Glycobiology vol. 9 no. 8 pp. 747–755, 1999

MINI REVIEW

Evolutionary considerations in relating oligosaccharide diversity to biological function

Pascal Gagneux and Ajit Varki[1]

Glycobiology Program and UCSD Cancer Center, Divisions of Hematology-Oncology and Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA 92093–0687, USA

Received on February 16, 1999; revised on April 2, 1999; accepted on April 5, 1999

[1]To whom correspondence should be addressed

The oligosaccharide chains (glycans) attached to cell surface and extracellular proteins and lipids are known to mediate many important biological roles. However, for many glycans, there are still no evident functions that are of obvious benefit to the organism that synthesizes them. There is also no clear explanation for the extreme complexity and diversity of glycans that can be found on a given glycoconjugate or cell type. Based on the limited information available about the scope and distribution of this diversity among taxonomic groups, it is difficult to see clear trends or patterns consistent with different evolutionary lineages. It appears that closely related species may not necessarily share close similarities in their glycan diversity, and that more derived species may have simpler as well as more complex structures. Intraspecies diversity can also be quite extensive, often without obvious functional relevance. We suggest one general explanation for these observations, that glycan diversification in complex multicellular organisms is driven by evolutionary selection pressures of both endogenous and exogenous origin. We argue that exogenous selection pressures mediated by viral and microbial pathogens and parasites that recognize glycans have played a more prominent role, favoring intra- and inter-species diversity. This also makes it difficult to appreciate and elucidate the specific endogenous roles of the glycans within the organism that synthesizes them.

Key words: oligosaccharide/glycan diversity/evolution/pathogens

Glycan diversity in nature

"Now every species of mammal and bird so far investigated has shown quite a surprising biochemical diversity by serological tests. The antigens concerned seem to be proteins to which polysaccharides are attached. We do not know their functions in the organism, though some of them seem to be part of the structure of the cell membrane. I wish to suggest that they may play a part in disease resistance, a particular race of bacteria or virus being adapted to individuals of a certain range of biochemical constitutions, while those of other constitutions are relatively resistant." (J.B.S.Haldane, 1949)

It is time to revisit Haldane's prescient words in the light of current knowledge of the structure, synthesis and function of the oligosaccharides (glycans) that are found attached to many cell surface and extracellular proteins and lipids. These abundant sugar chains are known to mediate many important biological phenomena (Rademacher *et al*., 1988; Paulson, 1989; Esko, 1991; Hart, 1992; Kobata, 1992; Lis and Sharon, 1993; Varki, 1993; Varki and Marth, 1995; Gahmberg and Tolvanen, 1996; Salmivirta *et al*., 1996) that include structural and physical roles, as well as specific recognition by lectin receptors of both endogenous and exogenous origin (Figure 1). A recent review (Drickamer and Taylor, 1998) pointed out that multiple forces must have driven the evolution of proteins that create and recognize glycans, and postulated that these forces were primarily the sequential development of valuable endogenous functions for these glycans. However, for many glycans that are found in a cell-type specific and developmentally regulated manner, there are still no obvious functions that are of known benefit to the organism that synthesizes them. In this regard, a fascinating and challenging issue is to find an explanation for the sheer complexity and diversity of glycans found on any given molecule or cell type (Rademacher *et al*., 1988; Hart, 1992; Kobata, 1992; Varki and Marth, 1995). The importance of this diversity is evident from conservation of the coding sequences of glycosyltransferases (the enzymes that assemble glycans), and from the deleterious consequences of genetic defects in several of them (Campbell *et al*., 1995; Kawagoe *et al*., 1996; Maly *et al*., 1996; Asano *et al*., 1997; Chui *et al*., 1997; Li *et al*., 1997; Tarutani *et al*., 1997; Hennet *et al*., 1998; Kornfeld, 1998; Ellies *et al*., 1998).

Since the enzymes involved in glycosylation are highly conserved, we can ask if the same glycoprotein carries the same type of glycosylation in taxonomically related species. There is relatively little data available concerning this issue. Comparing glycan structures among vertebrate species affords examples of either a high degree of conservation (e.g., plasma fibrinogen, Figure 2, minor species differences found in the type of sialic acids), or marked interspecies diversity (e.g., kidney γ-glutamyl transpeptidase; Figure 3). In the latter instance, there is not even an obvious relationship of structural diversification to the phylogenetic relationships amongst the species studied (the glycans on the mouse protein are more similar to those of the human than to those of the rat). Diversity in glycosylation is found at every level of biological organization, between species, within populations of the same species, and also among different molecules and cell types within the same organism. Marked changes in glycan structure can also occur during development, cellular activation, differentiation, cancer, and inflammation (Rademacher *et al*., 1988; Kobata, 1992; Varki, 1993; Varki and Marth, 1995; Maly *et al*., 1996; Tsuji, 1996). Here we discuss the significance of this remarkable diversity, mindful of the oft-repeated adage of Dobzhansky's that "nothing in biology makes sense, except in the light of evolution" (Dobzhnasky, 1973).

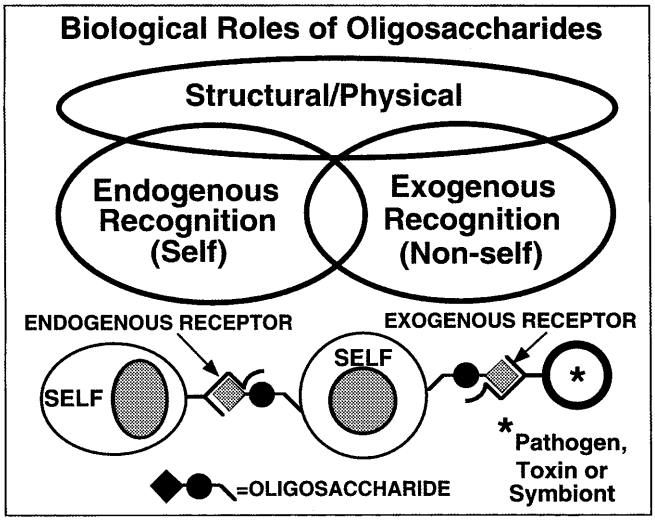
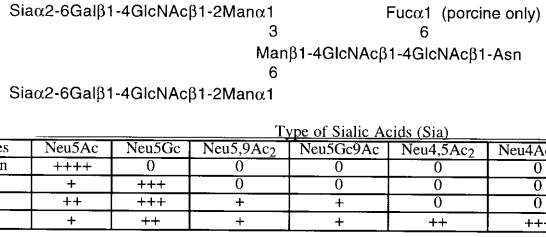


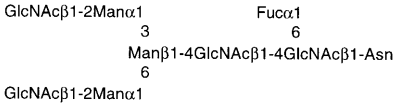
Fig. 1. Biological roles of oligosaccharides. The known biological roles of glycans are numerous and overlapping, but can be divided into two general groups: those based primarily upon the structural and physical effects of the molecules, and those involving recognition of specific oligosaccharides by cognate receptors. The latter falls into two general categories: self recognition, involving lectin receptors within the same organism; and non-self recognition, in which the receptors are mainly of microbial or parasitic origin (hemagglutinins, adhesins, toxins, etc.) but could also be involved in interactions with symbionts (e.g., gut flora). The question addressed in this mini review is: which category of glycan recognition is likely to be more common, based upon evolutionary considerations, and what does this predict regarding the significance of glycan diversity?



Type of Sialic Acids (Sia)

Species	Neu5Ac	Neu5Gc	Neu5,9Ac$_2$	Neu5Gc9Ac	Neu4,5Ac$_2$	Neu4Ac5Gc
Human	++++	0	0	0	0	0
Pig	+	+++	0	0	0	0
Cow	++	+++	+	+	0	0
Horse	+	++	+	+	++	+++

Fig. 2. Limited variation in the N-glycan structure of plasma fibrinogen amongst different species. The basic sialylated biantennary structure and the underlying linkages are conserved between species, expect for variations in the presence of a core fucose residue, and in the specific types of sialic acids found at the nonreducing terminus (Mizuochi *et al*., 1982; Damm *et al*., 1989; Debeire *et al*., 1985; Plummer *et al*., 1996).



Species	Terminal Sialylation	Poly-N-acetyl-Lactosamines	Outer α1-3 Fucose Residues	Tri/Tetra- antennary Branching
Human	++	++++	++++	0
Mouse	+	++	++++	+
Rat	++++	0	0	+++
Cow	++++	0	0	+++

Fig. 3. Extensive variation in N-glycan structures of kidney gamma-glutamyl transpeptidase among different species. Only the trimannosyl core structure common to complex N-glycans is conserved as shown, and many species-specific variations are seen in the rest of the molecules, as indicated (Yamashita *et al*., 1983a–c, 1985, 1986; Kobata, 1992). There does not even appear to be a relationship to phylogeny, e.g., the glycosylation of the mouse protein appears to be more similar to that of human than to that of the rat.

Glycan diversity can be specifically recognized by lectins of both endogenous and exogenous origin

The greatest diversity tends to be found among the outermost (nonreducing terminal) regions of glycans on cell surfaces and extracellular molecules (Rademacher *et al*., 1988; Hart, 1992; Kobata, 1992; Lis and Sharon, 1993; Varki and Marth, 1995; Tsuji, 1996). These regions of sugar chains are best positioned to mediate "self" recognition by endogenous carbohydrate binding proteins (lectins). Indeed, there is a growing list of vertebrate lectins that can recognize specific endogenous glycan structures and mediate major internal functions (Weis and Drickamer, 1996; Sharon, 1998). The preservation and diversification of such "self"-recognition certainly provides a partial explanation for the evolution of glycan diversity (Drickamer and Taylor, 1998). However, as indicated in Figure 1, most pathogens (broadly defined here as viruses, bacteria, protozoa, helminths, and fungi that infect multicellular organisms) also initiate and/or sustain interactions with vertebrate cells via lectins (hemagglutinins, fimbrial adhesins, and toxins) that can recognize glycan structures with exquisite specificity (Karlsson, 1995; Sharon, 1996). The same glycans at the outermost regions of host tissues and cells that mediate endogenous functions are also most prone to be used by pathogens for recognition and attachment. Such "non-self" recognition by pathogens makes the sugar chains particularly susceptible to the Red Queen Effect (Van Valen, 1974). This evolutionary concept is named for the Red Queen's comment to Alice in *Through the Looking Glass* that "it takes all the running you can do, to stay in the same place." It posits that multicellular organisms with long life cycles must constantly change, in order to survive the onslaught of potentially lethal pathogens which have much shorter life cycles and can thus evolve orders of magnitude faster. This provides one of the more plausible explanations for the predominance of sexual over asexual reproduction in vertebrates (Maynard Smith, 1978; Barton and Charlesworth, 1998; Partridge and Hurst, 1998), i.e., even though asexual reproduction is more efficient and less error-prone, sexual reproduction allows complex organisms to constantly generate diversity, evading rapidly evolving pathogens. Of course, besides pathogenic organisms, the symbionts that inhabit the external and internal epithelial surfaces of most vertebrates represent an additional exogenous factor influencing glycan diversity. For example, studies have shown dramatic differences in the fucosylation of glycans of the small intestinal epithelia of mice lacking the normal symbionts (Bry *et al*., 1996).

Keeping pace with pathogens: how much do we owe to the Red Queen Effect?

Since most pathogens and the toxins they produce bind to specific sugar sequences to initiate infection and disease, it is reasonable to assume that at least some glycan variation must have arisen from this selection pressure—the question is, how much? In discussing this matter further, we face a common limitation encountered by evolutionary biologists: since the glycan diversity observed today reflects both ongoing and past adaptation processes, it may not be possible to prove that a specific process has led to a certain structural outcome, because that particular selection pressure may have ceased to operate. We are therefore

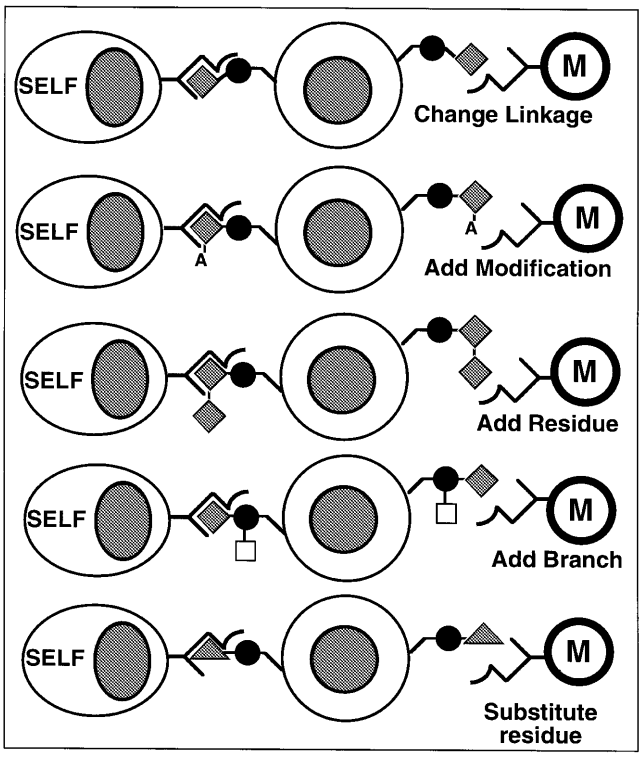


Fig. 4. Evading exogenous recognition without losing endogenous recognition. This schematic depicts in a simplified manner the various ways in which an organism might evade glycan recognition by a pathogen (M), without losing biologically important self-recognition mediated by a cognate lectin receptor within the same organism.

forced to rely on inferences in our reconstruction of possible evolutionary scenarios. Such processes can be experimentally studied only between rapidly coevolving model organisms.

Consider a lethal pathogen that specifically recognizes a glycan sequence which also happens to serve as a ligand for one or more endogenous lectins (Figure 1). As long as the latter interaction involves even one vital cell type or physiological process in the body, eliminating the cognate glycan structure to escape the pathogen is not a viable option. However, there are many potential ways to evade pathogen glycan recognition without losing endogenous glycan recognition function (Figure 4). As suggested in Figure 4, preservation of endogenous function in such circumstances may require some "looseness of fit" by the self-recognizing lectin. Such recognition properties are in fact very common with mammalian lectins, which often have shallow binding pockets and poor single site affinities (Weis and Drickamer, 1996), permitting significant variations in ligand structure without complete loss of endogenous recognition. Indeed, the inherent flexibility of glycosidic linkages in comparison to peptide bonds which may permit more of this "looseness of fit" may also add to the predisposition of glycans to the Red Queen Effect.

As discussed above, the dominance of sexual reproduction in vertebrates has been explained by the need to constantly diversify in order to escape from rapidly evolving pathogens (Maynard Smith, 1978). Similarly, the vertebrate immune system relies on somatic mechanisms for the generation of variation, i.e., somatic recombination, hypermutation, and gene conversion, followed by clonal selection (Tonegawa, 1983), allowing somatic vertebrate defense cells to evolve at rates comparable to those of microbes (Nesse and Williams, 1995). The generation of variability in offspring made possible by sexual reproduction comes at a cost of losing well-adapted genomes at each generation through recombination and of reduced reproductive potential with half the population being male. Similarly, the cost for allowing certain somatic cells to evolve at approximately the same rate as our pathogens comes with the price of uncontrolled proliferation (cancer), and misdirected short term somatic evolution (autoimmune diseases). We suggest here that the diverse and flexible nature of surface glycan structure may represent one more way in which long-lived vertebrates escape their rapidly adapting pathogens. The cost of this response remains to be elucidated, but the occurrence of aberrant glycosylation in diseases such as cancer and inflammatory diseases may provide some hints.

Learning from model animals with disrupted glycosylation

Disruptions of glycosyltransferase gene function in mice are instructive for exploring the function of the glycans. The ST6Gal I α2–6sialyltransferase is the only enzyme known to produce the well-characterized Siaα2–6Galβ1–4GlcNAc (6′-sialyllactosamine) termini on vertebrate glycans (Tsuji, 1996). While this sequence is a specific ligand for the B cell regulatory molecule siglec-2 (CD22) (Powell and Varki, 1995), it is also found on the surface of many other cell types, as well as on many circulating proteins. The level of ST6Gal I mRNA varies markedly amongst cell types, and transcription is extensively regulated by cell-type specific promoters that are modulated by hormones and cytokines (Jamieson *et al*., 1983, 1993; O'Hanlon *et al*., 1989; Wang *et al*., 1989, 1990, 1993; Shah *et al*., 1992; Vertino-Bell *et al*., 1994; Richardson and Jamieson, 1995; Lo and Lau, 1996). Despite these data suggesting diverse and complex roles for this gene product, the functional consequences of eliminating ST6Gal I expression in mice seem to be restricted to the B cell, with attenuated proliferative responses, and impaired antibody production (Hennet *et al*., 1998). These mice thrive in a specific pathogen-free vivarium, without any other obvious abnormalities in physiology, morphology, or behavior (Hennet *et al*., 1998). If the specific endogenous functions of the ST6Gal I oligosaccharide product are indeed restricted to B cells, why have vertebrates evolved to selectively regulate and express this structure in so many other locations? Also, why is expression of ST6Gal I upregulated so markedly during inflammatory responses (Jamieson *et al*., 1993; Tsuji, 1996)? One explanation is that the other molecules and cells carrying 6′-sialyllactosamine termini serve as "decoys," preventing an incoming cognate pathogenic organism from finding B cells, where the structure serves a critical role. Of course, the great majority of microbes actually bind preferentially to 3′-sialyllactosamine termini (Karlsson, 1995), often requiring high densities of these α2–3 linked sialic acid ligands to achieve functional avidity via multiple low-affinity binding sites (a "Velcro" effect). Thus, an additional purpose of upregulating ST6Gal I during inflammation may be to compete for the addition of α2–3 linked sialic acids in the Golgi, thereby decreasing overall densities of the latter. Limiting the existence of such decoy structures in time by turning down ST6Gal I expression after the inflammatory stimulus is past would prevent rapid adaptation by pathogens to such a defense mechanism. This is in fact what happens, with ST6Gal I expression returning to lower levels, when the inflammatory stimulus is past. In this regard, it is interesting that the hemagglutinin of one of the more virulent human viruses

Fucα1
GDP-Fuc 3
Galβ1-4GlcNAcβ1-R ----------> Galβ1-4GlcNAcβ1-R

Gene/Feature	FucT-III	FucT-V	FucT-VI	FucT-IV	FucT-VII
Linkages synthesized	3/4	4	4	4	4
Human expression	Mucosa	?	Liver	Gut,marrow	Marrow
Human Null allele	Common	?	Rare	No	No
Bovine homologues	Single gene			Same	Same
Murine homologues	Single pseudogene			Same	Same

Fig. 5. Evolutionary relationships among some mammalian α1–3 fucosyltransferase genes. The reaction catalyzed by the enzymes is shown, and the known phylogenetic distribution of the different genes is indicated. Note that one of the human enzymes (Fuc-TIII) can also generate α1–4 fucosyl linkages on Galβ1–3GlcNAc-R units. R = N-glycan, O-glycan, or glycosphingolipid.

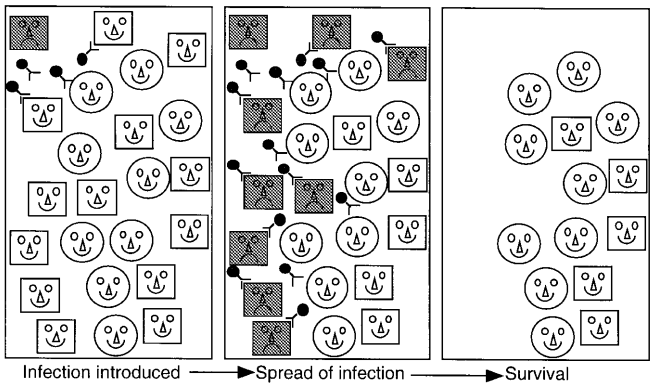


Fig. 6. Herd immunity—a possible evolutionary reason for generating and maintaining glycan diversity. A pathogen that can recognize glycans on the cells of only some individuals in the population (squares) cannot spread through the population because of the presence of others who have different glycan, and are therefore resistant (circles). The resistant individuals can even act as a barrier to spread of the pathogen to other susceptible individuals.

(Influenza A) is one of the rare pathogens that specifically recognizes α2–6 linked sialic acids.

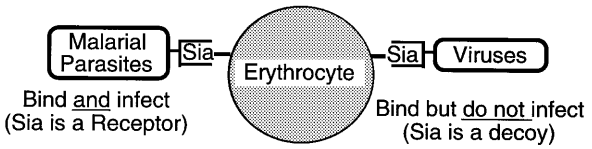
A more complex example can be found with the outer chain α1–3 linked fucose residues that are required components of the sialyl Lewis^x-based ligands for the selectin family of vascular adhesion molecules (Kansas, 1996). The two α1–3 fucosyltransferase genes (FucT IV and FucT VII) that are critical for forming these ligands (Maly *et al*., 1996; Costache *et al*., 1997) are highly conserved in expression between mice and humans. However, in humans this linkage can also be formed by at least three other fucosyltransferases (FucT III, FucT V, and FucT VI) (Figure 5). Comparative studies indicate that these three enzymes arose from recent duplications of an ancestral gene that remains single in the cow, and which appears to be inactivated (i.e., a pseudogene) in the mouse (Maly *et al*., 1996; Costache *et al*., 1997; Oulmouden *et al*., 1997). This cluster of duplicated fucosyltransferases shows highly variable expression amongst extravascular tissues in humans. Thus, FucT III is expressed primarily in epithelial cells (giving fucosylation of lumenally facing cell surfaces and of secreted mucins) and FucT VI in the liver (giving fucosylation of plasma glycoproteins). Despite such exquisite regulation of expression, inactivating mutations of both these genes exist in humans, without obvious functional consequence to the individuals missing them (Mollicone *et al*., 1994a,b; Brinkman-Van der Linden *et al*., 1996). The expression of these "Lewis blood group" structures on epithelial surfaces mediated by FucT-III is further complicated by interindividual variations in α1–2 fucosylation (generating "O" blood type structures) which can be differentially expressed in the blood and in the epithelia, with "nonsecretor" genotypes having uniquely absent fucosylation only in the epithelia (Kelly *et al*., 1995; Rouquier *et al*., 1995). The common theme appears to be that a critical endogenous function for a specific glycan structure in one cell type is accompanied by heterogenous expression of the same glycan in other cell types and tissues, with variations in the latter being found within and among species.

Protective glycan polymorphisms in populations?

Another factor favoring the creation and maintenance of glycan diversity may be herd immunity (Wills and Green, 1995). A pathogen recognizing a specific glycan cannot eliminate all members of a heterogenous population. Indeed, as modeled in Figure 6, glycan-negative individuals within a population can slow the general spread of the disease or even prevent the spread altogether. Density, population size, and frequency of contact between susceptible hosts are critical parameters for an epidemic to occur. The presence of as little as 10% nonsusceptible individuals in the population can make the difference between an epidemic and no epidemic at all. The evolution of such protective polymorphisms through individual selection is much more likely if the derived condition is the one observed in the smaller proportion of nonsusceptible individuals. As soon as the critical frequency of nonsusceptible individuals is reached, the reduced occurrence of epidemics will relax selection for the nonsusceptible allele, gene, or combination of genes. Such evolutionary scenarios however, depend critically on the costs resulting from changes in glycan structures and on the direct advantages to the individuals carrying the modification. If the cost is too high, a sufficient frequency within a species can never be reached. Differences in interspecies glycan expression could also represent selection stemming from cross-species transmission of pathogens (such as C-type retroviruses; see below).

Protective glycan variation within an organism?

We suggest that with regard to glycans, a variation of herd immunity may even operate amongst cells within a single organism. Erythrocytes are known to carry many complex glycan structures, many carried on the membrane protein glycophorin A. Even though this very complex glycoprotein is abundant (millions of copies per red blood cell), it does not seem essential for host functions, since individuals genetically lacking this protein suffer no apparent ill effects (Hadley *et al*., 1987). It thus appears that the primary function of glycophorin A is to serve as a polyvalent species-specific basis for flexible surface glycosylation (Rearden *et al*., 1990; Wilson and Planas, 1991). Since erythrocytes greatly outnumber other blood cells, an erythrocyte glycan-binding virus entering the bloodstream would be "sopped-up" by these non-nucleated cells that cannot serve as hosts for replication (Alford *et al*., 1994; Wybenga *et al*., 1996). Erythrocytes could thus serve as ideal decoys and "sinks" for such invading viruses. Of course, certain pathogens such as malaria-causing protozoans do not rely on a functioning nucleus and use the same glycan as receptors to invade erythrocytes (Orlandi *et al*., 1992; Sim *et al*., 1994; DeLuca *et al*., 1996) thus exerting the opposite selection



Types of Sialic Acids found on Erythrocytes

Species	Neu5Ac	Neu5Gc	Neu5,9Ac$_2$	Neu5Gc9Ac	Neu4,5Ac$_2$	Neu4Ac5Gc
Human	++++	0	0	0	0	0
Pig	+	+++	0	0	0	0
Cow	++	+++	+	+	0	0
Horse	+	++	+	+	++	+++
Mouse	++	+	++++	++	0	0

Fig. 7. Sialic acids on erythrocytes can act as pathogen receptors or decoy binding sites. Malarial parasites and viruses bind differentially to different types of sialic acids and/or sialic acid linkages that are present on cells from various species. As shown in this figure, the outcome of binding to erythrocytes will depend upon whether the cell is a host (malarial pathogen) or not (viruses, which require nucleated cells to replicate). In the latter case, the large number of erythrocytes in the bloodstream could potentially act as a "decoy sink" that could adsorb large amounts of virions, and commit them for eventual destruction by macrophages. These two types of binding represent opposing directions of selection, and could possibly explain the great variation in the types of sialic acids seen between different species (see lower part of the figure).

pressure. This combination of varying selection pressures may explain the marked intra- and interspecies variations of erythrocyte sialic acid types between different species (Figure 7).

Could herd immunity also operate within a single critical organ? The large family of complex brain glycosphingolipids show remarkable regional and cell-type-specific distribution patterns (Kotani *et al*., 1993, 1994), implying that they have highly specific roles in morphogenesis and development. Despite this, complete elimination of almost all complex glycolipids from the brains of mice by disruption of the G_{M2}/G_{D2} synthase enzyme gene does not give a severe developmental or morphological neural phenotype (Takamiya *et al*., 1996; Fukumoto *et al*., 1997). In this instance, evoking "redundancy" is not a reasonable explanation, since the two remaining gangliosides (G_{D3} and G_{M3}) cannot possibly mimic all of the complexities of structure of the many higher brain gangliosides that are eliminated in this mouse. An alternate explanation is that much of this intraorgan structural diversity is actually meant to provide protection from potent neurotoxins such as those generated by tetanus and botulinum bacteria, which work by binding specifically to certain neural glycolipid structures (Karlsson, 1995; Schengrund *et al*., 1996). If all cells in the brain expressed the same glycolipid structures at their surfaces, the chance of death from a single toxin would be much higher. Such protection by intraspecies diversity might be especially efficient if the spatial distribution of the different types of surface glycans was organized in an alternating pattern in which the closest neighboring cells all carry different glycans. Of course, the exposure to various neurotoxins is largely determined by the ecology of each organism. This would also fit with the observation that spatial patterns of brain gangliosides are variable among different species (Kappel *et al*., 1993; Wiegandt, 1995).

Marking enveloped viruses for destruction?

Another mystery is the complete loss of the Galα1–3Galβ1–4GlcNAc structure (α-Gal epitope) in the Old World primate lineage (Galili *et al*., 1988; Galili and Swanson, 1991; Galili, 1993). This major phylogenetic shift in glycosylation (note that the absence of α-Gal would lead to a compensatory increase in other terminal structures like sialic acids) does not explain any specific change in the overall phenotype of Old World primates. Furthermore, the experimental elimination of the α-Gal epitope in the mouse (which normally expresses it in abundance) gives no obvious major phenotype (Thall *et al*., 1995; Rother and Squinto, 1996). An example of intraspecies variation in glycosylation without obvious phenotypic change is the ABO-blood group polymorphism, which has evolved at least twice, once in the Old World monkeys and once in the hominoid apes (Doxiadis *et al*., 1998), and has been preserved for millions of years in these lineages (Martinko *et al*., 1993). With both α-Gal and the ABO system, "natural" IgM antibodies develop in individuals negative for a particular structure. The existence of such preformed antibodies recognizing surface glycans that are absent within an organism but found on the cells of other individuals or species is intriguing. It has been suggested that such antibodies can cause complement-mediated lysis of enveloped viruses generated within other individuals who can express the structure (Rother and Squinto, 1996; Takeuchi *et al*., 1996).

The unusually high titer (~1% of total circulating immunoglobulin) of antibodies against the α-Gal epitope in Old World primates suggests a critical function. This may represent a case where the loss of a specific glycan protects an entire evolutionary lineage (the Old World primates) from cross-species infection by certain C-type retroviruses produced in other animals. In practical terms, the absence of the α-Gal epitope and existence of natural antibodies in humans represents a major barrier to the xenotransplantation of pig tissues into humans. In the ongoing debate about xenotransplantation, it has been noted that if genetically engineered pigs lacking the α-Gal epitope were used as donors, this could augment the risk of transferring endogenous pig viruses and could result in new human transmissible viruses (Patience *et al*., 1997).

By the same reasoning, we suggest that blood group systems such as ABO could protect subsets of individuals within the same species from transmission of enveloped viral infections. For example, a virus generated within an A-positive individual should carry the A blood group structure on its surface, making it susceptible to lysis on first contact with the body fluids of a B or O-positive individual (Figure 8). In this scenario, the maintenance of intermediate frequencies of different blood groups would represent a case of balancing selection, preventing an excessively high frequency of any one type.

Glycans in host-pathogen coevolutionary arms races

There is other evidence for the importance of glycans in the coevolutionary arms race between vertebrate hosts and their microbial pathogens. The most common terminal family of sugars on vertebrate glycans are the sialic acids. These nine carbon sugars show a surprising taxonomic distribution, being found almost exclusively in animals of the deuterostome lineage (echinoderms, chaetognathes, hemichordates, and chordates), but also in some of their pathogens. The presence of sialic acids in these two phylogenetically distant groups of organisms is strongly suggestive of molecular mimicry. Indeed, a pathogen expressing sialic acids has a markedly reduced immunogenicity within the vertebrate organism, and would be protected from destruction by host complement. Evidence in favor of mimicry by convergent evolution comes from the fact that pathogen genes coding for generating and transferring sialic acids are generally not homologous to their animal counterparts, and thus seem to have been derived independently in the pathogens. The recent discovery of multiple "contingency loci" in pathogenic organisms, that function as rapidly reversible binary switches allowing the pathogen the creation of innumerable combinatorial phenotypes may be relevant. Some of these switches are adjacent to

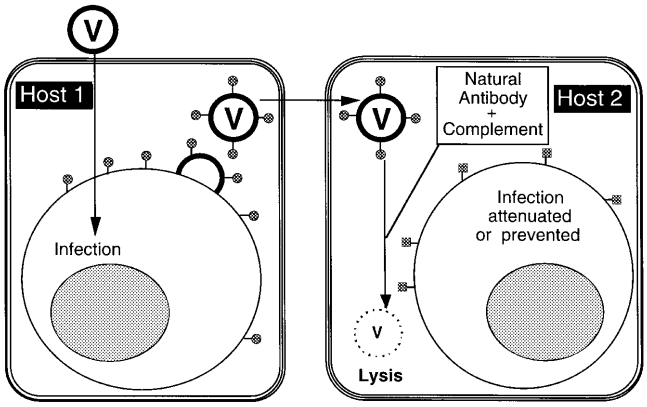


Fig. 8. A potential protective role for the natural antibodies directed against blood group glycans. An enveloped virus (V) replicating in the cells of an individual positive for a particular blood group type is likely to carry the corresponding carbohydrate structure on its surface. At the first contact with the body fluids of another individual of a different blood group type (especially if an inflammatory stimulus has allowed plasma proteins to enter tissue spaces) the natural antibody against the carbohydrate structure may cooperate with complement to neutralize virus particles before they can infect and undergo a new round of replication. Even a partial effect of this kind could serve to attenuate the severity of a given infection and eventually affect the spread of the pathogen through the population.

sialyltransferase genes, as well as to genes coding for the production of bacterial lipopolysaccharides that mimic the major human blood groups (Moxon *et al*., 1994).

Of course, many important pathogenic bacteria are themselves prone to infection and destruction by bacteriophages, which initiate infections by binding to the bacterial surface polysaccharides. Thus, while the surface structures of some microbes may allow them to successfully dodge the immune response of the vertebrate host, these may also be detected by the host and/or used as a receptor by a phage. Bacterial pathogens thus face the double task of succeeding in the invasion of a multicellular host while escaping their own pathogens. The antigenicity of pathogen phenotypes in both microbes and viruses is constantly being modulated by altered glycosylation of surface structures. Besides evading their own pathogens, many pathogens are likely to face competition from other pathogens. Competition between pathogens may also cause changes of glycan preference and/or pathogen surface glycans, e.g., multiple infections of *Neisseria menigitides* might lead to higher virulence due to short term competition between different strains (Van Baalen and Sabelis, 1995). Both plasmodia and trypanosomes are well known for causing multiple infections, and both rely on glycans for invading and/or evading host responses. Thus, it is conceivable that the coevolutionary arms race between vertebrate hosts and their pathogens is to a substantial degree taking place amongst their mutual, flexible "glycoforms."

In apparent contrast to other classes of glycans, the basic structural motifs and modifications of heparan sulfate glycosaminoglycans seem to have been conserved for several hundred million years of evolution (Cassaro and Dietrich, 1977; Dietrich *et al*., 1983). One suggested explanation is that endogenous heparan sulfate-binding proteins may have instead evolved different binding specificities with evolutionary time (Bernfield, in press). At first glance, this may seem an exception to the suggestion made here, that extensive glycan diversification has accompanied species evolution. However, heparan sulfates can generate numerous intrinsic structural variations, and there are currently inadequate data about the extent of species-specific differences in the specificities of the binding proteins and/or the expression of structural motifs in different cell lineages.

The difficulty of reconstructing the history of ancient pathogens and their interactions with hosts

Coevolutionary relationships between pathogens and their long-lived hosts are inherently difficult to study. The rapid rate of change of pathogens makes them moving targets, and the fact that the species concept cannot be applied to many of them makes reconstruction of their histories difficult. Unlike vertebrates where a certain amount of genetic divergence will result in the complete absence of gene flow (by reproductive isolation), many pathogenic microorganisms do not show a clear cut off point for a species definition (Roberts and Cohan, 1993). Some bacteria can exchange genes with others that are more than 20% divergent in their total DNA, and their evolution is better described as a network rather than as a phylogenetic tree (Smith *et al*., 1993). Horizontal transfer of genetic material between very divergent strains, cross-species transfer of pathogens, reassortment of viral genomes within vertebrate hosts and the importance of incorporation of host structural elements for viral evolution are all factors that drastically complicate phylogenetic reconstruction. Thus pathogens observed today may have been unrecognizably different in the not too distant past, while other past pathogens may have long gone extinct, evolved into unrecognizably benign or virulent forms or, in the case of viruses, even incorporated into the host genome. Taken together, these factors make it difficult to design conclusive experiments to support some of the evolutionary arguments presented above.

Exploring the relative importance of "self" and "non-self" recognition processes

While no single argument presented here might be considered conclusive, the aggregate of data available today suggest that "non-self" glycan recognition by pathogens is commoner than "self" recognition by endogenous lectins. Thus, a major fraction of the inter-species, intra-species, and intra-organismal glycan diversity observed today might be attributed to the Red Queen Effect. It is reasonable to assume that most highly conserved glycan structures have at least one critical endogenous function, somewhere, sometime, inside a multicellular organism. If such functions are mediated only by a small fraction of the total expression of the glycan within an organism, how does one find this "needle in a haystack"? More glycosyltransferase gene disruptions in intact animals will obviously be very helpful. More complex genetic manipulation of glycosylation in mice can then be used to test the effect of presence or absence of certain glycans on a variety of specific host tissues. Selective reexpression of viral ligands in conjunction with pathogenesis experiments will allow the testing of predictions about putative protective roles played by some glycans in certain tissues. Studying the comparative glycobiology of closely and distantly related species should also help, by ascertaining the rates of glycan diversification during evolution. Specific predictions about endogenous glycan function could also arise from examining species-specific variations in the tissue distribution of oligosaccharides (Figure 9). If the same glycan structure is found on the same molecules in the same cell type at the same point in development in many different vertebrate organisms, this is where it is likely to mediate an important function that could not be diversified by the Red Queen Effect. One established example of this principle is the high degree

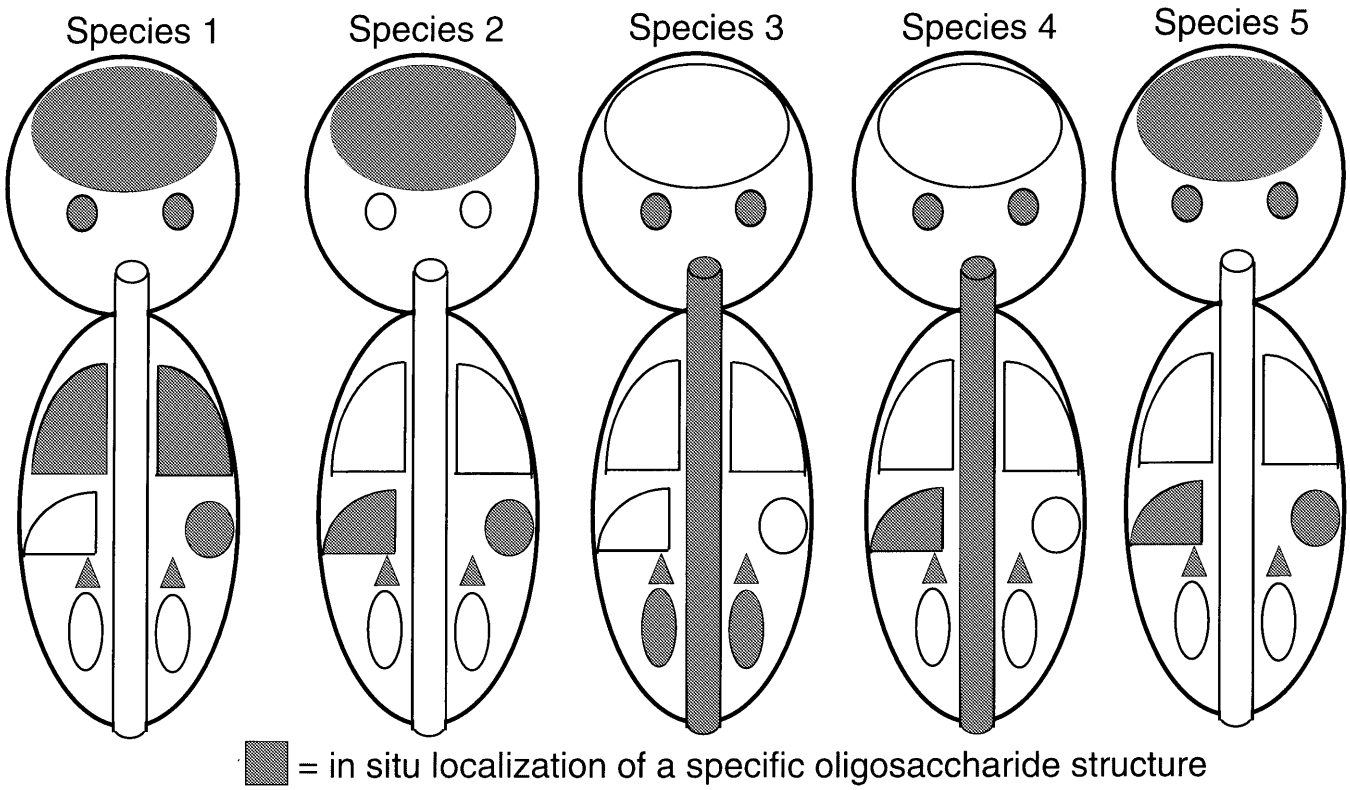


Fig. 9. If a particular glycan structure has a vital non-species-specific biological role, its distribution should be conserved between different species. In this hypothetical experiment, a probe (antibody or lectin) specific for a particular glycan structure is used to look for its expression in different organs. If the pattern of expression is highly variable between species, the glycan structure is unlikely to be mediating a vital role that is general to all the species studied. If expression is highly conserved in a particular organ/cell type (in the example shown here, in the adrenal gland), then it is more likely to be mediating a such a role.

of conservation in the vertebrate lineage of the endogenously important sulfated GalNAc residues on pituitary glycoprotein hormones (Baenziger, 1996; Hooper *et al*., 1996). Studies of pathogen–host interactions at the level of glycan recognition also promise to shed light on the importance of the most likely source of exogenous selection pressures. Despite general knowledge of the importance of glycan recognition by many pathogens, the details about their preference for specific structures as well their linkages to underlying structures remain inadequately characterized.

Quo vadis?

Much of the above discussion is obviously based upon broadly generalizable evolutionary reasoning and some reasonable conjectures. However, in the absence of a more extensive database of structural information about glycosylation in various taxa, many questions raised here about glycan diversity remain unanswered. What is the rate of oligosaccharide diversification during evolution? Is there a "molecular clock" for oligosaccharide diversification similar to that for diversion of gene sequences? If not, what is the extent of glycan diversification that can be tolerated in shorter evolutionary time scales, i.e., what are the constraints on the rate of glycan evolution? What are the selective forces driving the diversification and what are their relative roles? What is the functional significance of glycan diversification during evolution? Is species formation usually accompanied by changes in glycosylation?

Glycobiology remains a field where systematic descriptive nonmechanistic studies are needed, uncovering clues that will permit the formulation and testing of specific hypotheses about the critical endogenous biological roles of glycans. It is clear that the exploration of evolutionary diversification of glycosylation is worthwhile, and that may in turn educate us a great deal about oligosaccharide function. Potential paybacks range from increased insight into host susceptibility to infectious disease and better therapeutic design to the identification of glycans with critical endogenous roles during development, health, and disease.

Acknowledgments

This review is based partially on the Glyco XIV lecture, given at the XIVth. International Symposium on Glycoconjugates, in Zurich, Switzerland. The critical comments of Chris Wills, George Palade, Jeff Esko, and Hud Freeze are greatly appreciated, and the support of the NIH and the Mathers Foundation is gratefully acknowledged.

References

Alford,D., Ellens,H. and Bentz,J. (1994) Fusion of influenza virus with sialic acid-bearing target membranes. *Biochemistry*, **33,** 1977–1987.

Asano,M., Furukawa,K., Kido,M., Matsumoto,S., Umesaki,Y., Kochibe,N. and Iwakura,Y. (1997) Growth retardation and early death of β-1,4-galactosyltransferase knockout mice with augmented proliferation and abnormal differentiation of epithelial cells. *EMBO J.*, **16,** 1850–1857.

Baenziger,J.U. (1996) Glycosylation: to what end for the glycoprotein hormones? *Endocrinology*, **137,** 1520–1522.

Barton,N.H. and Charlesworth,B. (1998) Why sex and recombination? *Science*, **281,** 1986–1990.

Bernfield,M. (1999) Biological functions of cell surface heparan sulfate. *Annu. Rev. Biochem*., in press.

Brinkman-Van der Linden,E.C.M., Mollicone,R., Oriol,R., Larson,G., Van den Eijnden,D.H. and Van Dijk,W. (1996) A missense mutation in the *FUT6* gene results in total absence of α3-fucosylation of human α_1-acid glycoprotein. *J. Biol. Chem*., **271,** 14492–14495.

Bry,L., Falk,P.G., Midtvedt,T. and Gordon,J.I. (1996) A model of host–microbial interactions in an open mammalian ecosystem. *Science*, **273,** 1380–1383.

Campbell,R.M., Metzler,M., Granovsky,M., Dennis,J.W. and Marth,J.D. (1995) Complex asparagine-linked oligosaccharides in *Mgat1*-null embryos. *Glycobiology*, **5,** 535–543.

Cassaro,C.M. and Dietrich,C.P. (1977) Distribution of sulfated mucopolysaccharides in invertebrates. *J. Biol. Chem*., **252,** 2254–2261.

Chui,D., Oh-eda,M., Liao,Y.F., Panneerselvam,K., Lal,A., Marek,K.W., Freeze,H.H., Moremen,K.W., Fukuda,M.N. and Marth,J.D. (1997) Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell*, **90,** 157–167.

Costache,M., Apoil,P.A., Cailleau,A., Elmgren,A., Larson,G., Henry,S., Blancher,A., Iordachescu,D., Oriol,R. and Mollicone,R. (1997) Evolution of fucosyltransferase genes in vertebrates. *J. Biol. Chem*., **272,** 29721–29728.

Damm,J.B.L., Voshol,H., Hård,K., Kamerling,J.P. and Vliegenthart,J.F.G. (1989) Analysis of *N*-acetyl-4-*O*-acetylneuraminic-acid-containing N-linked carbohydrate chains released by peptide-N^4- (*N*-acetyl-β-glucosaminyl)asparagine amidase F—application to the structure determination of the carbohydrate chains of equine fibrinogen. *Eur. J. Biochem*., **180,** 101–110.

Debeire,P., Montreuil,J., Moczar,E., van Halbeek,H. and Vliegenthart,J.F. (1985) Primary structure of two major glycans of bovine fibrinogen. *Eur. J. Biochem*., **151,** 607–611.

DeLuca,G.M., Donnell,M.E., Carrigan,D.J. and Blackall,D.P. (1996) *Plasmodium falciparum* merozoite adhesion is mediated by sialic acid. *Biochem. Biophys. Res. Commun*., **225**, 726–732.

Dietrich,C.P., Nader,H.B. and Straus,A.H. (1983) Structural differences of heparan sulfates according to the tissue and species of origin. *Biochem. Biophys. Res. Commun*., **111,** 865–871.

Dobzhnasky,T. (1973) Nothing in biology makes sense, except in the light of evolution. *Am. Biol. Teacher*, **35,** 125–129.

Doxiadis,G.G., Otting,N., Antunes,S.G., de Groot,N.G., Harvey,M., Doxiadis,I.I., Jonker,M. and Bontrop,R.E. (1998) Characterization of the ABO blood group genes in macaques: evidence for convergent evolution. *Tissue Antigens*, **51,** 321–326.

Drickamer,K. and Taylor,M.E. (1998) Evolving views of protein glycosylation. *Trends Biochem. Sci*., **23,** 321–324.

Ellies,L.G., Tsuboi,S., Petryniak,B., Lowe,J.B., Fukuda,M. and Marth,J.D. (1998) Core 2 oligosaccharide biosynthesis distinguishes between selectin ligands essential for leukocyte homing and inflammation. *Immunity*, **9,** 881–890.

Esko,J.D. (1991) Genetic analysis of proteoglycan structure, function and metabolism. *Curr. Opin. Cell, Biol*., **3,** 805–816.

Fukumoto,S., Yamamoto,A., Hasegawa,T., Abe,K., Takamiya,K., Okada,M., Min,Z.J., Furukawa,K., Miyazaki,H., Tsuji,Y., Goto,G., Suzuki,M. and Shiku,H. (1997) Genetic remodeling of gangliosides resulted in the enhanced reactions to the foreign substances in skin. *Glycobiology*, **7,** 1111–1120.

Gahmberg,C.G. and Tolvanen,M. (1996) Why mammalian cell surface proteins are glycoproteins. *Trends Biochem. Sci*., **21,** 308–311.

Galili,U. (1993) Interaction of the natural anti-Gal antibody with α-galactosyl epitopes: a major obstacle for xenotransplantation in humans. *Immunol. Today*, **14,** 480–482.

Galili,U. and Swanson,K. (1991) Gene sequences suggest inactivation of α-1,3-galactosyltransferase in catarrhines after the divergence of apes from monkeys. *Proc. Natl. Acad. Sci. USA*, **88,** 7401–7404.

Galili,U., Shohet,S.B., Kobrin,E., Stults,C.L. and Macher,B.A. (1988) Man, apes and Old World monkeys differ from other mammals in the expression of alpha-galactosyl epitopes on nucleated cells. *J. Biol. Chem*., **263,** 17755–17762.

Hadley,T.J., Klotz,F.W., Pasvol,G., Haynes,J.D., McGinniss,M.H., Okubo,Y. and Miller,L.H. (198j7) Falciparum malaria parasites invade erythrocytes that lack glycophorin A and B (MkMk). Strain differences indicate receptor heterogeneity and two pathways for invasion. *J. Clin. Invest*., **80,** 1190–1193.

Haldane,J.B.S. (1949) Disease and evolution. In *Symposium sui fattori ecologi e genetici della specilazione negli animali, Supplemento a la Ricerca Scientifica Anno 19th*, pp. 68–75.

Hart,G.W. (1992) Glycosylation. *Curr. Opin. Cell Biol*., **4,** 1017–1023.

Hennet,T., Chui,D., Paulson,J.C. and Marth,J.D. (1998) Immune regulation by the ST6Gal sialyltransferase. *Proc. Natl. Acad. Sci. USA*, **95,** 4504–4509.

Hooper,L.V., Manzella,S.M. and Baenziger,J.U. (1996) From legumes to leukocytes: biological roles for sulfated carbohydrates. *FASEB J*., **10,** 1137–1146.

Jamieson,J.C., Kaplan,H.A., Woloski,B.M., Hellman,M. and Ham,K. (1983) Glycoprotein biosynthesis during the acute-phase response to inflammation. *Can. J. Biochem. Cell, Biol*., **61,** 1041–1048.

Jamieson,J.C., McCaffrey,G. and Harder,P.G. (1993) Sialyltransferase: a novel acute-phase reactant. *Comp. Biochem. Physiol. [B]*, **105B,** 29–33.

Kansas,G.S. (1996) Selectins and their ligands: current concepts and controversies. *Blood*, **88,** 3259–3287.

Kappel,T., Hilbig,R. and Rahmann,H. (1993) Variability in brain ganglioside content and composition of endothermic mammals, heterothermic hibernators and ectothermic fishes. *Neurochem. Int*., **22,** 555–566.

Karlsson,K.A. (1995) Microbial recognition of target-cell glycoconjugates. *Curr. Opin. Struct. Biol*., **5,** 622–635.

Kawagoe,K., Kitamura,D., Okabe,M., Taniuchi,I., Ikawa,M., Watanabe,T., Kinoshita,T. and Takeda,J. (1996) Glycosylphosphatidylinositol-anchor-deficient mice: mplications for clonal dominance of mutant cells in paroxysmal nocturnal hemoglobinuria. *Blood*, **87,** 3600–3606.

Kelly,R.J., Rouquier,S., Giorgi,D., Lennon,G.G. and Lowe,J.B. (1995) Sequence and expression of a candidate for the human *Secretor* blood group α (1,2)fucosyltransferase gene (*FUT2*). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J. Biol. Chem*., **270,** 4640–4649.

Kobata,A. (1992) Structures and functions of the sugar chains of glycoproteins. *Eur. J. Biochem*., **209,** 483–501.

Kornfeld,S. (1998) Diseases of abnormal protein glycosylation: an emerging area. *J. Clin. Invest*., **101,** 1293–1295.

Kotani,M., Kawashima,I., Ozawa,H., Terashima,T. and Tai,T. (1993) Differential distribution of major gangliosides in rat central nervous system detected by specific monoclonal antibodies. *Glycobiology*, **3,** 137–146.

Kotani,M., Kawashima,I., Ozawa,H., Ogura,K., Ishizuka,I., Terashima,T. and Tai,T. (1994) Immunohistochemical localization of minor gangliosides in the rat central nervous system. *Glycobiology*, **4,** 855–865.

Li,Q.X., Hasty,P. and Shur,B.D. (1997) Targeted mutation in β1,4-galactosyltransferase leads to pituitary insufficiency and neonatal lethality. *Dev. Biol*., **181,** 257–267.

Lis,H. and Sharon,N. (1993) Protein glycosylation—structural and functional aspects. *Eur. J. Biochem*., **218,** 1–27.

Lo,N.W. and Lau,J.T.Y. (1996) Transcription of the β-galactoside α2,6-sialyltransferase gene in B lymphocytes is directed by a separate and distinct promoter. *Glycobiology*, **6,** 271–279.

Maly,P., Thall,A.D., Petryniak,B., Rogers,G.E., Smith,P.L., Marks,R.M., Kelly,R.J., Gersten,K.M., Cheng,G.Y., Saunders,T.L., Camper,S.A., Camphausen,R.T., Sullivan,F.X., Isogai,Y., Hindsgaul,O., Von Andrian,U.H. and Lowe,J.B. (1996) The α (1,3)Fucosyltransferase Fuc-TVII controls leukocyte trafficking through an essential role in L-, E- and P-selectin ligand biosynthesis. *Cell*, **86,** 643–653.

Martinko,J.M., Vincek,V., Klein,D. and Klein,J. (1993) Primate ABO glycosyltransferases: evidence for trans-species evolution. *Immunogenetics*, **37,** 274–278.

Maynard Smith,J. (1978) *The Evolution of Sex*. Cambridge University Press, New York.

Mizuochi,T., Taniguchi,T., Asami,Y., Takamatsu,J., Okude,M., Iwanaga,S. and Kobata,A. (1982) Comparative studies on the structures of the carbohydrate moieties of human fibrinogen and abnormal fibrinogen Nagoya. *J. Biochem. (Tokyo)*, **92,** 283–293.

Mollicone,R., Reguigne,I., Fletcher,A., Aziz,A., Rustam,M., Weston,B.W., Kelly,R.J., Lowe,J.B. and Oriol,R. (1994a) Molecular basis for plasma α (1,3)-fucosyltransferase gene deficiency (FUT6). *J. Biol. Chem*., **269,** 12662–12671.

Mollicone,R., Reguigne,I., Kelly,R.J., Fletcher,A., Watt,J., Chatfield,S., Aziz,A., Cameron,H.S., Weston,B.W., Lowe,J.B. and Oriol,R. (1994b) Molecular basis for Lewis α (1,3/1,4)-fucosyltransferase gene deficiency (FUT3) found in Lewis-negative Indonesian pedigrees. *J. Biol. Chem*., **269,** 20987–20994.

Moxon,E.R., Rainey,P.B., Nowak,M.A. and Lenski,R.E. (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol*., **4,** 24–33.

Nesse,R.M. and Williams,G.C. (1995) *Why We Get Sick: The New Science of Darwinian Medicine*. Times Books, New York.

O'Hanlon,T.P., Lau,K.M., Wang,X.C. and Lau,J.T.Y. (1989) Tissue-specific expression of β-galactoside α-2,6-sialyltransferase. Transcript heterogeneity predicts a divergent polypeptide. *J. Biol. Chem*., **264,** 17389–17394.

Orlandi,P.A., Klotz,F.W. and Haynes,J.D. (1992) A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of *Plasmodium falciparum* recognizes the terminal Neu5Ac (α2–3)Gal-sequences of glycophorin A. *J. Cell. Biol*., **116,** 901–909.

Oulmouden,A., Wierinckx,A., Petit,J.M., Costache,M., Palcic,M.M., Mollicone,R., Oriol,R. and Julien,R. (1997) Molecular cloning and expression of a bovine α (1,3)-fucosyltransferase gene homologous to a putative ancestor gene of the human *FUT3-FUT5-FUT6* cluster. *J. Biol. Chem.*, **272,** 8764–8773.

Partridge,L. and Hurst,L.D. (1998) Sex and conflict. *Science*, **281,** 2003–2008.

Patience,C., Takeuchi,Y. and Weiss,R.A. (1997) Infection of human cells by an endogenous retrovirus of pigs [see comments]. *Nature Med.*, **3,** 282–286.

Paulson,J.C. (1989) Glycoproteins: what are the sugar chains for. *Trends Biochem. Sci.*, **14,** 272–276.

Plummer,T.H.,Jr., Phelan,A.W. and Tarentino,A.L. (1996) Porcine fibrinogen glycopeptides: substrates for detecting endo-β-*N*-acetylglucosaminidases F_2 and F_3. *Anal. Biochem.*, **235,** 98–101.

Powell,L.D. and Varki,A. (1995) I-Type lectins. *J. Biol. Chem.*, **270,** 14243–14246.

Rademacher,T.W., Parekh,R.B. and Dwek,R.A. (1988) Glycobiology. *Annu. Rev. Biochem.*, **57,** 785–838.

Rearden,A., Phan,H., Kudo,S. and Fukuda,M. (1990) Evolution of the glycophorin gene family in the hominoid primates. *Biochem. Genet.*, **28,** 209–222.

Richardson,K. and Jamieson,J.C. (1995) Release of sialyltransferases from rat liver Golgi membranes by a cathepsin D-like proteinase: Comparison of the release of Galβ1–4GlcNAcα2–6 sialyltransferase, Galβ1–3 (4)GlcNAcα2–3 sialyltransferase and lactosylceramide α2–3 sialyltransferase (SAT-1). *Comp. Biochem. Physiol. [B]* **110B,** 445–450.

Roberts,M.S. and Cohan,F.M. (1993) The effect of DNA sequence divergence on sexual isolation in Bacillus. *Genetics*, **134,** 401–408.

Rother,R.P. and Squinto,S.P. (1996) The α-galactosyl epitope: a sugar coating that makes viruses and cells unpalatable. *Cell*, **86,** 185–188.

Rouquier,S., Lowe,J.B., Kelly,R.J., Fertitta,A.L., Lennon,G.G. and Giorgi,D. (1995) Molecular cloning of a human genomic region containing the *H* blood group α (1,2)fucosyltransferase gene and two *H* locus-related DNA restriction fragments. Isolation of a candidate for the human *Secretor* blood group locus. *J. Biol. Chem.*, **270,** 4632–4639.

Salmivirta,M., Lidholt,K. and Lindahl,U. (1996) Heparan sulfate: a piece of information. *FASEB J.*, **10,** 1270–1279.

Schengrund,C.L., DasGupta,B.R., Hughes,C.A. and Ringler,N.J. (1996) Ganglioside-induced adherence of botulinum and tetanus neurotoxins to adducin. *J. Neurochem.*, **66,** 2556–2561.

Shah,S., Lance,P., Smith,T.J., Berenson,C.S., Cohen,S.A., Horvath,P.J., Lau,J.T.Y. and Baumann,H. (1992) *n*-Butyrate reduces the expression of β-galactoside α2,6-sialyltransferase in Hep G2 cells. *J. Biol. Chem.*, **267,** 10652–10658.

Sharon,N. (1996) Carbohydrate-lectin interactions in infectious disease. *Adv. Exp. Med. Biol.*, **408,** 1–8.

Sharon,N. (1998) Lectins: from obscurity into the limelight. *Protein Sci.*, **7,** 2042–2048.

Sim,B.K., Chitnis,C.E., Wasniowska,K., Hadley,T.J. and Miller,L.H. (1994) Receptor and ligand domains for invasion of erythrocytes by *Plasmodium falciparum*. *Science*, **264,** 1941–1944.

Smith,J.M., Smith,N.H., O'Rourke,M. and Spratt,B.G. (1993) How clonal are bacteria? *Proc. Natl. Acad. Sci. USA*, **90,** 4384–4388.

Takamiya,K., Yamamoto,A., Furukawa,K., Yamashiro,S., Shin,M., Okada,M., Fukumoto,S., Haraguchi,M., Takeda,N., Fujimura,K., Sakae,M., Kishikawa,M., Shiku,H. and Aizawa,S. (1996) Mice with disrupted GM2/GD2 synthase gene lack complex gangliosides but exhibit only subtle defects in their nervous system. *Proc. Natl. Acad. Sci. USA*, **93,** 10662–10667.

Takeuchi,Y., Porter,C.D., Strahan,K.M., Preece,A.F., Gustafsson,K., Cosset,F.L., Weiss,R.A. and Collins,M.K.L. (1996) Sensitization of cells and retroviruses to human serum by (α1–3) galactosyltransferase. *Nature*, **379,** 85–88.

Tarutani,M., Itami,S., Okabe,M., Ikawa,M., Tezuka,T., Yoshikawa,K., Kinoshita,T. and Takeda,J. (1997) Tissue-specific knockout of the mouse *Pig-a* gene reveals important roles for GPI-anchored proteins in skin development. *Proc. Natl. Acad. Sci. USA*, **94,** 7400–7405.

Thall,A.D., Maly,P. and Lowe,J.B. (1995) Oocyte Galα1,3Gal epitopes implicated in sperm adhesion to the zona pellucida glycoprotein ZP3 are not required for fertilization in the mouse. *J. Biol. Chem.*, **270,** 21437–21440.

Tonegawa,S. (1983) Somatic generation of antibody diversity. *Nature*, **302,** 575–581.

Tsuji,S. (1996) Molecular cloning and functional analysis of sialyltransferases. *J. Biochem. (Tokyo)*, **120,** 1–13.

Van Baalen,M. and Sabelis,M.W. (1995) The dynamics of multiple infection and the evolution of virulence. *Am. Nat.*, **146,** 881–910.

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. and Marth,J. (1995) Oligosaccharides in vertebrate development. *Semin. Dev. Biol.*, **6,** 127–138.

Vertino-Bell,A., Ren,J., Black,J.D. and Lau,J.T.Y. (1994) Developmental regulation of β-galactoside α2,6-sialyltransferase in small intestine epithelium. *Dev. Biol.*, **165,** 126–136.

Wang,X., O'Hanlon,T.P. and Lau,J.T.Y. (1989) Regulation of β-galactoside α2,6-sialyltransferase gene expression by dexamethasone. *J. Biol. Chem.*, **264,** 1854–1859.

Wang,X.C., Smith,T.J. and Lau,J.T.Y. (1990) Transcriptional regulation of the liver β-galactoside α2,6-sialyltransferase by glucocorticoids. *J. Biol. Chem.*, **265,** 17849–17853.

Wang,X.C., Vertino,A., Eddy,R.L., Byers,M.G., Jani-Sait,S.N., Shows,T.B. and Lau,J.T.Y. (1993) Chromosome mapping and organization of the human β-galactoside α2–6 sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.*, **268,** 4355–4361.

Weis,W.I. and Drickamer,K. (1996) Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.*, **65,** 441–473.

Wiegandt,H. (1995) The chemical constitution of gangliosides of the vertebrate nervous system. *Behav. Brain Res.*, **66,** 85–97.

Wills,C. and Green,D.R. (1995) A genetic herd-immunity model for the maintenance of MHC polymorphism. *Immunol. Rev.*, **143,** 263–292.

Wilson,D.J. and Planas,J.M. (1991) Species specificity in mouse glycophorin. *Biochem. Biophys. Res. Commun.*, **176,** 820–825.

Wybenga,L.E., Epand,R.F., Nir,S., Chu,J.W.K., Sharom,F.J., Flanagan,T.D. and Epand,R.M. (1996) Glycophorin as a receptor for Sendai virus. *Biochemistry*, **35,** 9513–9518.

Yamashita,K., Hitoi,A., Matsuda,Y., Tsuji,A., Katunuma,N. and Kobata,A. (1983a) Structural studies of the carbohydrate moieties of rat kidney gamma-glutamyltranspeptidase. An extremely heterogeneous pattern enriched with nonreducing terminal N-acetylglucosamine residues. *J. Biol. Chem.*, **258,** 1098–1107.

Yamashita,K., Tachibana,Y., Hitoi,A., Matsuda,Y., Tsuji,A., Katunuma,N. and Kobata,A. (1983b) Difference in the sugar chains of two subunits and of isozymic forms of rat kidney gamma-glutamyltranspeptidase. *Arch. Biochem. Biophys.*, **227,** 225–232.

Yamashita,K., Tachibana,Y., Shichi,H. and Kobata,A. (1983c) Carbohydrate structures of bovine kidney gamma-glutamyltranspeptidase. *J. Biochem. (Tokyo)*, **93,** 135–147.

Yamashita,K., Hitoi,A., Tateishi,N., Higashi,T., Sakamoto,Y. and Kobata,A. (1985) The structures of the carbohydrate moieties of mouse kidney gamma-glutamyltranspeptidase: occurrence of X-antigenic determinants and bisecting N-acetylglucosamine residues. *Arch. Biochem. Biophys.*, **240,** 573–582.

Yamashita,K., Hitoi,A., Matsuda,Y., Miura,T., Katunuma,N. and Kobata,A. (1986) Structures of sugar chains of human kidney gamma-glutamyltranspeptidase. *J. Biochem. (Tokyo)*, **99,** 55–62.