while proteolysis at the C-terminus occurs more seldom, when a single trans-membrane domain is cleaved to transfer the remaining protein onto a glycosylphosphatidylinositol (GPI) lipid anchor. It has been estimated that approximately 250 proteins are prone to transfer onto GPI-anchors (11). The oxidative environment in the ER lumen promotes the stability of disulfide bridges, which are enzymatically formed by protein disulfide isomerase. An important aspect of protein assembly is the formation of multimeric proteins from single polypeptide subunits. Excess subunits often expose their more hydrophobic interaction surfaces and are thus recognized in the same way as misfolded proteins. A functional relationship between different classes of post-translational events, like formation of oligomers and glycosylation has been reported to influence apical sorting of some proteins (12-14), and will be discussed later.

The altogether most common modification of proteins is addition of N-glycans in the ER. A pre-assembled glycan moiety, rich in mannoses and with terminal glucose units is built onto the isoprene lipid dolichol and transferred to an appropriate asparagine-X-serine/threonine site by an oligosaccharyl transferase, as the polypeptide moves through the Sec 61 translocon into the ER lumen. Glycans added in this manner are used as handles for protein folding investigations carried out by chaperones, and the status of their terminal sugar residues also indicates when and where to depart from the ER (1, 15). The N-glycans acquired and subjected to initial trimming in the ER can function as mediators of forward transport in the early secretory pathway for several proteins. L-type lectins, like ERGIC-53, VIP-36 and VIP-like lectins bind immature N-glycan groups in a Ca²⁺ dependent manner (16) and mediate efficient transport of glycoproteins towards the intermediate compartment (IC) and cis-Golgi region in this way (17). In the Golgi apparatus, N-glycans are trimmed further by removal of more mannose units, before addition of novel sugar units takes place. A variable number of antennae receive N-acetyl-glucosamine, galactose, and sialic acid units, or only some of these sugars. The glycans may also receive fucose units and more rare
modifications, variable from tissue to tissue. N-glycans may appear at the cell surface as hybrid structures, where some antennae have maintained a mannose-rich branch, while others have been rebuilt to a more complex variant. Some glycoproteins arrive at the plasma membrane with only high-mannose branches in their N-glycans. In some cases, at least, this correlates to a Golgi-independent transport route to the cell surface (18), but a good correlation between N-glycan structures at the cell surface and the transport route followed to get there has not been established. From a standard high-mannose N-linked glycan structure in the ER, an almost unlimited number of carbohydrate structures may be generated after departure from the ER, during the subsequent passage through the Golgi apparatus. In fact, the glycan modification of a particular protein depends among other factors on the tissue where expression takes place and the developmental stage of the organism.

The Golgi apparatus is localized in a central, perinuclear position in vertebrate cells, where stacks of membrane-limited cisternae are joined together in a membrane ribbon. In other animals, like Drosophila melanogaster (19), and in plant cells, Golgi stacks are spread in the cytoplasm without ribbon-forming interconnections. Still, general mechanisms of protein exit from the ER and transport towards and through the Golgi apparatus are very similar. Glycan modification of proteins and lipids, and their further transport through the secretory pathway share similar principles across the plant and animal worlds, although the glycans acquired are structurally different. Significant attention has been given to the trans-most cisterna or region of the Golgi apparatus, the trans-Golgi network, as an important site for segregation of newly synthesized proteins on their way to apical and basolateral membrane domains of polarized epithelial cells (2). Early studies of epithelial polarity were performed with the MDCK cell line (20-23), where viral proteins, like influenza hemagglutinin (HA) and the vesicular stomatitis virus (VSV) G protein, were used as markers for apical and basolateral protein delivery, respectively. Initially, experiments were carried out with double infection of epithelial cells by viruses with opposite budding polarity. The polarized epithelial cells could be studied in a subsequent time window until the intracellular transport machinery had been shut down (20, 21). By electron microscopy, both influenza HA and VSV G protein were observed in the trans-Golgi network of the same cell, contributing, together with biochemical assays, to the conclusion that apical and basolateral proteins travel together through the Golgi apparatus and are only sorted and segregated upon exit from the trans-Golgi network (2). Later, as molecular biology techniques developed, epithelial cells were transfected to express the viral envelope proteins, including manipulated variants that could shed light on where in the proteins sorting information is localized (24-26). The early studies of polarized sorting of viral proteins did not indicate a role for glycans as apical sorting information, but rather implicated that sorting was mediated by the transmembrane and cytoplasmic regions of apically transported proteins like (27, 28, 25). A similar sorting outcome as that reported for epithelial MDCK cells was reported for hippocampal neurons, where influenza HA localized mainly to the axon, while the VSV G protein localized to the dendrites (6). The GPI-linked protein Thy-1 confirmed the expectation of being sorted exclusively to the axon (29), in line with the view that the axon resembles the apical surface of the epithelium, since a large number of GPI-linked proteins were observed to localize predominantly apically in MDCK cells (30).
4. Sorting sites in polarized cells

The trans-Golgi network is, as previously noted, not equally important as a sorting site of the secretory pathway in all polarized cell types, since hepatocytes deliver most apical proteins via the basolateral surface domain (9). Eventually, endosomes have been found to be transit compartments between the trans-Golgi network and the cell surface both in the secretory (31-33) and transcytotic pathways (34-37), also in epithelial MDCK cells. Thus, both direct and indirect routes towards the cell surface generally exist for apical and basolateral cargo transport, and numerous types of carriers have been implicated in post-Golgi transport (38, 39). Clearly, the trans-Golgi network is observed as an important sorting site for apical and basolateral cargo molecules, based on studies of fluorescently labeled model proteins in living cells (40). However, the observation of tubulo-saccular membrane structures emanating from the trans-Golgi network, carrying incompletely segregated apical and basolateral proteins, supports the existence of additional post-Golgi sorting stations (41, 42). The choice of model proteins for apical and basolateral cargo transport routes, evidently influences the observations made.

More recently, a number of studies have suggested that protein sorting in the secretory pathway is not exclusively a late-and post-Golgi affair. For instance, the trans-most Golgi cisterna is not the only one in a Golgi stack that displays clathrin coats, indicating the possibility of departure from the Golgi apparatus earlier than the trans-Golgi network. These coats may drive the formation of vesicles that are targeted for the endosomal/lysosomal pathway, but also for the basolateral membrane (43, 44). Other coat types have been observed even at the rim of the cisternae preceding the clathrin-coated cisternae (45). Such coats were shown to engage in transport towards the plasma membrane. Thus, several cisternae could constitute the exit interface of Golgi membranes (46). In mammalian cells, the intermediate compartment (IC), consisting of vacuolar and tubular membrane domains and functionally positioned between the ER and the Golgi apparatus, is also stably localized in a perinuclear position at the cell center, close to the Golgi cisternae. Although the trans region of the Golgi apparatus has been regarded as the primary sorting area for proteins and lipids in the secretory pathway, the IC has been suggested as a protein and lipid sorting station operating prior to entry into the Golgi apparatus (47-49). In fact, lateral segregation of proteins in the secretory pathway can occur already in the ER (50-53, 3, 18).

Where in the secretory pathway sorting takes place is an interesting question in relation to the role played by glycans, since the input into and output from the Golgi apparatus of glycoproteins, proteoglycans (PGs) and glycolipids is very different with respect to glycan structures. As far as we know today, the glycan structures of glycoproteins entering the Golgi apparatus are identical or very similar, while the action of Golgi glycosidases, glycosyltransferases, epimerases and sulfotransferases generate the astonishing diversity observed for glycan structures, even on individual proteins. This is important to consider when discussing possible targeting information in N-glycan structures, since the terminal sugars of N-glycans are likely to be too diverse to serve as general recognition units, while
the high-mannose structures existing early in the secretory pathway have proven to serve as secretory cargo ligands for protein lectins like ERGIC-53 and VIP-36 (54, 55).

In addition to N-linked glycosylation, another important protein glycosylation mechanism exists that is initiated with the addition of an N-acetyl-galactosamine (GalNAc) unit to a serine or threonine in an acceptor protein core early in the Golgi apparatus, or just before entry into the cis-Golgi lumen. Further extension by the action of several glycosyltransferases results in a branched O-glycan, sometimes sulfated. Some O-glycosylated secretory proteins are classified as mucins. These are heavily glycosylated and are mainly secreted apically from different epithelial tissues to constitute the mucus layer lining the body cavities. Mucins have attracted much attention because their glycan structures are variable from tissue to tissue and change during transformation of normal epithelial cells into cancer cells (56, 57).

The third of the major glycosylation mechanism in metazoans is the polymerization of long, linear glycosaminoglycan (GAG) chains onto protein acceptor cores, resulting in proteoglycans (PGs). The major classes of GAG chains that decorate PGs are chondroitin sulfate (CS)/dermatan sulfate (DS) and heparan sulfate (HS)/heparin. These polymers are all connected to a protein core via a linker tetrasaccharide consisting of xylose-galactose-galactose-glucuronic acid, extending from -serine-glycine-sites in the polypeptide, with the xylose attached to the serine. While the initiation of the linker region might take place before entry into the Golgi apparatus, GAG chains are shown to be enzymatically polymerized in the Golgi apparatus itself and subsequently modified to a variable extent by epimerization and sulfation. Keratan sulfate PGs are less abundant and extend from N- or O-linked glycans. In addition, hyaluronic acid is classified as a GAG, but is not attached to a protein core (58). Some PGs are defined as hybrids, since the protein core may carry more than one type of GAG. In addition GAGs may co-exist together with N-glycans and O-glycans on the same PG protein core.

5. Glycans and protein sorting

Early speculations on protein sorting in epithelial cells suggested that apical protein transport would most likely require sorting signals, since the apical surface is the specialized domain of epithelial plasma membranes. Basolateral transport could then occur by default, since this surface domain bears resemblance to non-polarized cells in which anterograde sorting was not discussed (59). Since a number of GPI-linked proteins had been detected at the apical surface of epithelial MDCK cells (30) such lipid anchors were proposed to possess apical sorting information. The subsequent discovery of basolateral sorting signals in the cytoplasmic domain of several transmembrane protein receptors (60-66) indicated that polarized basolateral transport also requires sorting. Later it was discovered that basolateral targeting does not have a monopoly on cytoplasmic protein domain sorting signals (67-69). The situation for GPI-linked proteins was later shown to be quite complex. The GPI-anchor alone is not sufficient to direct a protein to the apical surface. There is usually an additional or a sole requirement for N-glycans (12, 70) or also for additional oligomerization (71).
5.1. N-glycans and apical sorting

N-glycans were introduced as apical sorting signals in the 1990-ies. When polarized MDCK cells were transfected to express human erythropoietin, the glycoprotein that possesses three sites for N-glycosylation, was secreted from the apical membrane. Mutagenesis studies demonstrated, however, that only one of the N-glycosylation sites was critically important for apical sorting (72). In another study, the non-glycosylated protein rat growth hormone (rGH) was expressed in polarized MDCK cells and the secretion pattern was compared to that of rGH variants with one or two N-glycosylation sites. Addition of glycosylation sites stimulated apical secretion compared to the random secretion pattern observed for unglycosylated rGH (73). Such examples of apical targeting mediated by N-glycan groups lead to a search for sorting lectins that could mediate apical sorting at the required site in the Golgi apparatus, the trans-Golgi network. A promising candidate was VIP36, a protein lectin which was extracted in the detergent insoluble fraction, proposed to contain lipid raft associated proteins that are destined for the apical membrane (74). Later studies revealed that VIP36 in fact is localized more to the early secretory pathway (75), fitting with the observed binding affinity for high-mannose glycans (54). Still, manipulation of the expression level of VIP36 affected the apical sorting stringency of gp80 (or clusterin), the major endogenous glycoprotein secreted from MDCK cells. An early indication of the importance of N-glycans for apical secretion came from a study of this protein, which was secreted randomly from polarized MDCK cells in the presence of the glycosylation inhibitor tunicamycin (76). Enhanced expression of VIP36 in MDCK cells directed gp80 more strictly towards the apical medium (77). The fact that VIP36 and other lectins in the secretory pathway recognize high-mannose glycan structures (54, 16), indicates that their site of action is early in the pathway, or that their substrates are proteins that maintain a high mannose structure throughout the pathway (78). Thus, the role of VIP36 and related lectins in polarized sorting in epithelial cells requires more research to become fully understood.

Cholesterol and glycosphingolipid-rich lipid “rafts” have been proposed to function as sorting platforms in the TGN for vectorial delivery of proteins and lipids to the apical membrane of epithelial cells. The lipid raft concept describes regions of a membrane that are stable in time due to a higher content of cholesterol and more saturated fatty acids than surrounding membrane areas, with more freely moving lipids (79, 80). Since direct experimental contact with intracellular membranes is difficult, much of the evidence has been generated by indirect methods. Cholesterol depletion from cells by synthesis inhibition and extraction from cellular membranes has been shown to reduce apical transport efficiency of influenza HA and gp80 (81). A criterion for a particular protein to be regarded as raft associated has been an observable insolubility to treatment with cold detergent. In agreement with this view has cholesterol depletion been shown to both reduce apical transport efficiency and detergent insolubility (81). A reduction in apical transport is, however, not necessarily specifically interfering with apical sorting mechanisms, but could mediate a general reduction of apical transport capacity, since cholesterol depletion also reduces apical secretion of non-sorted proteins (82). A class of proteins proposed to require lipid rafts to reach the apical surface, is constituted by GPI-linked proteins (83, 84). However, a GPI anchor as such is not sufficient to drive the protein exclusively in the apical
Glycosylation direction. This has been shown in many cases to require the presence of N-glycans (12, 71, 70) and also oligomerization of the apically sorted protein (71, 13). An additional level of complexity was added, when it also was shown that the structure of the GPI-anchor itself influences whether the GPI-linked protein is transported apically or basolaterally (85, 86). Some GPI-linked proteins are transported basolaterally, presumably because their GPI-anchors do not favour oligomerization (85). The GPI-linked PG glypican carries both heparan sulfate (HS) chains and N-glycans, and is mainly transported basolaterally (87). However, when the sites for HS attachment are removed, glypican is mainly transported apically. Since the GPI-anchor and the N-glycans are the same in both cases, this would indicate that the presence of HS chains reduces the ability of glypican to oligomerize. Alltogether, the complexity of this field underlines one important fact: it is easier to study secretory proteins than proteins with membrane attachment.

Apically targeted proteins that are not associated with lipid rafts, meaning they are not associated with detergent-resistant microdomains, do not seem to require oligomerization to reach the apical surface (13). Still some clustering mechanism can be required. A group of small lectins with one or two carbohydrate binding domains are called galectins. Although galectins are synthesized in the cytoplasm and are devoid of known signals for membrane translocation, several members of this family of proteins have been shown to be of the outmost importance for epithelial cell differentiation and glycoprotein sorting mechanisms (88-92). The galectins are to a variable extent able to cross intracellular membranes or the plasma membrane to gain access to the lumen of intracellular compartments like endosomes, or to the outside of the cell, where (in both cases) the glycans of glycoproteins and glycolipids are localized. The translocation process has been shown to require glycan counter-receptors at the opposite side of the plasma membrane in the case of galectin-1 translocation (93). This might also be a requirement for other galectins and underscores that galectin translocation could be a highly regulated process, because the proper counter-receptors might be required. Galectin-3 has been shown to promote apical sorting of glycoproteins that are not raft-associated. Depletion of galectin-3 from MDCK cells resulted in missorting of non-raft glycoproteins, but did not affect the apical sorting of raft-lipid associated proteins (88). The effect on sorting seems to be that apical transport requires galectin-3 mediated glycoprotein clustering to be efficient (94). In polarized epithelial HT-29 cells, galectin-4 recruits N-linked glycoproteins to detergent resistant membrane fractions and further to the apical membrane (89). Knock-down of galectin-4 did not inhibit Golgi exit of glycoproteins, but reduced their surface appearance. These, and other indications point to endosomes as the site of the galectin-mediated glycoprotein clustering required for efficient apical transport. Galectin-9 has affinity for a class of glycosphingolipids at the apical surface of epithelial MDCK cells and is essential to the organisation of the apical membrane (90). Whether galectin-9 also contributes to apical transport of these glycolipids is not yet known.

5.2. O-glycans

O-glycans are branched and diverse, tissue-specific glycan structures initiated by the addition of an N-acetyl-galactosamine unit to a serine or threonine residue in an acceptor protein core (95). Although some bioinformatic tools have been developed for the purpose,
these sites are not easily predicted from analysis of the amino acid sequence. As described for N-linked glycans, O-linked sugars may also direct apical sorting in polarized MDCK cells, as demonstrated for the neurotrophin receptors, where the stalk region rich in O-glycans was required for apical transport (96). O-glycans are also important for apical sorting of the human intestinal sucrase-isomaltase in MDCK cells (97). The sorting domain could induce apical sorting after transfer to rat growth hormone (rGH) in MDCK cells (98). In some cases, both N-linked and O-linked sugars play essential roles in apical targeting of a protein, as shown for bovine enteropeptidase (99, 100). The heavily glycosylated transmembrane mucin glycoprotein MUC1 is efficiently delivered to the apical membrane of epithelial MDCK cells in a glycan-dependent manner (101). The apical targeting was not affected by galectin-3 knockdown, but was blocked when O-glycan synthesis was inhibited. Thus, the mechanistic basis for apical sorting of MUC1 is in all likelihood a glycan recognition event, but this remains to be characterized.

5.3. Glycosaminoglycans (GAGs) – The glycans decorating proteoglycans (PGs)

Proteoglycans are expressed in vertebrate cells, in *Drosophila melanogaster* (102-105), in *C. elegans* (106, 107) and in simple organisms as the starlet sea anemone *Nematostella vectensis* (108). Synthesis, sorting and transport of proteoglycans can therefore be studied in a number of different model organisms. When arriving at the cell surface, proteoglycans can participate in binding and uptake of signalling molecules, such as fibroblast growth factors. Proteoglycans at the cell surface may be endocytosed and pass via endosomes to lysosomes (109), but have also been observed in the nucleus (110-112), co-localising with growth factors. The route taken from endosomes to the nucleus is essentially unknown. In rat hepatocytes, heparan sulfate proteoglycans are transported from the trans-Golgi network to the cell surface in vesicles different from those that transports serum albumin, apolipoprotein E and fibrinogen (113, 114). Frequently, proteoglycans are sorted to the regulated secretory pathway, being released together with other contents of storage granules upon a proper stimulus (115-117). The negatively charged proteoglycans may bind small, positively charged molecules, such as histamine (118-120) and proteases (121, 122).

The definition of a proteoglycan is a protein core that is modified with one or several glycosaminoglycan (GAG) chains in the Golgi apparatus. Three different classes of protein-linked GAG chains are classified as heparan sulfate (HS)/heparin, chondroitin sulfate (CS)/dermatan sulfate (DS) or keratan sulfate (KS). The classification is based on the sugar units that constitute the repeating disaccharides in each type of the long, linear GAG chains (58). The most widespread GAGs are those of the CS/DS type and the HS/heparin type.

The synthesis of these GAG chains is initiated by the stepwise enzymatic addition of one xylose, two galactoses and a glucuronic acid, which makes up the linker region of the GAG chain. The fifth sugar added decides whether a GAG chain enters the HS/heparin pathway or the CS/DS track. The xylose is thought to be added to a serine with a neighboring glycine residue already in the ER by xylosyl transferase I or II (123, 124). The two galactoses are subsequently added by galactosyl transferase I and II, which are enzymes localized to the
cis-Golgi region (125). The addition of glucuronic acid takes place in the medial/trans Golgi (126). A protein core that carries several GAG chain binding sites can be modified with both CS and HS chains and would then be called a hybrid proteoglycan (127, 128). The polymerization is catalyzed by enzyme complexes adding disaccharides of hexosamine and hexuronic acid nature. In the case of CS and DS, the hexosamine is N-acetyl-galactosamine, while in the case of HS and heparin it is N-acetyl-glucosamine. The hexuronic acid added is in all cases glucuronic acid, but this may subsequently be epimerized to iduronic acid in DS, HS and heparin by specific epimerases. All these classes of GAG chains are modified by sulfotransferases in various positions on the disaccharides and to a variable extent (58). All in all, although linear structures, the GAG chains may contain a variety of negatively charged domains that have affinity for proteins with clusters of positively charged amino acids.

Epithelial cells transport proteoglycan components of the extracellular matrix (ECM) and receptor proteins that attach the basolateral membrane domain of the epithelium to the ECM by vectorial basolateral routes. Basolateral localization of syndecan-1 in MDCK cells requires a signal positioned in the 12 most distal amino acids of its cytoplasmic tail (129), while basolateral secretion of the major basement membrane heparan sulfate (HS) PG occurs by a pH-dependent sorting mechanism (130). The fact that different types of proteoglycans are secreted apically and basolaterally indicates that these molecules are actively sorted. While HSPGs are mainly secreted basolaterally from MDCK cells, CSPGs are mainly secreted apically (131). The observation that endogenously synthesized CSPGs of high molecular mass and individual hexyl-β-D xylosides that have acquired CS chains, are mainly secreted apically in MDCK cells, indicated that CS chains may contain essential apical sorting determinants (132). Still, the interpretation of the CS data is not straightforward, since the presence of a CS chain in amyloid precursor-like protein 2 did not alter the predominant basolateral secretion observed for the CS-deficient variant (133). Addition of a CS chain did also not alter the basolateral sorting pattern of the H1 subunit of the asialoglycoprotein receptor, when it was expressed in MDCK cells (134). Since the H1 subunit is a trans-membrane protein, the lack of effect of the CS chain could be explained by the fact that glycan signals often are recessive to signals in the protein core, as observed for transmembrane receptors like the polymeric IgA receptor (135) and the low density lipoprotein receptor (64). HS chains, on the other hand, could possess basolateral sorting information. The GPI-linked HSPG glypican was detected mainly at the basolateral surface of both CaCo-2 and MDCK cells (87). A variant of glypican lacking sites for HS attachment was transported predominantly to the apical surface in MDCK cells, presumably directed there by the GPI membrane anchor and/or the N-linked glycan groups. However, the latter question was not addressed by removal of the N-glycans. HS chains could either promote basolateral sorting of glypican, or simply interfere with the recognition of the apical sorting information present in the molecule, which comes into play when the HS chains have been removed. As discussed above, the effect of the HS chains could merely be blocking of oligomerization.
A number of studies have been carried out in polarized MDCK cells with the small CSPG serglycin as a model protein. Serglycin was expressed with green fluorescent protein fused to the C-terminus. The expressed serglycin-GFP fusion proteoglycan obtained mainly CS chains and was mainly secreted into the apical medium (Figure 2) of epithelial MDCK cells (53).

![Polarized secretion of serglycin-GFP in MDCK cells](image)

**Figure 2.** Polarized secretion of serglycin-GFP. A fusion protein of serglycin and GFP was expressed in polarized, filter-grown MDCK cells. Secretion of serglycin-GFP to the apical and basolateral medium was quantified from Western blots and expressed as % distribution in the apical and basolateral medium.

However, the minor fraction secreted into the basolateral medium was several times more intensely sulfated and showed a different sensitivity to the sulfation inhibitor chlorate (136). A sulfation intensity difference was also observed for the linker region sugars, indicating that the apical and basolateral serglycin-GFP molecules were segregated at an early stage of the secretory pathway, where the linker region is formed. A fraction of the serglycin constructs studied had the ability to bypass the Golgi apparatus and appear at the cell surface as a variant without GAG chains (137). Also this variant was secreted predominantly apically, indicating that a sorting event had already occurred prior to Golgi entry. This is in line with the observation that apical and basolateral glycoproteins in MDCK cells possess differential detergent extractability already in their high-mannose form (52), and raises the question whether glycans are signals directing a protein into the right pathway or merely a product of the pathway followed, due to a signal present in the immature protein or proteoglycan. The difference in sulfation intensity observed for apically and basolaterally secreted serglycin-GFP was essentially abolished upon neutralization of the secretory pathway with the macrolide Bafilomycin A1. The effects of neutralization were by far most prominent in the apical route (138). The differences in sulfation intensity in the apical and basolateral secretory route were observed also for endogenous proteins and could be
overcome by transfection leading to over-expression of the PAPS-transporter (139). This indicates that the difference observed in the sulfation intensity in the apical and basolateral secretory routes was due to a reduced concentration of nucleotide sulfate (PAPS) in the apical route, a difference that was subsequently counteracted by expression of more PAPS transporter. In a recent study, the GAG binding domain of serglycin, which has the ability to carry several GAG chains, was transferred to the non-glycosylated model protein rat growth hormone (rGH). The GAG domain still gave rise to CS chains in the new protein context, and redirected the randomly secreted protein rGH more towards the apical surface (Figure 3) of MDCK cells (140). This is the first published example of transplantable sorting information in a GAG binding site with chondroitin sulfate chains. An interesting aspect of this study was that, although the secretion polarity of the serglycin GAG domain was maintained in the new protein environment provided by rGH, the sulfation intensity was no longer different in the apical and basolateral secretory pathways, indicating that different routes have been followed, or different enzymatic regimes have been recruited for GAG synthesis (140).

![Polarized secretion of rGH variants in MDCK cells](image)

**Figure 3.** Polarized secretion of rat growth hormone (rGH) variants. Secretion from transfected polarized MDCK cells of a fusion protein of rGH with GFP was compared to that of an rGH-GFP fusion protein with the glycosaminoglycan (GAG) domain of serglycin inserted between rGH and GFP (rGH-GAG). Secretion of rGH and rGH-GAG to the apical and basolateral medium was quantified from Western blots and expressed as % distribution in the apical and basolateral medium.
6. Glycans are not the whole story

Although there evidently are many examples of glycan-mediated sorting to the apical surface of epithelia, there are many examples of apical sorting not requiring glycans. In fact, quite a few proteins are transported apically in a glycan-independent manner (141), like CD3-epsilon which has no N-glycan modification (142-146). It is evident that both proteins with (99) and without (142) N-linked sugars are transported to the apical surface, while some proteins carrying N-glycans are transported basolaterally (147). This could depend on variability in the structure of the N-glycans on different proteins, on the surrounding context in the protein (72), or the glycans could be recessive to other sorting signals in the molecule. The lectins localized to the early secretory pathway has affinity for high-mannose ligands, while the galectins generally has the better affinity for galactose containing glycans, found in both glycoproteins and glycolipids. While terminal processing of N-glycan structures seemed important for endolyn, gp114 and DPPIV (100, 148), there was little misorting of glycoproteins in a mutant MDCK cell line deficient in UDP-galactose delivery to the Golgi apparatus (149). A detailed investigation of the structure of the N-glycans attached to a dually glycosylated variant of rGH revealed only minor differences in these glycan structures after apical and basolateral transport and secretion, indicating that, although the glycosylated variant was mainly secreted apically, there were no observable differences that could mediate sorting based on terminal glycan structure (150).

Reduction in the sulfation level of GAG chains and attempts to generate shorter GAG chains have not resulted in alterations in the transport polarity of PGs. As long as the sorting machineries have not been outlined, the role of CS chains as apical sorting mediators and HS chains as basolateral sorting mediators remains incompletely understood.

Two classes of lectins have been implicated as mediators of apical sorting of N-glycanated proteins in polarized epithelial cells, the high mannose binding lectin VIP36, and possibly some related lectins, and the galectin family. While the former are mostly localized to the early secretory pathway, the latter family members are found in the lumen of endosomes. Thus, any suggestion for a glycan based sorting mechanism in the trans-Golgi network, a major sorting site in the epithelial MDCK cell line, lacks a description at the molecular level. How and when glycans influence the sorting and transport of their host glycoproteins and proteoglycans is therefore open to future investigations.

Author details

Kristian Prydz*, Gro Live Fagereng and Heidi Tveit
Department of Molecular Biosciences,
University of Oslo, Oslo, Norway

*Corresponding Author
7. References


The Role of Glycans in Apical Sorting of Proteins in Polarized Epithelial Cells


