CSH
PERSPECTIVES

Cold Spring Harbor
Perspectives in Biology

Evolutionary Forces Shaping the Golgi Glycosylation Machinery: Why Cell Surface Glycans Are Universal to Living Cells

Ajit Varki

Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**
Subject collections	Articles on similar topics can be found in the following collections The Golgi (11 articles)

Early Release Articles are published online ahead of the issue in which they appear. The online first release date serves as the official date of publication.

To subscribe to *Cold Spring Harbor Perspectives in Biology* go to:
http://cshperspectives.cshlp.org/site/misc/subscribe.xhtml

Evolutionary Forces Shaping the Golgi Glycosylation Machinery: Why Cell Surface Glycans Are Universal to Living Cells

Ajit Varki

Glycobiology Research and Training Center, Center for Academic Research and Training in Anthropogeny, University of California at San Diego, La Jolla, California 92093

Correspondence: a1varki@ucsd.edu

Despite more than 3 billion years since the origin of life on earth, the powerful forces of biological evolution seem to have failed to generate any living cell that is devoid of a dense and complex array of cell surface glycans. Thus, cell surface glycans seem to be as essential for life as having a DNA genetic code, diverse RNAs, structural/functional proteins, lipid-based membranes, and metabolites that mediate energy flux and signaling. The likely reasons for this apparently universal law of biology are considered here, and include the fact that glycans have the greatest potential for generating diversity, and thus evading recognition by pathogens. This may also explain why in striking contrast to the genetic code, glycans show widely divergent patterns between taxa. On the other hand, glycans have also been coopted for myriad intrinsic functions, which can vary in their importance for organismal survival. In keeping with these considerations, a significant percentage of the genes in the typical genome are dedicated to the generation and/or turnover of glycans. Among eukaryotes, the Golgi is the subcellular organelle that serves to generate much of the diversity of cell surface glycans, carrying out various glycan modifications of glycoconjugates that transit through the Golgi, en route to the cell surface or extracellular destinations. Here I present an overview of general considerations regarding the selective forces shaping evolution of the Golgi glycosylation machinery, and then briefly discuss the common types of variations seen in each major class of glycans, finally focusing on sialic acids as an extreme example of evolutionary glycan diversity generated by the Golgi. Future studies need to address both the phylogenetic diversity the Golgi and the molecular mechanisms for its rapid responses to intrinsic and environmental stimuli.

Every eukaryotic cell is covered with a dense and complex array of glycans, which also feature prominently in extracellular matrix and secreted soluble molecules (Varki and Sharon 2009). The bulk of these glycans are synthesized by the Golgi (Emr et al. 2009; Varki et al. 2009a), and a significant percentage of the genes in the typical eukaryotic genome are dedicated to these glycosylation functions (Henrissat et al. 2009). Other articles on this topic address the primary biochemical pathways involved (Stanley 2011), transport mechanisms that expose the transiting cargo to the ordered sequence of glycosidases and glycosyltransferases that carry out the modifications to the bound oligosaccharides (Glick and Luini

A. Varki

2011; Banfield 2011), genetic diseases that can affect the Golgi glycosylation system (Freeze and Ng 2011), and the evolution and diversity of Golgi apparatus structure in eukaryotes (Klute et al. 2011). Here I consider the Golgi glycosylation machinery from an evolutionary perspective, asking why it is universal to eukaryotes, how it has changed and been modified over evolutionary time, and why this machinery has diverged so much between different taxa. Some figures in this article are from *Essentials of Glycobiology*, the first open access textbook (http://www.ncbi.nlm.nih.gov/bookshelf/br. fcgi?book=glyco2). The reader is referred to the relevant articles in *Essentials of Glycobiology* (Varki and Sharon 2009; Varki et al. 2009a; Freeze and Elbein 2009; Varki and Lowe 2009) for further information about some of the points made here, as well as to some other articles in the textbook on related topics (Stanley et al. 2009; Brockhausen et al. 2009; Schnaar et al. 2009; Ferguson et al. 2009; Freeze and Haltiwanger 2009; Stanley and Cummings 2009; Varki and Schauer 2009; Hascall and Esko 2009; Esko et al. 2009).

CELL SURFACE GLYCOSYLATION IS UNIVERSAL TO LIVING CELLS

In striking contrast to mathematics, physics, and chemistry, biology has very few absolute laws. This is because biology is the product of undirected evolutionary processes, which cannot generate universal designs, and indeed often generates "unintelligent design" (Avise 2010). Aside from a genetic code (DNA contained within a genome) (Crick 1970) and diverse RNAs comprising a transcriptome (Wilusz et al. 2009) (Fig. 1), all living cells have structural and functional proteins (contained in the proteome) (Yates et al. 2009), key metabolites that mediate energy flux and signaling (Gehlenborg et al. 2010; Vinayavekhin et al. 2010; Wallace 2010) (contained in the metabolome); and lipid-based membranes (a part of the lipome, or lipidome) (Shevchenko and Simons 2010) (Fig. 1). In addition to these apparent universals, it is a remarkable rule that all living cells in nature are covered (see Fig. 2 for an example) with a dense and complex array of glycans (a part of the glycome) (An et al. 2009; North et al. 2009; Alley and Novotny 2010; Cohen and Varki 2010). Even nonnucleated enveloped viruses are coated with glycans derived from the cells from which they bud. More than 3 billion years of evolutionary "experimentation" has apparently failed to select for any cell devoid of surface glycans. Thus, this feature of having a cell surface "glycocalyx" seems to be an absolute requirement for survival as a cellular life form. However, although the existence of a glycocalyx seems as essential to life as having a genetic code, there is a striking difference with regard to diversity
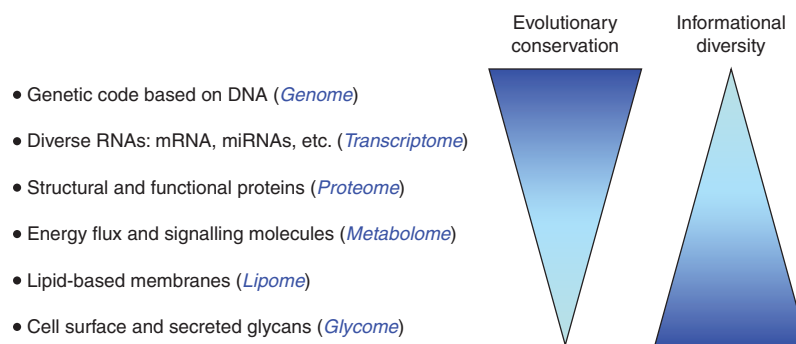
Figure 1. Universal characteristics of all living cells. As indicated in the figure and discussed in the text, cell surface glycosylation is among the key features that are universal to all living cells. However, in contrast to the genetic code, the degree of evolutionary conservation of glycans between taxa is the least. The likely reasons for this difference are discussed in the text. In eukaryotes, most of the resulting structural diversity of cell surface glycosylation is generated by the Golgi apparatus.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a005462

Figure 2. An example of the thickness of the cell surface glycocalyx. Shown is an electron microscopic overview of a rat left ventricular myocardial capillary stained with Alcian blue 8GX. Alcian Blue will react with acidic glycans such as glycosaminoglycans and sialic acids, as well as the nonGolgi derived glycan hyaluronan (bar = 1 μm). It can be seen that the glycocalyx on the luminal surface of the vessel wall is almost as thick as the underlying endothelial cell, which synthesizes it (Used with permission from van den Berg et al. 2003).

and conservation (Fig. 1). As a rule of thumb, evolutionary conservation is roughly in the order genetic code > RNA sequences > primary protein sequence > metabolic pathways > cellular lipid composition > surface glycan structures. The reverse order seems to apply to structural diversity (Fig. 1). Particularly striking is the remarkable structural diversity of cell surface glycans, as well as those of the extracellular matrix and secreted molecules, most of which are also glycoconjugates. In addition, when it comes to eukaryotic cells, most of these diverse forms of glycosylation are generated by the Golgi (Varki et al. 2009a).

WHY IS CELL SURFACE GLYCOSYLATION APPARENTLY UNIVERSAL TO ALL LIVING CELLS IN NATURE?

If a cell could exist in nature without cell surface glycosylation, surely 3 billion years of evolution would have selected for one that is covered only with proteins and lipids, featuring minimal or no glycans on it? The fact that this has apparently never happened indicates very strong

evolutionary selective forces that favor the presence of dense and complex glycocalyx on cell surfaces. Although glycosylation is required for many important biological processes that are intrinsic to multicellular organisms (Varki and Lowe 2009), it is unlikely that this was the only evolutionary selection force, as all unicellular organisms also have cell surface glycosylation. The more prominent universal evolutionary force has likely been interactions between organisms, in pathogen or symbiont relationships (Varki 1993; Drickamer and Taylor 1998; Gagneux and Varki 1999; Varki 2006; Bishop and Gagneux 2007; Varki et al. 2009b).

If cells were coated only with proteins, it would be more difficult for an organism to escape from a pathogen that evolves to bind to one of these proteins (Gagneux and Varki 1999; Bishop and Gagneux 2007). The reason for this is that random amino acid changes are not usually well tolerated by most proteins. The majority of such changes are likely to affect folding and/or stability, rendering the protein partially or completely dysfunctional. Thus, it would be difficult for an organism besieged by a protein-recognizing pathogen to rapidly evolve tolerable changes to the structure of the target protein, without losing its functions. In keeping with this, the great majority of pathogens initiate contacts with host cells via glycan targets (for examples, see Ofek and Sharon 1990; Ilver et al. 2003; Esko and Sharon 2009). Indeed, the minority of pathogens that bind only to surface proteins tend to be relatively successful in their selected niche. In contrast to proteins, glycans have many features that make it more difficult for a pathogen to gain and maintain a hold. First, most glycans do not have rigid structures and instead have significant freedom of motion in aqueous solution (Poppe and van 1991; Fadda and Woods 2010; Woods and Tessier 2010). Thus, the common analogy of a cell surface being like a forest is actually incorrect—it is more like a coral reef or kelp bed on the ocean floor (a suggestion by Pascal Gagneux). A second likely reason for the universality of surface glycosylation is that most intrinsic glycan functions are mediated not by a single absolutely required

A. Varki

sequence, but by an ensemble of glycan structures, spanning a continuum that can carry out the functions (Varki and Lowe 2009). In other words, glycan functions tend to be more "analog" than "digital." Thus, it would be easier for a host to escape a pathogen by partially or subtly changing its glycosylation, without drastically altering its own intrinsic functions. Thirdly, given the complex chemistry of glycans, there are many more structural variations possible in glycan evolution, such as change of a glycan linkage, a change in branching, variations in glycan modification, a change of a terminal glycan, or even altered clustering of an existing glycan (Gagneux and Varki 1999). These are just some of the likely reasons why glycans dominate the cell surface biology of all living organisms, even when showing wide variations in structural patterns between

species, and even within species. Of course, all of these considerations also apply to symbiotic relationships, an aspect much less studied, compared to pathogenic interactions. Figure 3 considers the interplay between various forces driving the evolutionary diversification of glycans (Varki 2006), as addressed in several sections that follow.

MOST PATHOGENS AND SYMBIONTS RECOGNIZE HOST CELL GLYCANS AS A FIRST STEP OF ATTACHMENT

As discussed above, the majority of microorganisms use glycans as primary or secondary targets for recognition of host cells. This is not surprising, given the universality of glycosylation on the cell surface. The generally more rapid evolution of the pathogens in comparison to slower
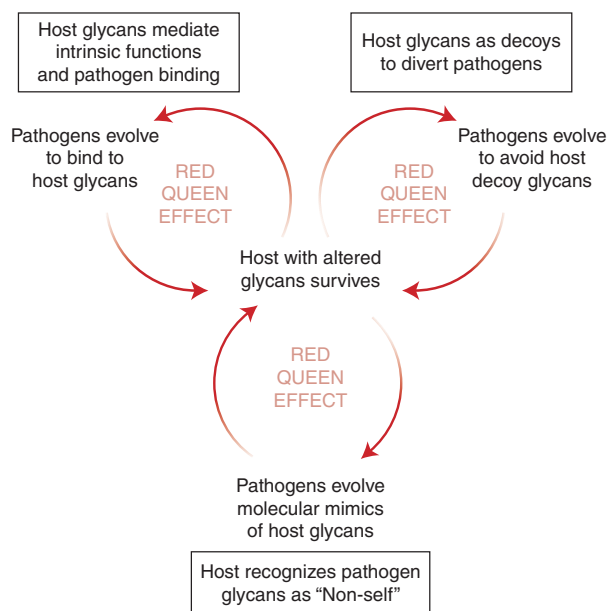
Figure 3. Forces driving the evolutionary diversification of glycans. Each arrowed circle represents a potential vicious cycle, driven by a "Red Queen" effect, in which hosts are constantly trying to evade the more rapidly evolving pathogens that infect them. Hosts require glycans for critical cellular functions but must constantly change them to evade glycan-binding pathogens, yet without impairing their own survival. Hosts also produce soluble glycans such as mucins, which act as decoys to divert pathogens from cell surfaces; but pathogens are constantly adjusting to these defenses. Hosts recognize pathogen-specific glycans as markers of "non-self," but pathogens can modify their glycans to more closely mimic host glycans. There are also possible secondary Red Queen effects involving host glycan binding proteins that recognize "self." In each of these cycles, hosts with altered glycans that can still carry out adequate cellular functions are most likely to survive. (Modified with permission from Varki [2006].)

reproducing hosts then results in survival of the "fittest" pathogen, explaining why such glycan-binding phenomena are often exquisitely specific, and of relatively or absolutely high affinity (because of multivalent avidity or univalent good fit, respectively). Indeed, pathogen and symbiont glycan binding is often of higher affinity and specificity than glycan binding phenomena mediating intrinsic functions.

HOSTS MUST CONSTANTLY CHANGE THEIR GLYCANS, TO LIMIT PATHOGEN INTERACTIONS

Meanwhile, hosts in turn must constantly change their glycans to survive, by limiting pathogen interactions. In this regard, Golgi apparatus-mediated static or dynamic cell-type specific variations in glycosylation within an organism could play a key role in limiting the success of a pathogen. This is currently a poorly studied area. However, given that pathogens evolve more rapidly, hosts face a "Red Queen" situation (Van Valen 1974) wherein they must "keep running to stay in place" (Gagneux and Varki 1999; Varki 2006). Thus, dynamic changes mediated by the Golgi might not be sufficient, and permanent species-specific changes may be needed, mediated by selection of genetic changes that ultimately affect Golgi structure and function (Bishop and Gagneux 2007).

PATHOGENS OR SYMBIONTS MUST ALSO EVOLVE RAPIDLY, TO MAINTAIN HOST GLYCAN RECOGNITION

As discussed above, the rapidity of multiplication of most pathogens enhances the rate at which they can continually optimize their binding sites for host target glycans. This likely explains the exquisite specificity and relatively high affinity of such interactions. Of course, this variability also represents the "Achilles Heel" of such pathogens, because increasing specificity of binding provides an opportunity for hosts to evade recognition by slightly changing the glycan binding site, without altering its intrinsic functions within the host (Gagneux and Varki 1999). Of course, the pathogens

usually win such an "arms race" in the end. Far less is known about how symbiont binding of glycans evolves, but the issues involved are likely to be relatively similar.

MANY HOST GLYCANS HAVE IMPORTANT OR ESSENTIAL INTRINSIC FUNCTIONS

Although pathogen and/or symbiont relationships may have primarily driven the universality and rapid divergence of eukaryotic cell surface glycans, there have also been hundreds of millions of years of opportunity for hosts to use these glycans to evolve useful intrinsic functions. Indeed, given the high density, complexity, and ubiquity of cell surface and extracellular matrix glycans, it is not surprising that many intercellular interactions in eukaryotes are mediated or modulated by the structure and variations of glycans (Varki and Lowe 2009). These intrinsic glycan functions will not be discussed in detail here, but broadly include: protective, stabilizing, organizational, and barrier functions; modulating the interaction of proteins or lipids with one another; acting as a protective storage depot for biologically important molecules; preventing diffusion of critical factors (sometimes generating gradients), allowing release of factors under specific conditions; acting as specific ligands for certain intrinsic lectins; involvement in cell–cell recognition and cell–matrix interactions, and so on (Varki and Lowe 2009). It should also be kept in mind that once a new glycan or glycan modification is expressed in a multicellular eukaryote, many distinct functions could evolve in different tissues and at different times in development. In addition, if any of these functions enhance survival and reproductive fitness of that species, the mechanisms involved would tend to become conserved, to the extent possible under prevailing circumstances.

HOSTS MAY DISCARD NONCRITICAL GLYCANS TO ALLOW SURVIVAL OF THE SPECIES

Some intrinsic glycan functions are critical enough that they cannot be discarded to avoid a pathogen, and still allow survival of a species.

A. Varki

However, the Golgi can assist the host response by dynamically changing the presentation of the glycan on the cell surface or the underlying glycosylation pattern, as happens, for example, during the "acute phase response" of hepatocytes to inflammation, which changes the glycosylation patterns of many plasma proteins (Baumann and Gauldie 1994). Occasionally a host species may be able to discard the Golgi machinery for generation of a specific noncritical glycan, to allow survival from pathogen assault. Of course in this situation "noncritical" is really a relative statement; as such a host is faced with multiple intrinsic consequences of the change in glycosylation. However, as long as it is not markedly detrimental to reproductive fitness, such a change may be the only way out of an evolutionary corner created by extreme selection pressure from a pathogen. Of course, depending on the length of time that the selection pressure is applied and how detrimental it is, one could end up with a glycan population polymorphism, rather than a permanent fixed change in the host species (Bishop and Gagneux 2007). Many examples of both situations can be found in nature, such as the ABO blood group polymorphisms in humans (Milland and Sandrin 2006), the elimination of the α-Gal epitope in Old World monkeys (Galili 2005), and loss of the sialic acid Neu5Gc in humans (Varki 2010).

HOSTS CAN USE FOREIGN GLYCANS ON PATHOGENS FOR IMMUNE RECOGNITION

It is well known that the innate immune system has multiple receptors that recognize non-self glycans and trigger immune responses (Linehan et al. 2000; Takeuchi and Akira 2010). Indeed, prior to the emergence of the adaptive immune system ∼500 million years ago, this mechanism formed the major component of the immune response against foreign invaders. Thus, once a species has discarded or substantially modified a preexisting glycan, the opportunity then arises to use such a glycan as a flag to recognize foreign invaders (Bishop and Gagneux 2007). This of course requires the evolution of binding proteins within the host that recognize the foreign glycan, to mediate immune responses. It

is unclear how long such a process might take in evolutionary time. Meanwhile, the somatic genetic diversity that can be engendered by the humoral immune response allows antibodies to provide the earliest such responses, in animals with an adaptive immune system. Of course, the adaptive immune response can sometimes do more harm than good (Hedrick 2004).

PATHOGENS CAN EVOLVE MOLECULAR MIMICS OF HOST GLYCANS

Although hosts recognize pathogen glycans as being foreign (using specific innate immune receptors and/or antibodies), some pathogens can evade this outcome, by either terminating their glycans with a host-like structure, or going even further, to generate complete molecular mimics of the glycan structures generated by the host Golgi apparatus (Harvey et al. 2001; Tsai 2001; Coyne et al. 2005; Lewis et al. 2009). It is truly remarkable how precise such molecular mimicry can be, such as the almost perfect mimicking of complex vertebrate ganglioside glycans by *Campylobacter jejuni* (Yuki et al. 2004), or of sialylated N-glycan antennae by Group B *Streptococcus* (Wessels et al. 1989). The initial assumption was that this mimicry was accomplished by bacterial "stealing" of the relevant host genes that would normally encode Golgi enzymes. Remarkably, in every case elucidated to date this mimicry turns out to have actually occurred by convergent evolution, involving evolutionary reinventions of almost all the necessary components. To my knowledge to date no claim of vertebrate-to-bacterial transfer of genetic material has held up to rigorous scrutiny. Overall, this suggests that pathogens are under very strong selection pressure to independently evolve these molecular mimics of the host Golgi glycan products.

HOSTS CAN USE GLYCANS AS MOLECULAR DECOYS TO EVADE PATHOGENS

Pathogens that bind host glycans must recognize them at the correct cell surface, to initiate invasion. Hosts can protect themselves by presenting molecular glycan decoys such as soluble mucins,

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a005462

circulating plasma proteins or nonnucleated red blood cells, providing alternate binding sites for the pathogens, and diverting them away from their true cellular targets (Gagneux and Varki 1999). Again, it is the Golgi of the relevant cells that must generate the decoys. As one might imagine, such decoys are also undergoing rapid evolution, to adjust to the suite of pathogens that a given species is exposed to. This may explain the extreme variations in types of outer sequences of Golgi-derived glycans seen on potential decoys such as mucins, plasma proteins, and red blood cells. The Golgi can also assist in a host response by rapidly changing the glycan structures on such decoys, such as the "acute phase response" discussed above, (Baumann and Gauldie 1994), or in the mucosal secretory response to infections (Voynow et al. 2006).

ROLE OF THE GOLGI IN MEDIATING RED QUEEN EFFECTS IN HOST GLYCAN EVOLUTION

Although most prokaryotic glycans tend to be relatively simple in their structure (often with repeating units), eukaryotic glycans tend to have vast degrees of complexity and heterogeneity. Aside from genome size, the prime reason for this difference is the existence of the Golgi in eukaryotes. This perhaps also explains why the Golgi glycosylation machinery has evolved and diverged so much over time. Indeed, this machinery provides a mechanism by which glycan changes can occur rapidly, without changing the genome of the host. Given the nontemplate driven mechanisms of Golgi glycosylation in which multiple enzymes compete for multiple substrates over short period of time, marked glycan changes can be generated on the cell surface simply by changing the presence or localization of one or two glycosyltransferases or nucleotide sugar donors (Nilsson et al. 2009; Varki et al. 2009a; Tu and Banfield 2010). Indeed, those who work in biotechnology and biopharma know that even slight changes in the medium of their high production fermenter facilities can result in marked changes in glycosylation of their products (Hooker et al. 1995; Rivinoja et al. 2009).

## PRINCIPLES UNDERLYING GOLGI-DERIVED GLYCAN VARIATION IN NATURE

Having provided an overview of the evolutionary backdrop against which the Golgi evolved, we can now briefly consider the major classes of glycans that are generated or modified in the animal Golgi apparatus, and examples of variations that have occurred. The fact that evolution is characterized by increasing glycan complexity over time is not surprising, because life could only begin in simple forms, and complexity is thus the inevitable outcome of many evolutionary processes. However, the concept of "higher" and "lower forms" is no longer a valid way of thinking about evolution. This actually harks back to the old concept of a "scala naturae" or "Great Chain of Being" leading upward to humans. Rather, we should consider evolution as simply having occurred via descent by various selection mechanisms in the background of neutral drift. In this regard, a general rule is that the core structures of most glycan types tend to be conserved amongst various taxa (Fig. 3). Thereafter, within each glycan class one sees varying degrees of modifications. It appears that this diversification is particularly marked in the *Deuterostome* lineage of animals (so-called "higher" invertebrates and vertebrates) (Freeze et al. 2009). Some classes of glycans such as sialic acids and hyaluronan are even restricted to this clade. On the other hand, other complex multicellular organisms such as flowering plants have also evolved complex Golgi glycan modification machineries (Etzler and Mohnen 2009; Gomord et al. 2010). Considered below are some examples of evolutionary variations in specific classes of glycans generated by the Golgi apparatus.

N- and *O*-linked Glycans

The core pathways of initial *N*-linked glycosylation (Fig. 4, left lower) are well conserved in all eukaryotes (Stanley et al. 2009). These steps mostly take place in the endoplasmic reticulum. Thereafter, processing and rebuilding of the N-glycans in sequential Golgi cisternae generates a diversity of possible structural outcomes, which can vary widely between different species

A. Varki
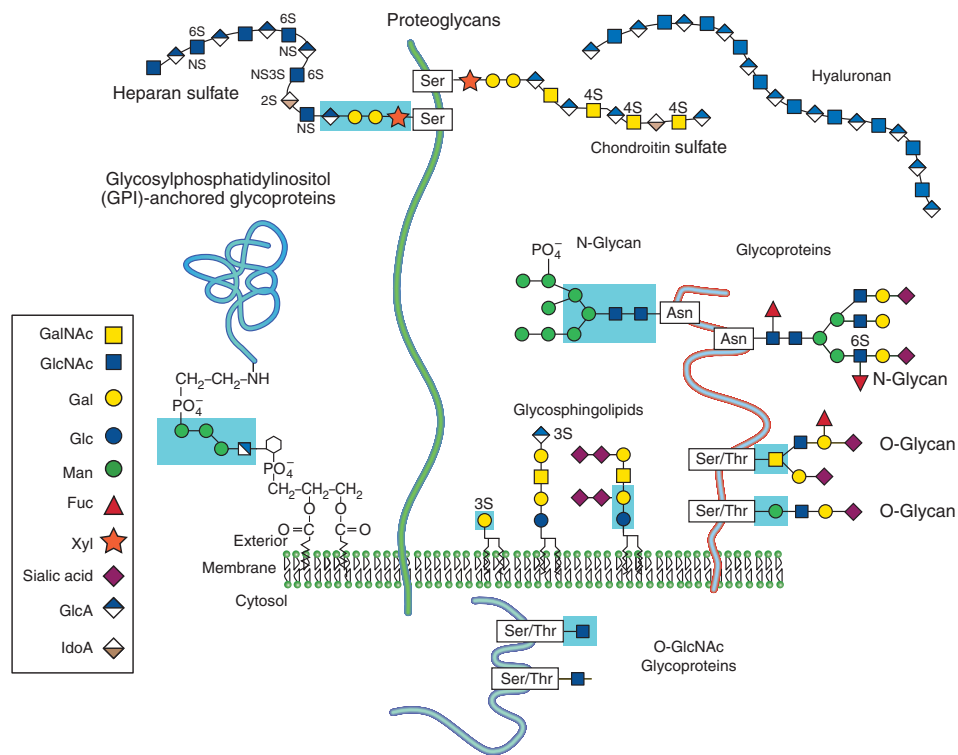
Figure 4. Common classes of animal glycans. The major classes of animal glycans are shown, with an emphasis on typical vertebrate sugar chains. The shaded blue boxes indicate the core glycan sequences that are conserved in most or all taxa that express such glycan classes. The monosaccharide abbreviations are GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; Gal, Galactose; Glc, Glucose; Man, Mannose; Fuc, Fucose; Xyl, Xylose; GlcA, Glucuronic acid; and, IdoA, Iduronic acid. (Reproduced with permission from Varki and Sharon [2009].)

(Fig. 4) and to a lesser extent between cell types within a given species (Dennis et al. 2009; Varki et al. 2009b). With the common kinds of O-glycans, the core glycosylation event is initiated mainly within the Golgi, and then undergoes various extension and branching reactions to yield final products. However, even the core monosaccharides involved in the initial O-linkage to polypeptides can vary amongst taxa (Brockhausen et al. 2009).

Glycosphingolipids

These complex glycolipids share a common core ceramide sphingolipid component, which can vary in terms of the *N*-acyl chain structure. The commonest type of glycosphingolipid core structure in vertebrates is glucosylceramide,

which is generated on the cytosolic leaflet of ER and Golgi membranes and then flipped over into the lumen of these organelles to be further extended in a variety of ways. The second most common variation in vertebrates is galactosylceramide. Although there is conservation of such classes of core glycosphingolipid glycans within related species, variations are seen in divergent clades, and there can be much complexity in the outer sequences (Schnaar et al. 2009).

Outer Extensions of Glycoprotein and Glycolipid Glycans

Once N-glycans have been trimmed in the ER and early Golgi, and O-glycan and glycosphingolipids core structures have been generated within the Golgi apparatus, each can be

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a005462

modified with outer extensions, which can be highly similar among these three classes of glycans (Stanley and Cummings 2009). In several instances, the same glycosyltransferases can actually be competing for extending the termini of these three types of glycan chains. One can imagine that the extent of this competition is affected by many factors, including the rate of transport of the molecules through the Golgi, their relationship to the each other and to the glycosyltransferases, and the availability of nucleotide sugar donors (Caffaro and Hirschberg 2006; Varki et al. 2009a). This is a remarkable example of the functions of the Golgi in generating glycan diversity. Very recently, it has been shown that altering glycosylation in the terminal part of the Golgi can affect earlier steps, and even induce an ER Stress response (Xu et al. 2010).

Glycosaminoglycans

This class of typically linear acidic sugar chains are initiated within the Golgi apparatus by a sequence of highly conserved monosaccharides, followed by divergence in the types of outer chains, which can then be extensively modified (Esko et al. 2009). In contrast to the glycan classes mentioned above, glycosaminoglycan core sequences tend to be relatively conserved, with variations primarily seen in the nature of the modifications on the glycans, such as sulfate esters. However, further studies may well show some relatively species-specific modifications. A major exception is hyaluronan, an unmodified linear acidic glycan chain that is only found in the *Deuterostome* lineages of animals, and in some bacteria that apparently make the structure via convergent evolution. Hyaluronan is not synthesized in the Golgi apparatus, but is rather extruded directly into the extracellular space from plasma membrane hyaluronan synthases (Hascall and Esko 2009).

Glycophospholipid Anchors

These phosphatidylinositol-based glycolipid anchors for a subset of cell surface proteins have been highly conserved in their core regions

for a long time in evolution, although they show some variations in different taxa (Ferguson et al. 2009). Even the phosphatidylinositol component is not invariant, as yeast can build such an anchor on a ceramide lipid tail. Other major variations are in the glycan side chains that extend from the common core structure. It is unclear what exactly are the functions of these side chain modifications, which are generated in the Golgi apparatus. The fact that some pathogens do bind to the GPI anchors suggests that this may be a mechanism of evasion.

Sialic Acids

Sialic acids are the outermost monosaccharides on most *N*-linked, *O*-linked and glycosphingolipid glycan chains on the cell surface and secreted glycoconjugates of animals of the *Deuterostome* lineage (Varki and Schauer 2009). From the evolutionary perspective, it appears that following the emergence of these core glycan structures, various other monosaccharides are then added sequentially in different taxa, finally terminating in the most complex monosaccharide, sialic acid. It is rare to find other monosaccharides extending beyond sialic acids (other than additional sialic acids, as in polysialic acids). Instead of being capped by other sugars, sialic acids instead become highly modified with more than 50 variants, which can be presented in many different kinds of linkages (Fig. 5) (Harduin-Lepers et al. 2005; Varki and Schauer 2009; Cohen and Varki 2010). Given their location and their frequent contact with pathogens and symbionts as well as their many intrinsic functions, it is not surprising that there are also substantial sialic acids variations within and between species. Again, it is the Golgi that is responsible for generating this remarkable diversity in the sialome, utilizing more than 20 known sialyltransferases, and a variety of sialic acid modifying enzymes which mostly remain to be identified.

CONCLUDING REMARKS

Much of the glycan density, complexity, and diversity found on eukaryotic cell surface, extracellular matrix and secreted molecules is
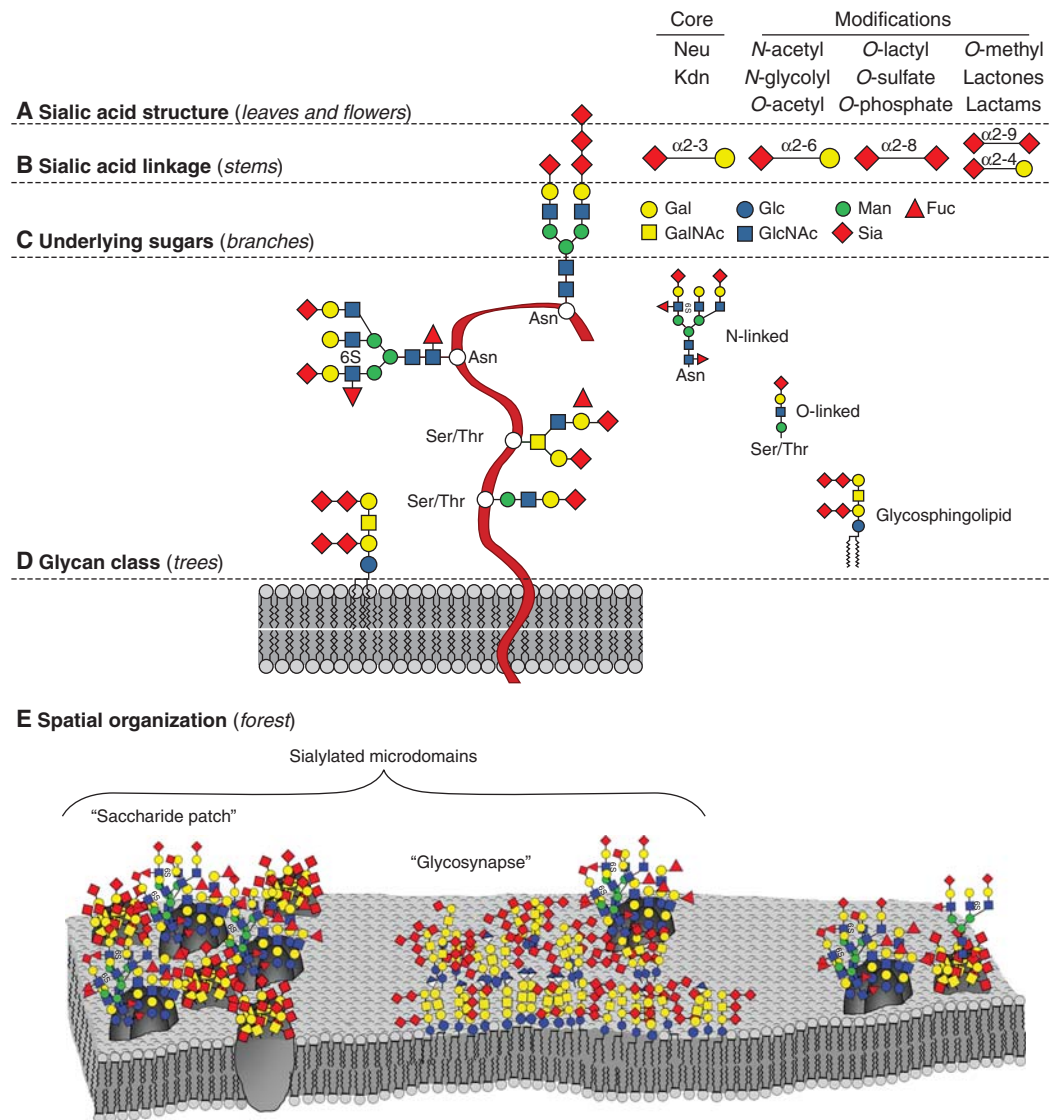
A. Varki
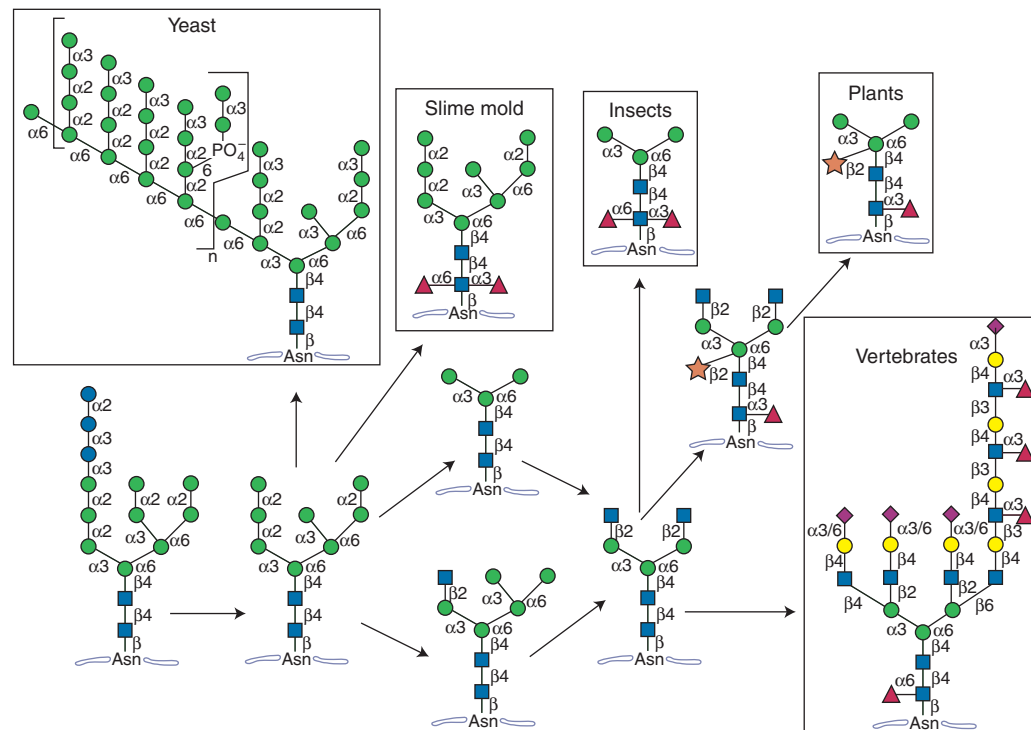
Figure 5. Dominant pathways of *N*-glycan processing in various taxa. See text for further discussion. Monosaccharide abbreviations are as in Figure 4. (Reproduced with permission from Varki et al. [2009b].)

generated by the glycosylation functions of the Golgi. This evolutionarily conserved feature of eukaryotic cells is able to generate a wide diversity of glycans, ranging from rapid dynamic changes in the Golgi of individual cells, all the way to genetic changes in the glycosylation machinery of the Golgi, generating intraspecies polymorphisms or interspecies differences. In addition, the Golgi is ideally suited to handle such tasks, because of its nontemplate-driven, assembly line-like mechanisms for generating glycan diversity. The likely evolutionary forces behind these phenomena include pathogen exploitation of glycans, as well as intrinsic functions of glycans, both of which can be critical for survival. Much more work needs to be performed, to fully understand the structural and functional evolution of the Golgi, under such evolutionary selection pressures. Other aspects deserving intensive study are the molecular mechanisms by which dynamic changes in the

Golgi structure and function occur, when a cell is faced with environmental changes such as injury, pH variation, inflammation, infectious agents, and so on.

In the future, Golgi-derived cell surface glycans also need to be studied at different hierarchical levels of complexity. In a recent review (Cohen and Varki 2010), we defined five such levels of complexity for just one subclass of the glycome, (i.e., the sialome) (See Fig. 6). In analogy to the canopy of a forest, the sialome covers the cell with diverse array of sialylated structures (Fig. 5). This complexity includes modifications of sialic acid core structures (the leaves and flowers), the possible linkages to the underlying monosaccharide (the stems), the identity and precise arrangement of the underlying glycans (the branches), the detailed structural attributes of underlying glycans (the trees), and last but not least, the spatial organization of the sialoglycans in relation to

Figure 6. Diversity in the sialic acids. All known sialic acids have a common nine-carbon backbone attached in the α configuration to the underlying sugar chain (Varki and Schauer 2009). The sialome can be analyzed at the following complexity levels: (A) Sia core and core modifications: esterification (with various groups), o-methylation, lactonization or lactamization yielding >50 different structures, (B) Linkage to the underlying sugar (three major and many minor linkages), (C) Identity and arrangement of the underlying sugars that can also be further modified by fucosylation or sulfation, (D) Glycan class (N-linked, O-linked or glycosphingolipids), (E) Spatial organization of the Sia in sialylated microdomains (including glycosynapses and clustered saccharide patches), and higher levels of the cellular and organismal milieu. Sia, Sialic acid; other monosaccharide abbreviations are as in Figure 4. (Reproduced with permission from Cohen and Varki [2010].)

A. Varki

components of the intact cell surface (the forest). As shown in an initial study of red cell surface sialoglycan display (Cohen et al. 2009), understanding the full complexity of the Golgi glycosylation requires combined analyses at all these levels—in other words, the glycome is far more than the sum of its parts.

ACKNOWLEDGMENTS

REFERENCES

Alley WR, Novotny MV. 2010. Glycomic analysis of sialic acid linkages in glycans derived from blood serum glycoproteins. *J Proteome Res* **9:** 3062–3072.

An HJ, Kronewitter SR, de Leoz ML, Lebrilla CB. 2009. Glycomics and disease markers. *Curr Opin Chem Biol* **5:** 601–607.

Avise JC. 2010. Colloquium paper: footprints of nonsentient design inside the human genome. *Proc Natl Acad Sci* **107 Suppl 2:** 8969–8976.

Banfield DK. 2011. Mechanisms of protein retention in the golgi. *Cold Spring Harb Perspect Biol* doi:10.1101/cshperspect.a005264.

Baumann H, Gauldie J. 1994. The acute phase response. *Immunol Today* **15:** 74–80.

Bishop JR, Gagneux P. 2007. Evolution of carbohydrate antigens–microbial forces shaping host glycomes? *Glycobiology* **17:** 23R–34R.

Brockhausen I, Schachter H, Stanley P. 2009. O-GalNAc glycans. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 115–128. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Caffaro CE, Hirschberg CB. 2006. Nucleotide sugar transporters of the golgi apparatus: from basic science to diseases. *Acc Chem Res* **39:** 805–812.

Cohen M, Varki A. 2010. The sialome-far more than the sum of its parts. *OMICS* **14:** 455–464.

Cohen M, Hurtado-Ziola N, Varki A. 2009. ABO blood group glycans modulate sialic acid recognition on erythrocytes. *Blood* **114:** 3668–3676.

Coyne MJ, Reinap B, Lee MM, Comstock LE. 2005. Human symbionts use a host-like pathway for surface fucosylation. *Science* **307:** 1778–1781.

Crick F. 1970. Central dogma of molecular biology. *Nature* **227:** 561–563.

Dennis JW, Nabi IR, Demetriou M. 2009. Metabolism, cell surface organization, and disease. *Cell* **139:** 1229–1241.

Drickamer K, Taylor ME. 1998. Evolving views of protein glycosylation. *Trends Biochem Sci* **23:** 321–324.

Emr S, Glick BS, Linstedt AD, Lippincott-Schwartz J, Luini A, Malhotra V, Marsh BJ, Nakano A, Pfeffer SR, Rabouille C, et al. 2009. Journeys through the Golgi–taking stock in a new era. *J Cell Biol* **187:** 449–453.

Esko JD, Sharon N. 2009. Microbial Lectins: Hemagglutinins, Adhesins, and Toxins. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 489–500. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Esko JD, Kimata K, Lindahl U. 2009. Proteoglycans and sulfated glycosaminoglycans. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 229–248. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Etzler ME, Mohnen D. 2009. Viridiplantae. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 321–332. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Fadda E, Woods RJ. 2010. Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov Today* **15:** 596–609.

Ferguson MAJ, Kinoshita aT, Hart GW. 2009. Glycosylphosphatidylinositol Anchors. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 143–162. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Freeze HH, Vacquier VD, Esko JD. 2009. Deuterostomes. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 363–373. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Freeze HH, Elbein AD. 2009. Glycosylation precursors. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 363–373. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Freeze HH, Haltiwanger RS. 2009. Other classes of ER/Golgi-derived glycans. In *Essentials of Glycobiology* (eds. Varki A., Cummings R.D., Esko J.D., Freeze H.H., Stanley P., Bertozzi C.R., Hart G.W., Etzler M.E.), pp. 363–373. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Freeze HH, Ng BG. 2011. Golgi glycosylation and human inherited diseases. *Cold Spring Harb Perspect Biol* doi:10.1101/cshperspect.a005371.

Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9:** 747–755.

Galili U. 2005. The α-gal epitope and the anti-Gal antibody in xenotransplantation and in cancer immunotherapy. *Immunol Cell Biol* **83:** 674–686.

Glick BS, Luini A. 2011. Models for golgi traffic: a critical assessment. *Cold Spring Harb Perspect Biol* doi:10.1101/cshperspect.a005215.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a005462

Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, et al. 2010. Visualization of omics data for systems biology. *Nat Methods* **7:** S56–68.

Gomord V, Fitchette AC, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol J* **8:** 564–587.

Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. 2005. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology* **15:** 805–817.

Harvey HA, Swords WE, Apicella MA. 2001. The mimicry of human glycolipids and glycosphingolipids by the lipooligosaccharides of pathogenic neisseria and haemophilus. *J Autoimmun* **16:** 257–262.

Hascall V, Esko JD. 2009. Hyaluronan. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 219–228. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hedrick SM. 2004. The acquired immune system: a vantage from beneath. *Immunity* **21:** 607–615.

Henrissat B, Surolia A, Stanley P. 2009. A Genomic View of Glycobiology. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 89–100. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hooker AD, Goldman MH, Markham NH, James DC, Ison AP, Bull AT, Strange PG, Salmon I, Baines AJ, Jenkins N. 1995. N-glycans of recombinant human interferon-γ change during batch culture of Chinese hamster ovary cells. *Biotechnol Bioeng* **48:** 639–648.

Ilver D, Johansson P, Miller-Podraza H, Nyholm PG, Teneberg S, Karlsson KA. 2003. Bacterium-host protein-carbohydrate interactions. *Methods Enzymol* **363:** 134–157.

Klute MJ, Melancon P, Dacks JB. 2011. Evolution and diversity of the golgi apparatus. *Cold Spring Harb Perspect Biol* doi:10.1101/cshperspect.a007849.

Lewis AL, Desa N, Hansen EE, Knirel YA, Gordon JI, Gagneux P, Nizet V, Varki A. 2009. Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. *Proc Natl Acad Sci* **106:** 13552–13557.

Linehan SA, Martínez-Pomares L, Gordon S. 2000. Macrophage lectins in host defence. *Microbe Infect* **2:** 279–288.

Milland J, Sandrin MS. 2006. ABO blood group and related antigens, natural antibodies and transplantation. *Tissue Antigens* **68:** 459–466.

Nilsson T, Au CE, Bergeron JJ. 2009. Sorting out glycosylation enzymes in the Golgi apparatus. *FEBS Lett* **583:** 3764–3769.

North SJ, Hitchen PG, Haslam SM, Dell A. 2009. Mass spectrometry in the analysis of N-linked and O-linked glycans. *Curr Opin Struct Biol* **19:** 498–506.

Ofek I, Sharon N. 1990. Adhesins as lectins: Specificity and role in infection. *Curr Top Microbiol Immunol* **151:** 91–114.

Poppe L, van HH. 1991. Nuclear magnetic resonance of hydroxyl and amido protons of oligosaccharides in aqueous solution: Evidence for a strong intramolecular hydrogen bond in sialic acid residues. *J Am Chem Soc* **113:** 363–365.

Rivinoja A, Hassinen A, Kokkonen N, Kauppila A, Kellokumpu S. 2009. Elevated Golgi pH impairs terminal N-glycosylation by inducing mislocalization of Golgi glycosyltransferases. *J Cell Physiol* **220:** 144–154.

Schnaar RL, Suzuki A, Stanley P. 2009. Glycosphingolipids. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 129–142. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shevchenko A, Simons K. 2010. Lipidomics: coming to grips with lipid diversity. *Nat Rev Mol Cell Biol* **11:** 593–598.

Stanley P, Cummings RD. 2009. Structures Common to Different Glycans. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 175–198. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Stanley P, Schachter H, Taniguchi N. 2009. N-Glycans. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 101–114. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Stanley P. 2011. Golgi glycosylation. *Cold Spring Harb Perspect Biol* doi:10.1101/cshperspect.a005199.

Takeuchi O, Akira S. 2010. Pattern recognition receptors and inflammation. *Cell* **140:** 805–820.

Tsai CM. 2001. Molecular mimicry of host structures by lipooligosaccharides of Neisseria meningitidis: characterization of sialylated and nonsialylated lacto-N-neotetraose (Galβ1-4GlcNAcβ1-3Galβ1-4Glc) structures in lipooligosaccharides using monoclonal antibodies and specific lectins. *Adv Exp Med Biol* **491:** 525–542.

Tu L, Banfield DK. 2010. Localization of Golgi-resident glycosyltransferases. *Cell Mol Life Sci* **67:** 29–41.

van den Berg BM, Vink H, Spaan JA. 2003. The endothelial glycocalyx protects against myocardial edema. *Circ Res* **92:** 592–594.

Van Valen L. 1974. Two modes of evolution. *Nature* **252:** 298–300.

Varki A. 1993. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology* **3:** 97–130.

Varki A. 2006. Nothing in glycobiology makes sense, except in the light of evolution. *Cell* **126:** 841–845.

Varki A. 2010. Colloquium paper: uniquely human evolution of sialic acid genetics and biology. *Proc Natl Acad Sci* **107 Suppl 2:** 8939–8946.

Varki A, Lowe JB. 2009. Biological Roles of Glycans. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 75–88. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Varki A, Schauer R. 2009. Sialic Acids. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart,

A. Varki

M.E. Etzler), pp. 199–218. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Varki A, Sharon N. 2009. Historical Background and Overview. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 1–22. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Varki A, Esko JD, Colley KJ. 2009a. Cellular Organization of Glycosylation. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 37–46. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Varki A, Freeze HH, Gagneux P. 2009b. Evolution of Glycan Diversity. In *Essentials of Glycobiology* (eds. A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler), pp. 281–292. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Vinayavekhin N, Homan EA, Saghatelian A. 2010. Exploring disease through metabolomics. *ACS Chem Biol* **5:** 91–103.

Voynow JA, Gendler SJ, Rose MC. 2006. Regulation of mucin genes in chronic inflammatory airway diseases. *Am J Respir Cell Mol Biol* **34:** 661–665.

Wallace DC. 2010. Colloquium paper: bioenergetics, the origins of complexity, and the ascent of man. *Proc Natl Acad Sci* **107 Suppl 2:** 8947–8953.

Wessels MR, Rubens CE, Benedí V-J, Kasper DL. 1989. Definition of a bacterial virulence factor: Sialylation of the group B streptococcal capsule. *Proc Natl Acad Sci* **86:** 8983–8987.

Wilusz JE, Sunwoo H, Spector DL. 2009. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* **23:** 1494–1504.

Woods RJ, Tessier MB. 2010. Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr Opin Struct Biol* **5:** 575–583.

Xu YX, Liu L, Caffaro CE, Hirschberg CB. 2010. Inhibition of Golgi apparatus glycosylation causes endoplasmic reticulum stress and decreased protein synthesis. *J Biol Chem* **285:** 24600–24608.

Yates JR, Ruse CI, Nakorchevsky A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* **11:** 49–79.

Yuki N, Susuki K, Koga M, Nishimoto Y, Odaka M, Hirata K, Taguchi K, Miyatake T, Furukawa K, Kobata T, et al. 2004. Carbohydrate mimicry between human ganglioside GM1 and Campylobacter jejuni lipooligosaccharide causes Guillain-Barre syndrome. *Proc Natl Acad Sci* **101:** 11404–11409.

Cold Spring Harbor Perspectives in Biology

www.cshperspectives.org