74. E. K. Fan *et al.*, *J. Am. Chem. Soc.* **122**, 2663 (2000).
75. P. I. Kitov *et al.*, *Nature* **403**, 669 (2000).
76. W. F. Goebel, O. T. Avery, *J. Exp. Med.* **50**, 521 (1929).
77. R. U. Lemieux, D. R. Bundle, D. A. Baker, *J. Am. Chem. Soc.* **97**, 4076 (1975).
78. V. Pozsgay, *Angew. Chem. Int. Ed.* **37**, 138 (1998).
79. S. J. Danishefsky, J. R. Allen, *Angew. Chem. Int. Ed.* **39**, 837 (2000).
80. E. J. Gordon, W. J. Sanders, L. L. Kiessling, *Nature* **392**, 30 (1998).
81. E. J. Gordon, L. E. Strong, L. L. Kiessling, *Bioorg. Med. Chem. Lett.* **6**, 1293 (1998).
82. S. D. Rosen, *Am. J. Pathol.* **144**, 1013 (1999).
83. T. K. Kishimoto, M. A. Jutila, E. L. Berg, E. C. Butcher, *Science* **245**, 1238 (1989).
84. W. J. Sanders, E. J. Gordon, P. J. Beck, R. Alon, L. L. Kiessling, *J. Biol. Chem.* **274**, 5271 (1999).
85. S. Goon, C. R. Bertozzi, "Metabolic substrate engineering as a tool for glycobiology," in *Glycochemistry: Principles, Synthesis and Applications*, P. G. Wang, C. R. Bertozzi, Eds. (Dekker, New York), in press.
86. A. K. Sarkar, T. A. Fritz, W. H. Taylor, J. D. Esko, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3323 (1995).
87. F. N. Lugemwa, J. D. Esko, *J. Biol. Chem.* **266**, 6674 (1991).
88. H. Kayser *et al.*, *J. Biol. Chem.* **267**, 16934 (1992).
89. O. T. Keppler *et al.*, *J. Biol. Chem.* **270**, 1308 (1995).
90. O. T. Keppler *et al.*, *Biochem. Biophys. Res. Commun.* **253**, 437 (1998).
91. J. R. Wieser, A. Heisner, P. Stehling, F. Oesch, W. Reutter, *FEBS Lett.* **395**, 170 (1996).
92. B. E. Collins, T. J. Fralich, S. Itonori, Y. Ichikawa, R. L. Schnaar, *Glycobiology* **10**, 11 (2000).
93. L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **276**, 1125 (1997).
94. E. Saxon, C. R. Bertozzi, *Science* **287**, 2007 (2000).
95. K. J. Yarema, L. K. Mahal, R. Bruehl, E. C. Rodriguez, C. R. Bertozzi, *J. Biol. Chem.* **273**, 31168 (1998).
96. G. A. Lemieux, K. J. Yarema, C. L. Jacobs, C. R. Bertozzi, *J. Am. Chem. Soc.* **121**, 4278 (1999).
97. J. H. Lee *et al.*, *J. Biol. Chem.* **274**, 21878 (1999).
98. H. C. Hang, C. R. Bertozzi, *J. Am. Chem. Soc.* **123**, 1242 (2001).
99. The authors thank J. Gestwicki, L. Strong, E. Saxon, and K. Yarema for their comments on the manuscript, and the NIH and NSF for research support.

REVIEW

Intracellular Functions of N-Linked Glycans

Ari Helenius[1]* and Markus Aebi[2]

N-linked oligosaccharides arise when blocks of 14 sugars are added cotranslationally to newly synthesized polypeptides in the endoplasmic reticulum (ER). These glycans are then subjected to extensive modification as the glycoproteins mature and move through the ER via the Golgi complex to their final destinations inside and outside the cell. In the ER and in the early secretory pathway, where the repertoire of oligosaccharide structures is still rather small, the glycans play a pivotal role in protein folding, oligomerization, quality control, sorting, and transport. They are used as universal "tags" that allow specific lectins and modifying enzymes to establish order among the diversity of maturing glycoproteins. In the Golgi complex, the glycans acquire more complex structures and a new set of functions. The division of synthesis and processing between the ER and the Golgi complex represents an evolutionary adaptation that allows efficient exploitation of the potential of oligosaccharides.

In mature glycoproteins, N-linked glycan moieties are structurally diverse. The sugar composition and the number and size of branches in the sugar tree varies among glycans bound to a protein, among glycoproteins, and among cell types, tissues, and species (*1*, *2*). However, when initially added in the ER to growing nascent polypeptides, the glycans do not display such heterogeneity. The "core glycans" are homogeneous and relatively simple (Fig. 1).

The trimming and processing that the glycans undergo when the glycoprotein is still in the ER introduce only limited additional diversity, because the alterations are shared by all glycoproteins. Thus, the spectrum of glycoforms remains rather uniform until the glycoproteins reach the medial stacks of the Golgi apparatus, where structural diversification is introduced through a series of nonuniform modifications. Particularly in vertebrate and plant cells, it is the terminal glycosylation in the Golgi complex that gives rise to the tremendous diversity seen in glycoconjugates that reach the cell surface.

The switch from structural uniformity in the ER to diversification in the Golgi complex coincides with a marked change in glycan function. In the early secretory pathway, the glycans have a common role in promoting protein folding, quality control, and certain sorting events. Later, Golgi enzymes prepare them for the spectrum of novel functions that the sugars display in the mature proteins (*3*). Here, we mainly address events in the early secretory pathway. We focus on observations that are starting to unmask the logic of the various early trimming and modification events. We also discuss glycan structure and function in light of fundamental differences between the two biosynthetic organelles, the ER and the Golgi complex.

N-Linked Glycan Synthesis and Modification

During the synthesis of N-linked glycans in mammalian cells (Fig. 2), a 14-saccharide

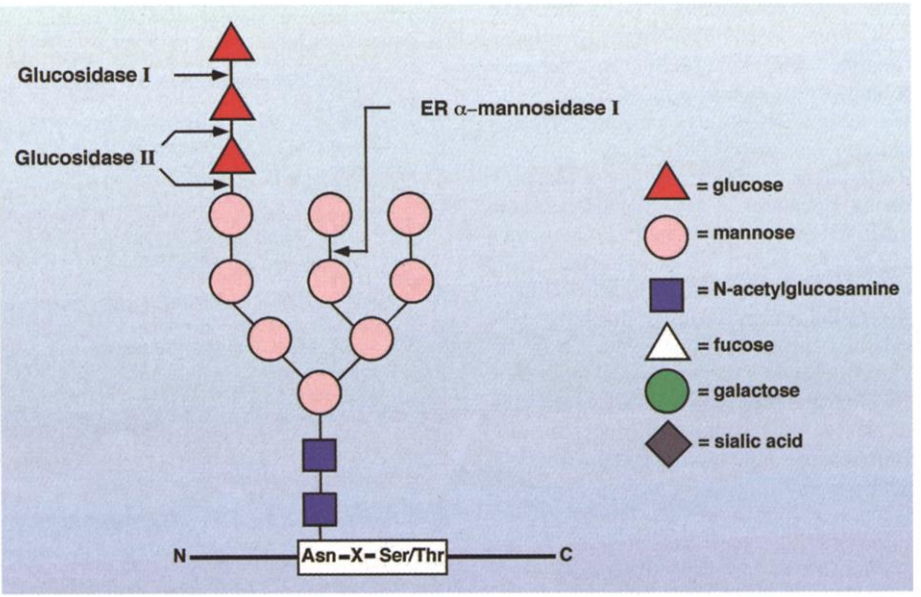

Fig. 1. The N-linked core oligosaccharide. N-linked glycans are added to proteins in the ER as "core oligosaccharides" that have the structure shown. These are bound to the polypeptide chain through an N-glycosidic bond with the side chain of an asparagine that is part of the Asn-X-Ser/Thr consensus sequence. Terminal glucose and mannose residues are removed in the ER by glucosidases and mannosidases. The symbols for the different sugars are used in the following figures.

[1]Institute of Biochemistry, [2]Institute of Microbiology, Eidgenössische Technische Hochschule Zürich, Universitätstrasse 16, CH-8092 Zürich, Switzerland.

*To whom correspondence should be addressed.

"core" unit is assembled as a membrane-bound dolichylpyrophosphate precursor by enzymes located on both sides of the ER membrane (*1, 4, 5*). The completed core oligosaccharide is transferred from the dolichylpyrophosphate carrier to a growing, nascent polypeptide chain, and is coupled through an N-glycosidic bond to the side chain of an asparagine residue. The oligosaccharyltransferase responsible for this transfer is a complex enzyme with its active site in the ER lumen (*6*). It recognizes a specific conformation of the glycosylation sequon (the sequence Asn-X-Ser/Thr) transiently formed when the growing, nascent polypeptide chain emerges from the translocon (*7, 8*). Of all sequons, it has been estimated that 90% are glycosylated (*9*). Sometimes there is variability in the efficiency by which individual sequons are recognized, giving rise to heterogeneously glycosylated products.

Immediately after coupling to the polypeptide chain, terminal glucose and mannose residues are removed by ER glucosidases and mannosidases (*1, 10*). When the glycoprotein moves to the Golgi complex, the glycan chains undergo further trimming of mannoses. In many cases, new sugars are added during terminal glycosylation to produce complex N-linked glycans. One of many possible terminal glycosylation pathways is shown in Fig. 2. Mature glycoproteins carry a mixture of complex glycans as well as some high-mannose glycans that have escaped terminal glycosylation.

This biosynthesis pathway is puzzling in several respects. Why does the cell assemble a large oligosaccharide, and then trim it down in order to rebuild it again in a different form? Why is the synthesis and modification shared between two separate biosynthetic organelles? Why does the ER contain enzymes both to remove and to restore glucose residues?

Operating Principles of the ER and the Golgi Complex

To better understand the switch that occurs in N-linked glycan biosynthesis midway through the maturation of a glycoprotein, it is first useful to discuss the division of labor between the ER and the Golgi complex. The two organelles follow each other in series, so that molecules originating in the ER pass through the Golgi complex on their way to their final intra- or extracellular locations. Although they are partners in the same pathway, the ER and the Golgi complex are fundamentally different with respect to operational principle, overall architecture, and probably evolutionary origin.

In the ER, most of the biosynthetic machinery faces the cytosol and can directly use precursors such as sugar nucleotides and amino acids provided by cytosolic enzymes. After synthesis, the products are translocated across the membrane into the lumen. This

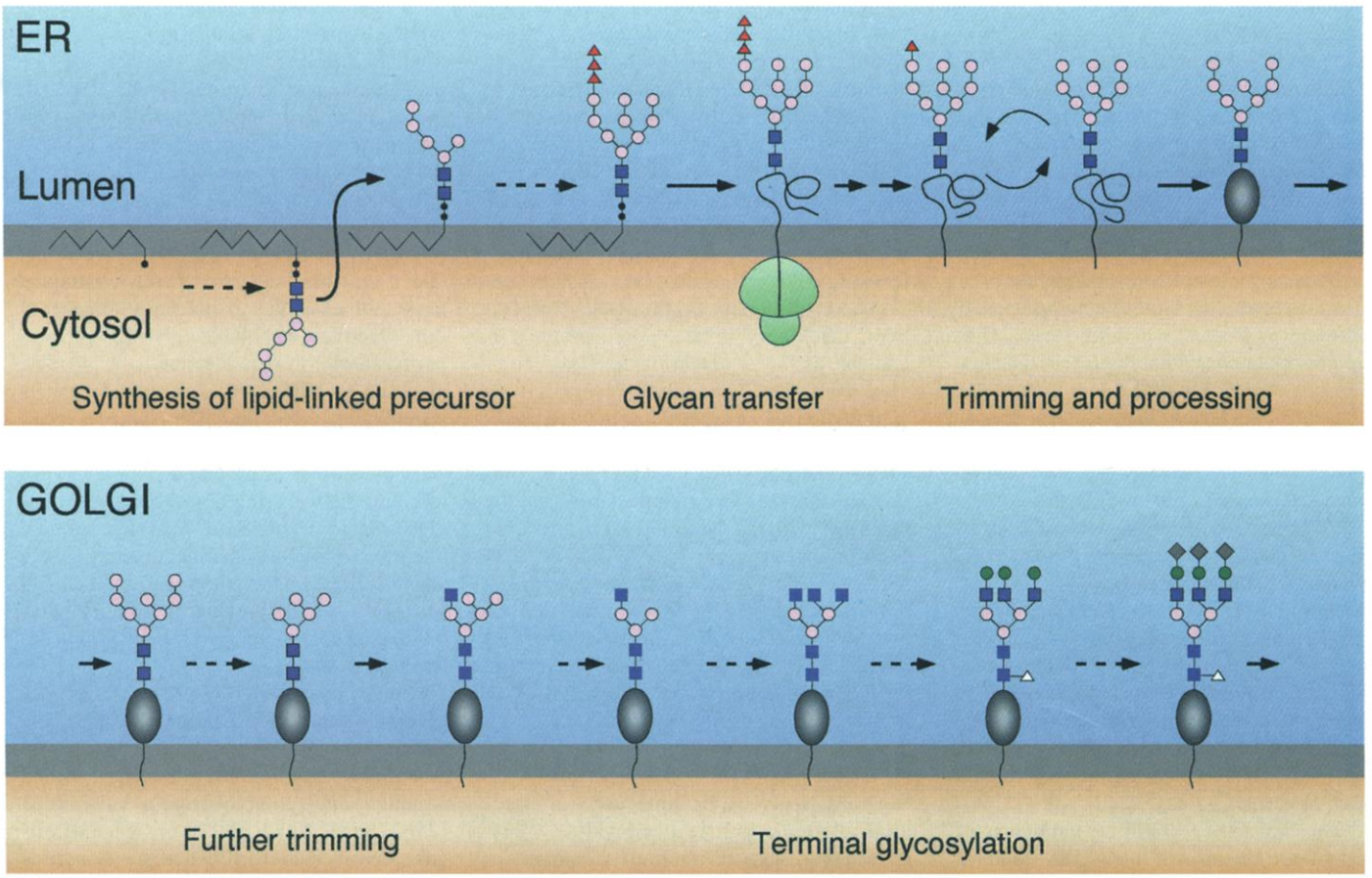

Fig. 2. Biosynthesis of the N-linked core oligosaccharide. Synthesis starts on the cytosolic surface of the ER membrane by the addition of sugars, one by one, to dolichylphosphate. When two N-acetylglucosamines and five mannoses have been added, the oligosaccharide is flipped to the lumenal side of the membrane, and seven further sugars are added from lipid precursors. After the last of the three glucoses have been added, the oligosaccharyltransferase enzyme complex catalyzes the transfer of the core oligosaccharide to the asparagine residues of nascent, growing polypeptide chains. The three glucoses are trimmed away by glucosidase I and II, and terminal mannoses by one or more different ER mannosidases. The ER also contains a glucosyltransferase that can reglucosylate glucose-free chains and thus establish, with glucosidase II, a deglucosylation-reglucosylation cycle. When the glycoprotein has folded (gray oval) and reached the Golgi complex, further mannose trimming occurs. The addition of a GlcNAc residue is followed by trimming of two additional mannoses. During subsequent terminal glycosylation there is addition of new terminal sugars including GlcNAc, galactose, sialic acid, and fucose. Of the original core glycan, just five sugars remain. Only one of many possible terminal glycosylation pathways is shown; the number of branches generated is variable, as are the number and identity of sugars added. Whereas the glycoforms in the ER are homogeneous, the Golgi-generated forms are highly diverse and differ widely between species.

topological arrangement applies to the polypeptide synthesis that occurs on membrane-bound ribosomes, as well as to the synthesis of phospholipids. It also applies, in part, to the synthesis of dolichol precursors for N-linked glycans (Fig. 2). In those few cases where precursors are transported to the lumenal side, transport occurs almost exclusively through lipid-linked intermediates (such as dolichylphosphomannose or dolichylphosphoglucose) rather than as soluble precursors.

In contrast, the Golgi complex assembles its synthetic products on the lumenal side. This is dictated by the fact that the proteins or protein domains to be modified have already been translocated to the lumenal side in the ER. Another major difference is that synthesis in the Golgi complex makes use of soluble precursors (nucleotide sugars and other nucleotide precursors). These are imported from the cytosol through the Golgi membrane by specific transporters (*11*). Lipid-linked intermediates, such as those in the ER, are not used.

The ER contains a high concentration of soluble molecular chaperones and folding enzymes. Substrate proteins constitute a minority population in this large organelle (*12*). In the Golgi complex, resident proteins (glycosidases, glycosyltransferases, proteolytic enzymes, permeases, lectins, and transport receptors) are membrane-bound. Synthetic events thus occur on the inside surface of the membrane, with the lumen being occupied by substrate molecules and other cargo. These are present in considerably higher concentrations than in the ER (*13*).

Unlike the ER, the Golgi complex does not have a rigorous system for controlling the fidelity of its biosynthetic processes. For example, when cells are treated with glycosidase inhibitors, or when specific glycosyltransferases are mutated, glycoproteins carrying defective glycan moieties are readily exported (*14, 15*). Even severe glycosylation defects or deficiencies are, as a rule, tolerated by the Golgi without retention or degradation. But analysis of the secreted products of normal cells shows that they are appropriately modified. Thus, although lacking ER-like quality control and degradation systems, the Golgi complex does not export unfinished products.

The explanation lies in the multicompartmental architecture of the Golgi complex. Substrate molecules are progressively modified as they move from the entry side (cis) to the exit side (trans) (*16*). Glycoproteins must traverse the whole stack before they can be exported. Most glycoproteins end up spending about the same length of time (usually 5 to 15 min) in the Golgi exposed to the full panel of modifying enzymes. The programmed cis-to-trans movement—whether by vesicular transport, percolation, or cisternal progression (*17*)—ensures a high degree of completeness in the biosynthetic processes without imposing direct quality control. This, in turn, makes it possible for the Golgi to produce such a wide diversity of structures. It is tempting to speculate that the need to ensure full processing without imposing stringent, ER-like exit criteria on individual molecules is one of the reasons why the Golgi complex is divided into subcompartments.

Like protein translocation, N-linked glycosylation clearly belongs to the functions that the ER has inherited from the prokaryotic, most likely archaeal, plasma membrane (*18, 19*). In archaea, N-linked glycans occur in the S-layer of the cell wall. As in eukaryotes, synthesis makes use of dolichylphosphate- and dolichylpyrophosphate-linked oligosaccharides. This makes sense because the coupling of protein and carbohydrate is likely to occur extracellularly, and, if exported, soluble sugar nucleotides would diffuse away. Analysis at the genomic level confirms the presence of a biosynthetic pathway for dolichylphosphate-linked oligosaccharides in *Archaeoglobus fulgidus* (*5*). In addition, all archaeal genomes sequenced so far contain one or more homologs of the eukaryotic oligosaccharyltransferase subunit STT3 (*20, 21*). In archaea, the oligosaccharides are transferred to asparagine side chains in the same Asn-X-Ser/Thr sequence motif that is used by the ER oligosaccharyltransferase in eukaryotic cells (*19*).

The evolutionary origin of the glycosylation machinery in the Golgi complex is unclear, as is the origin of the Golgi complex itself. The sequences of Golgi glycosyltransferases suggest homology with cytosolic enzymes responsible for the use of sugar nucleotides (*22, 23*). This implies that the enzyme system in the Golgi and the machinery in the ER may have different evolutionary origins. It would also seem reasonable to assume that the machinery now present in the Golgi complex evolved after an intracellular membrane system was already in place, as nucleotide sugars and soluble precursors would diffuse away if transported out of the cell.

N-Linked Glycans Help Proteins Fold

Whereas N-glycosidic links serve a structural role in the archaeal cell wall (*19, 24*), a variety of other functions dominate in eukaryotes. The single most important one is the promotion of proper folding of newly synthesized polypeptides in the ER (*2, 25*). This eukaryotic adaptation of glycan function allows cells to produce and secrete larger and more complex proteins at higher levels. It also explains why the addition of N-linked glycans must occur cotranslocationally in the ER (i.e., before the folding process has begun).

When glycosylation is inhibited, the most commonly observed effect is the generation of misfolded, aggregated proteins that fail to reach a functional state (*2, 25, 26*). The importance of the added glycans varies between proteins and depends on the physiological context. Some proteins are completely dependent on glycosylation, whereas many display no dependence at all. Some are partially dependent, some become temperature-sensitive for folding, and some are glycan-dependent in one cell type but not in another. When the importance of individual glycosylation sites is evaluated, it is often found that some are more important than others. This implies that oligosaccharide appendices have local effects on protein folding. However, when multiple sites are modified, folding may be compromised even though none of the glycans are needed individually. Thus, the oligosaccharides seem to have both "local" and "global" effects on the folding process.

It is well known that glycans can have a direct effect on the folding process (*27, 28*). Studies performed using peptides indicate that although an N-linked glycan does not induce permanent secondary structure, it alters the conformational preferences close to the glycosylation site, resulting in more compact conformations (*28*). A truncated N-linked glycan [the disaccharide chitobiose ($GlcNAc_{\beta 1\text{-}4}GlcNAc$, where Ac is acetyl and Glc is glucose)] attached to an asparagine tends, for example, to induce a compact β turn. One-third of N-linked glycans in glycoproteins occur in such locations (*29, 30*). The presence of a large polar saccharide unit is also likely to affect the folding process locally by orienting the polypeptide segment toward the surface of protein domains. In vitro refolding studies using glycosylated and nonglycosylated versions of the same protein confirm that oligosaccharide moieties have a positive effect on the folding process (*27, 31*) and can keep the proteins in solution, thus mimicking the role played by many molecular chaperones (*32*).

Before discussing the indirect effects of glycans on the folding process, it is important to point out that glycans are not usually essential for maintaining the overall folded structure once a glycoprotein has folded (*25, 27, 33*). Although the presence of oligosaccharides does influence the properties of the polypeptide moiety—for example, by increasing stability (*27, 28*)—the effect is usually rather small. Consequently, glycans can, as a rule, be removed from a folded glycoprotein without major effects on protein conformation.

The Calnexin-Calreticulin Cycle

The most important indirect effect of glycans on folding involves a unique chaper-

one system found in the ER of nearly all eukaryotes, the so-called calnexin-calreticulin cycle (*34–36*) (Fig. 3). Glycans serve as an "admission ticket" to this cycle. Calnexin (membrane-bound) and calreticulin (soluble) are homologous ER lectins that bind transiently to virtually all newly synthesized glycoproteins (*37, 38*). They are highly asymmetric molecules with a long, curved, hydrophilic peptide arm formed by the P-domain (*39–41*). This domain may interact with the cochaperone ERp57, or it may form a region of the chaperone that protects the bound substrate molecule during folding; it is unlikely to form the lectin binding site. Calnexin and calreticulin interact with the glycan moieties of substrate glycoproteins after they have been trimmed by glucosidases I and II to the monoglucosylated form ($Glc_1Man_{9\text{-}6}GlcNAc_2$, where Man is mannose) (*38, 42, 43*). If there is a glycosylation site within the first 50 amino acids of a nascent glycoprotein, interaction with calnexin and calreticulin begins cotranslationally (*44*). After chain termination, binding to the lectins continues for a period ranging from a few minutes to several hours, depending on the rate of folding.

Calnexin and calreticulin form a complex with ERp57, a thiol oxidoreductase homolog of protein disulfide isomerase (*45*). Short-lived disulfide bonds occur between ERp57 and cysteines in calnexin- and calreticulin-bound substrate glycoproteins (*46*). Such bonds serve as intermediates in oxidation and isomerization reactions and lead to the formation of correctly paired disulfide bonds (*47*). For many glycoproteins, the interaction with calnexin, calreticulin, and ERp57 slows down the rate of folding but increases efficiency. Even though not absolutely necessary for the folding of a protein, the lectins retain misfolded conformers in the ER and thus mediate the quality control process that regulates ER-to-Golgi transport, and possibly the ER degradation process as well (see below). There are also reports that calnexin and calreticulin can interact with polypeptide chains devoid of oligosaccharides (*48, 49*), but it is unclear whether this is important for protein maturation in cells.

Association of glycoproteins with calnexin and calreticulin involves a binding-and-release cycle driven by the opposing actions of two soluble ER enzymes, uridine diphosphate (UDP)–glucose:glycoprotein glucosyltransferase (GT) and glucosidase II (*36, 38, 50*). The former adds a glucose residue to high-mannose glycans, and the latter removes it (Fig. 3). In this cycle, GT serves as the folding sensor. In vivo and in vitro studies have shown that this remarkable enzyme only reglucosylates incompletely folded glycoproteins; that is, glucose addition depends on the folding status of the protein moiety (*36, 51, 52*). It is apparent that GT specifically reglucosylates glycans in glycoprotein domains that have a partially folded structure (*53*). Glycans in nearby folded domains, or in domains that have a random coil conformation, are not glucosylated (*54*). Substrate recognition by GT is triggered not by specific signal sequences but by general, biophysical properties common to partially folded glycoproteins. These may include exposed hydrophobic patches, or features such as excessive dynamic mobility. By recognizing the folding status, GT forces incompletely folded conformers to remain in the calnexin-calreticulin cycle, whereas folded conformers are allowed to proceed further in the secretory pathway.

The calnexin-calreticulin cycle seems to be essential in vivo. Transgenic mice devoid of calreticulin die on embryonic day 18 (*55*). The essential nature of the pathway is also

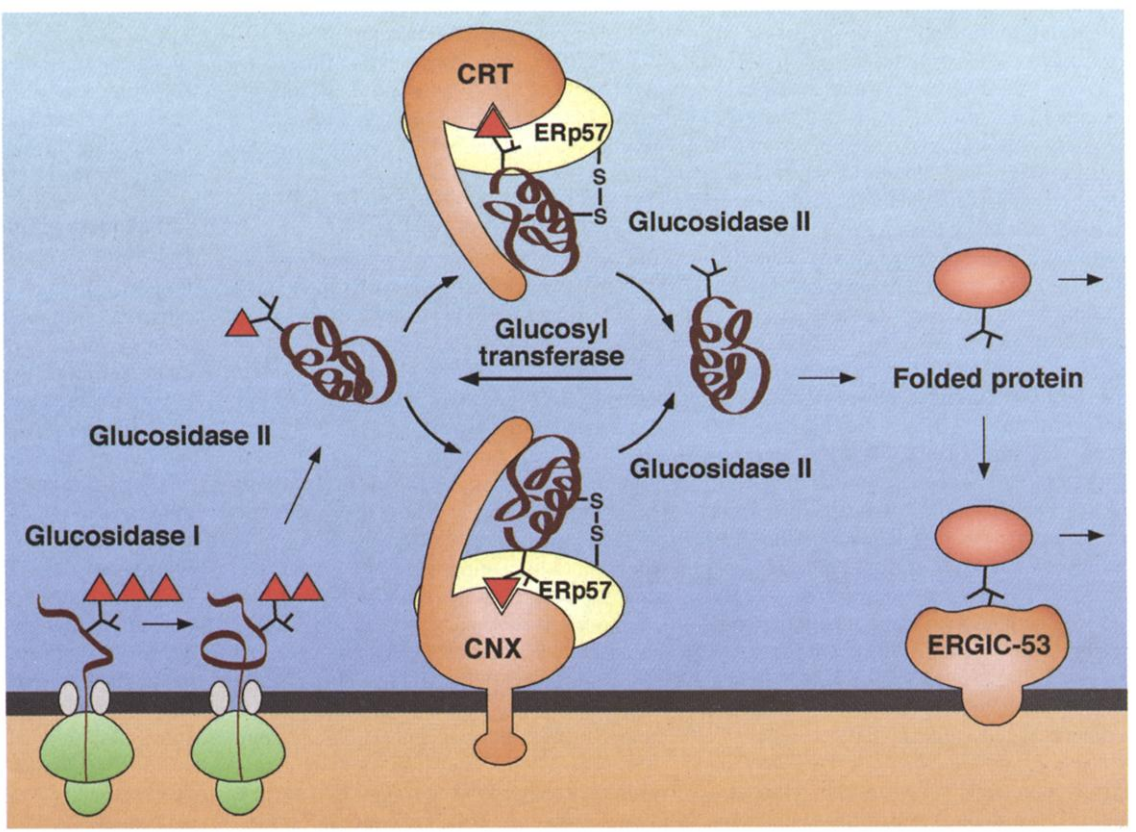

Fig. 3. The calnexin-calreticulin cycle. When two of the glucoses in the N-linked core glycans have been trimmed away by glucosidases I and II, the nascent or newly synthesized glycoproteins bind to calnexin (CNX) and/or calreticulin (CRT), two homologous ER lectins specific for monoglucosylated core oligosaccharides. The protein is thereby exposed to another folding factor, ERp57, a thiol oxidoreductase that binds to both calnexin and calreticulin. If the glycoproteins have cysteines, the formation of disulfide bonds is catalyzed through the formation of transient mixed disulfides with ERp57. When the remaining third glucose residue is trimmed by glucosidase II, the complexes dissociate. If the glycoprotein is not folded at this time, the oligosaccharides are reglucosylated by an ER glucosyltransferase, and the protein reassociates with the lectins. The cycle is repeated until the protein is either folded or degraded. Once correctly folded, a glycoprotein is no longer recognized by the glucosyltransferase, and because it is no longer reglucosylated, it will not bind back to calnexin and/or calreticulin. It can now leave the ER. Exit of certain glycoproteins from the ER to the Golgi complex is assisted by another membrane-bound lectin, ERGIC-53, which binds to mannose residues. The calnexin-calreticulin cycle promotes correct folding, inhibits aggregation of folding intermediates, blocks premature oligomerization, regulates ER degradation, and provides quality control by preventing incompletely folded glycoproteins from exiting to the Golgi complex.

illustrated by the lethal outcome of an inherited glucosidase I deficiency involving a neonate born with severe generalized hypotonia and dysmorphic features (*56*). The defects were likely caused by a shutdown of the calnexin-calreticulin cycle. That glycan modifications in the Golgi complex could still occur normally without glucose removal in the ER is explained by the presence of an endomannosidase enzyme in the Golgi complex (*57*).

Oligosaccharides as Signals in ER-Associated Protein Degradation

Trimming of the N-linked glycans also plays a role in the sorting process leading to glycoprotein degradation in the ER. Proteins that fail to reach their native conformation in the ER are selectively eliminated by ER-associated degradation (ERAD) (*58, 59*). This fate is shared by misfolded and mutant proteins, by orphan subunits of oligomers, and by some heterologously expressed proteins. Because misfolded side products are common even under unstressed conditions, ERAD has a central clearance function in the cell. Instead of using a lumenal degradation system that might endanger protease-sensitive folding intermediates, cells transport proteins that need to be destroyed to the cytosol, where they are ubiquitinated and degraded by 26S proteasomes (*59–61*).

When trimming by ER–mannosidase I is prevented by inhibitors or genetic manipulation, degradation of glycoproteins essentially stops (*62–64*). This mannosidase removes a single α-1,2–linked mannose residue from the α-1,3 branch of the core oligosaccharide, resulting in a $Glc_{0-3}Man_8GlcNAc_2$ structure (B isoform) (Fig. 1). It is apparent that the resulting Man_8 structures serve as part of the signal needed for ERAD. Removal of the mannose is not sufficient, however, because most proteins that have folded normally are mannose-trimmed before leaving the ER. Given that the mannosidase is relatively slow-acting, it has been suggested that it provides a timer function that awards protection against premature degradation to the most recently synthesized glycoproteins (*25, 64, 65*). With the mannose gone from the glycans, the glycoprotein can be degraded, provided that it is incompletely folded.

How the system works is not clear. Recent studies with mutant α_1-antitrypsin have led to a model in which calnexin plays a central role by sequestering $Glc_1Man_8GlcNAc_2$-containing glycoproteins (*64*). In this case, GT would be the folding sensor for degradation. This cannot be the only pathway, however, because mannosidase-dependent degradation of some glycoproteins occurs without involvement of either calnexin or GT (*66, 67*). Moreover, *Saccharomyces cerevisiae* has an effective mannosidase I–dependent ERAD system (*65*) but lacks a calnexin cycle and GT activity (*68*). The available data suggest multiple, parallel pathways leading to degradation, with mannose trimming as one of the most conspicuous sorting criteria.

Role in Transport and Targeting

ERGIC-53 and VIP36 are homologous, mannose-specific lectins in the Golgi complex and the early secretory pathway (*69–73*). They are both membrane proteins with a lumenal domain homologous to leguminous plant lectins. It has been proposed that ERGIC-53 serves as a cargo capture and transport receptor for ER-to-Golgi traffic of glycoproteins (Fig. 3) (*69*). VIP36 may play a similar function.

ERGIC-53 (also called p58) is the better characterized of the two (*69, 74*). It is a homo-oligomeric membrane protein that shuttles among the ER, the intermediate compartment, and the Golgi complex. Although it is not essential for the secretion or maturation of any known glycoprotein, ERGIC-53 has been shown to associate with human cathepsin Z and accelerate its transport to the Golgi in a glycan-dependent fashion (*69*). Moreover, patients with nonfunctional ERGIC-53 have been found to suffer from an autosomal recessive bleeding disorder with combined deficiency of coagulation factors V and VIII (F5F8D) (*75*). The plasma levels of these two highly glycosylated coagulation factors are reduced to 5 to 30% of normal, suggesting a hepatic secretion defect.

Mannose-6-Phosphate and Targeting of Lysosomal Hydrolases

The selective targeting of lysosomal hydrolases from the trans-Golgi network (TGN) to endosomes and lysosomes also makes use of lectins (*76*). The key event occurs in the cis-Golgi compartment where a resident enzyme (UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase) specifically modifies N-linked glycan moieties in lysosomal hydrolases by adding GlcNAc-phosphate to terminal or subterminal mannoses. After removal of the GlcNAc residue, the remaining mannose-6-phosphate group is recognized by mannose-6-phosphate receptors (M-6-P receptors) in the TGN. These include the dimeric cation-dependent M-6-P receptor (CD-MPR) and the much larger, multifunctional insulin-like growth factor II/cation-independent M-6-P receptor (IGF-II/CI-MPR). Their role is to sequester lysosomal enzymes in the TGN by associating with one or more of the oligosaccharides, escort them via clathrin-coated vesicles to endosomes, and return empty to the TGN. That these receptors serve a cargo capture, targeting, and transport function is beyond doubt.

In this well-characterized process, the phosphotransferase, a large, multisubunit enzyme in the cis-Golgi membrane, serves as the primary sensor (*77*). It selects the lysosomal enzymes among all cargo coming from the ER. Like GT, it relies on conformation-dependent substrate recognition, but differs in that it only modifies folded forms of the appropriate substrate proteins. Recognition requires critically spaced lysine residues in the surface of the protein and involves, in addition, a wide surface of contact (*78*). In some of the substrates, additional lysines direct the enzyme to one of multiple N-linked glycans.

The M-6-P receptors are homologous type I membrane proteins that differ in size, oligomerization state, and pH dependence of ligand binding. Together they constitute the P-type family of animal lectins, and are thus clearly different from the other intracellular lectins described above. The crystal structure of the lumenal domain of CD-MPR shows that the protein is a homodimer with each monomer composed of a nine-stranded, avidin-like, flattened β barrel with the mannose-6-phosphate binding site in a cavity between two loops (*79*). The pH dependence of the receptors is critical because net transport depends on the higher affinity for the ligand at the pH prevailing in the loading compartment, the TGN (pH 6.3 to 6.5), than in the release compartment, the late endosome (pH 5). IGF-II/CI-MPR also binds mannose-6-phosphate at neutral pH and is therefore, unlike CD-MPR, able to endocytose lysosomal enzymes from the cell surface.

Other Intracellular Functions of N-Linked Glycans

Resident proteins of the ER, the Golgi complex, endosomes, and lysosomes are themselves often glycoproteins. Their sugar composition reflects their location; early in the pathway they carry mainly high-mannose glycans, whereas in the Golgi complex or in post-Golgi organelles they contain complex sugars. On the whole, they are more sparingly glycosylated than glycoproteins in the plasma membrane and extracellular space.

Lysosomes and related organelles constitute an exception. The limiting membrane contains a collection of highly glycosylated membrane proteins, the most abundant of which are the homologs LAMP-1 (lgp110) and LAMP-2 (lgp120) (*80–82*). These have up to 20 N-linked glycans each, and it has been estimated that they occur in sufficient concentrations to form a nearly continuous, carbohydrate-covered coat over the lumenal surface of the lysosomal membrane (*81*). The glycans have been shown to protect the LAMPs against degradation by lysosomal proteases (*83*). LAMP-2 was recently dem-

onstrated to be required for efficient lysosomal degradation after autophagocytosis (*84, 85*).

Summary

Glycosylation differs from most other covalent protein modifications with respect to the size and complexity of the added group and the magnitude of the cellular machinery devoted to synthesis and modulation. Why do cells need such an elaborate and costly system? Why are so many proteins in the eukaryotic cell glycosylated? Why do the N-linked glycans undergo so many changes during glycoprotein maturation?

At a conceptual level, it is important to recognize that one of the limitations imposed by the linear nature of polypeptide chains is the lack of possibilities to generate branched structures. Addition of oligo- and polysaccharides (N-linked or O-linked) provides a way to circumvent this limitation. In the case of N-linked glycans, the branches reach more than 3 nm from the protein surface as bulky, mobile carbohydrate clusters that are themselves branched. The outer parts are so far from the protein surface that they can act as essentially independent domains.

One of the explicit advantages of adding these polar branches is that cells can produce and secrete proteins of greater complexity and with better efficiency than would otherwise be the case. As we have seen, many plasma membrane proteins and secretory proteins are not able to fold without added glycans. The evolution of proteins into better folders has been facilitated not only by the addition of glycans, but probably also by the generation and elimination of glycosylation sequons in the primary sequence. This can occur by single point mutation, allowing easy modulation of folding parameters through the shuffling of glycosylation sites.

Because carbohydrates differ in biosynthesis and overall properties from amino acids, the joining of polypeptide and carbohydrate elements into hybrid molecules provides an opportunity to add new functionalities and specificities. Differential terminal glycosylation pathways allow, for example, fine-tuning of protein properties in a cell- and tissue-specific manner without a change in amino acid sequence.

The large number of possible ways to link sugars to each other makes oligosaccharides ideal as compact and versatile recognition markers. It is not uncommon, as shown by calnexin and the M-6-P receptors, that a single sugar residue or sugar-associated group suffices as a distinctive marker for lectin binding. The more complex a eukaryotic organism, the more numerous are the ways in which it seems to make use of this opportunity for intra- and extracellular recognition. Intracellularly, N-linked glycans with minor molecular differences are used as a universal molecular sorting determinant to guide a heterogeneous cohort of glycoproteins through sorting processes such as quality control, ERAD, and lysosomal targeting. Although the system of lectins and other factors involved is still incompletely characterized, it is apparent that N-linked glycans allow the ER and the Golgi complex to keep track of information such as folding status, time spent in the ER, and final destination. Like luggage handling in an airport, the task is greatly simplified by the use of a uniform "tag" applied to virtually every piece of cargo.

References and Notes

1. R. Kornfeld, S. Kornfeld, *Annu. Rev. Biochem.* **54**, 631 (1985).
2. J. C. Paulson, *Trends Biochem. Sci.* **14**, 272 (1989).
3. A. Varki *et al.*, *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999).
4. C. G. Gahmberg, M. Tolvanen, *Trends Biochem. Sci.* **21**, 308 (1996).
5. P. Burda, M. Aebi, *Biochim. Biophys. Acta* **1426**, 239 (1999).
6. S. Silberstein, R. Gilmore, *FASEB J.* **10**, 849 (1996).
7. E. Bause, *Biochem. J.* **209**, 331 (1983).
8. B. Imperiali, K. W. Rickert, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 597 (1995).
9. Y. Gavel, H. G. von Heijne, *Protein Eng.* **3**, 433 (1990).
10. K. W. Moremen, R. B. Trimble, A. Herscovics, *Glycobiology* **4**, 113 (1994).
11. C. B. Hirschberg, P. W. Robbins, C. Abeijon, *Annu. Rev. Biochem.* **67**, 49 (1998).
12. P. Quinn, G. Griffiths, G. Warren, *J. Cell Biol.* **98**, 2142 (1984).
13. W. E. Balch, J. M. McCaffery, H. Plutner, M. G. Farquhar, *Cell* **76**, 841 (1994).
14. A. D. Elbein, *FASEB J.* **5**, 3055 (1991).
15. P. Stanley, *Annu. Rev. Genet.* **18**, 525 (1984).
16. G. E. Palade, *Science* **189**, 347 (1975).
17. H. R. Pelham, J. E. Rothman, *Cell* **102**, 713 (2000).
18. T. D. Bugg, P. E. Brandish, *FEMS Microbiol. Lett.* **119**, 255 (1994).
19. J. Lechner, F. Wieland, *Annu. Rev. Biochem.* **58**, 173 (1989).
20. U. Spirig, *Mol. Gen. Genet.* **256**, 628 (1997).
21. R. Knauer, L. Lehle, *J. Biol. Chem.* **274**, 17249 (1999).
22. I. M. Saxena, R. M. Brown Jr., M. Fevre, R. A. Geremia, B. Henrissat, *J. Bacteriol.* **177**, 1419 (1995).
23. C. A. Wiggins, S. Munro, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7945 (1998).
24. P. Messner *et al.*, *FEMS Microbiol. Rev.* **20**, 25 (1997).
25. A. Helenius, *Mol. Biol. Cell* **5**, 253 (1994).
26. A. Varki, *Glycobiology* **3**, 97 (1993).
27. B. Imperiali, S. E. O'Connor, *Curr. Opin. Chem. Biol.* **3**, 643 (1999).
28. M. R. Wormald, R. A. Dwek, *Structure Fold. Des.* **7**, R155 (1999).
29. J. J. Beintema, *Biosci. Rep.* **6**, 709 (1986).
30. S. E. O'Connor, B. Imperiali, *Chem. Biol.* **3**, 803 (1996).
31. G. Kern, D. Kern, R. Jaenicke, R. Seckler, *Protein Sci.* **2**, 1862 (1993).
32. R. Jaenicke, *Biochemistry* **30**, 3147 (1991).
33. K. Olden, J. B. Parent, S. L. White, *Biochim. Biophys. Acta* **650**, 209 (1982).
34. A. Helenius, E. S. Trombetta, D. N. Hebert, J. F. Simons, *Trends Cell Biol.* **7**, 193 (1997).
35. A. Zapun, C. A. Jakob, D. Y. Thomas, J. J. Bergeron, *Structure Fold. Des.* **7**, R173 (1999).
36. A. J. Parodi, *Biochem. J.* **348**, 1 (2000).
37. W.-J. Ou, P. H. Cameron, D. Y. Thomas, J. J. M. Bergeron, *Nature* **364**, 771 (1993).
38. C. Hammond, I. Braakman, A. Helenius, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 913 (1994).
39. M. Bouvier, W. F. Stafford, *Biochemistry* **39**, 14950 (2000).
40. L. Ellgaard *et al.*, *FEBS Lett.* **488**, 69 (2000).
41. L. Ellgaard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3133 (2001).
42. F. E. Ware *et al.*, *J. Biol. Chem.* **270**, 4697 (1995).
43. R. G. Spiro, Q. Zhu, V. Bhoyroo, H. D. Soling, *J. Biol. Chem.* **271**, 11588 (1996).
44. M. Molinari, A. Helenius, *Science* **288**, 331 (2000).
45. J. D. Oliver, F. J. van der Wal, N. J. Bulleid, S. High, *Science* **275**, 86 (1997).
46. M. Molinari, A. Helenius, *Nature* **402**, 90 (1999).
47. J. B. Huppa, H. L. Ploegh, *Cell* **92**, 145 (1998).
48. Y. Ihara, M. F. Cohen-Doyle, Y. Saito, D. B. Williams, *Mol. Cell* **4**, 331 (1999).
49. Y. Saito, Y. Ihara, M. R. Leach, M. F. Cohen-Doyle, D. B. Williams, *EMBO J.* **18**, 6718 (1999).
50. E. S. Trombetta, J. F. Simons, A. Helenius, *J. Biol. Chem.* **271**, 27509 (1996).
51. M. C. Sousa, M. A. Ferrero-Garcia, A. J. Parodi, *Biochemistry* **31**, 97 (1992).
52. K. S. Cannon, A. Helenius, *J. Biol. Chem.* **274**, 7537 (1999).
53. E. S. Trombetta, A. Helenius, *EMBO J.* **18**, 3282 (1999).
54. C. Ritter, A. Helenius, *Nature Struct. Biol.* **7**, 278 (2000).
55. N. Mesaeli *et al.*, *J. Cell Biol.* **144**, 857 (1999).
56. C. M. De Praeter *et al.*, *Am. J. Hum. Genet.* **66**, 1744 (2000).
57. R. G. Spiro, in *Carbohydrate in Chemistry and Biology*, B. Ernst, G. W. Hart, P. Sinay, Eds. (Wiley-VCH, Weinheim, Germany, 2000), vol. 3, pp. 65–79.
58. R. D. Klausner, R. Sitia, *Cell* **62**, 611 (1990).
59. R. K. Plemper, D. H. Wolf, *Trends Biochem. Sci.* **24**, 266 (1999).
60. J. S. Bonifacino, A. M. Weissman, *Annu. Rev. Cell Dev. Biol.* **14**, 19 (1998).
61. R. R. Kopito, *Cell* **88**, 427 (1997).
62. K. Su, T. Stoller, J. Rocco, J. Zemsky, R. Green, *J. Biol. Chem.* **268**, 14301 (1993).
63. M. Knop, N. Hauser, D. H. Wolf, *Yeast* **12**, 1229 (1996).
64. Y. Liu, P. Choudhury, C. M. Cabral, R. N. Sifers, *J. Biol. Chem.* **272**, 7946 (1997).
65. C. A. Jakob, P. Burda, J. Roth, M. Aebi, *J. Cell Biol.* **142**, 1223 (1998).
66. J. Wang, A. L. White, *Biochemistry* **39**, 8993 (2000).
67. F. Tokunaga, C. Brostrom, T. Koide, P. Arvan, *J. Biol. Chem.* **275**, 40757 (2000).
68. A. Parodi, *Annu. Rev. Biochem.* **69**, 69 (2000).
69. H. Hauri, C. Appenzeller, F. Kuhn, O. Nufer, *FEBS Lett.* **476**, 32 (2000).
70. K. Fiedler, R. G. Parton, R. Kellner, T. Etzold, K. Simons, *EMBO J.* **13**, 1729 (1994).
71. S. Hara-Kuge, T. Ohkura, A. Seko, K. Yamashita, *Glycobiology* **9**, 833 (1999).
72. J. Fullekrug, P. Scheiffele, K. Simons, *J. Cell Sci.* **112**, 2813 (1999).
73. C. Arar *et al.*, *J. Biol. Chem.* **270**, 3551 (1995).
74. U. Lahtinen, B. Dahllof, J. Saraste, *J. Cell Sci.* **103**, 321 (1992).
75. W. C. Nichols *et al.*, *Cell* **93**, 61 (1998).
76. S. Kornfeld, *Biochem. Soc. Trans.* **18**, 367 (1990).
77. M. Bao, J. L. Booth, B. J. Elmendorf, W. M. Canfield, *J. Biol. Chem.* **271**, 31437 (1996).
78. A. Nishikawa, A. Nanda, W. Gregory, J. Frenz, S. Kornfeld, *J. Biol. Chem.* **274**, 19309 (1999).
79. D. L. Roberts, D. J. Weix, N. M. Dahms, J. J. Kim, *Cell* **93**, 639 (1998).
80. Y. Cha, S. M. Holland, J. T. August, *J. Biol. Chem.* **265**, 5008 (1990).
81. B. L. Granger *et al.*, *J. Biol. Chem.* **265**, 12036 (1990).
82. D. M. Fambrough, K. Takeyasu, J. Lippincott-Schwarz, N. R. Siegel, *J. Cell Biol.* **106**, 61 (1988).
83. R. Kundra, S. Kornfeld, *J. Biol. Chem.* **274**, 31039 (1999).
84. Y. Tanaka *et al.*, *Nature* **406**, 902 (2000).
85. I. Nishino *et al.*, *Nature* **406**, 906 (2000).

86. We thank J. Helenius, K. Simons, J. Brunner, and T. Hennet for discussions and R. Moser for the illustrations. Supported by grants from the Swiss National Science Foundation to both authors.