Evolutionary Considerations in Studying the Sialome: Sialic Acids and the Host–Pathogen Interface

4

Amanda L. Lewis and Ajit Varki

Glycobiology Research and Training Center, Departments of Medicine, Biological Sciences and Cellular and Molecular Medicine, University of California at San Diego, La Jolla, CA 92093-0687, USA

4.1 Introduction and Summary

The apparent variation in cell surface oligosaccharide (glycan) structures within and between species has long been an interesting and yet puzzling aspect of glycobiology. It has been suggested that this diversity reflects the often-conflicting pressures of evading pathogens, while simultaneously maintaining endogenous functions [1–4]. Since most pathogens replicate much faster than their hosts, they can rapidly evolve different ways to target or mimic structures that are critical for host processes – a feature called the "Red Queen" effect[1], that may be especially relevant to glycans [2–5]. Two categories of pathogen molecules are relevant to protein–glycan interactions at the host–pathogen interface: (1) the pathogen receptors/toxins that recognize and *bind to* host glycans, and (2) pathogen surface molecules that *mimic* host glycans. A great number and variety of pathogens use host glycan structures for targeted adherence, invasion, or cytotoxicity. In fact, the vast majority of known glycan-binding lectins are those used by pathogens [6, 7]. Host population heterogeneity of the targeted glycan structures may also ensure survival of some individuals and reduce the chance of epidemic spread, a phenomenon referred to in other contexts as "herd immunity" [2, 8, 9]. The extent of population heterogeneity (of a targeted glycan structure) is constrained by the existence and stringency of internal host functions in which the same structure participates. In a similar yet opposing pathogenic mechanism, many microbes decorate themselves with glycans that are similar to or identical with structures expressed in the host (i.e. "molecular mimicry") [10, 11]. Microbial hijacking of host lectins via molecular mimicry is a virulence mechanism likely broader in scope than currently recognized.

Despite frequent pathogen targeting and mimicry of carbohydrate structures, *all* cells in nature are covered with a dense coating of glycans [12]. This remarkable "rule"

---

[1] The "Red Queen" effect in evolutionary processes is based on the observation to Alice by the Red Queen in Lewis Carroll's *Through the Looking Glass* – that "it takes all the running you can do, to keep in the same place."

---

suggests that glycans afford the most flexible way to adapt a cell surface away from pathogen recognition, and/or that certain cell surface glycans are required for non-dispensable host functions. Broadening our understanding of host–pathogen co-evolution requires further study of glycans and lectins on *both* sides of the host–pathogen equation. This chapter focuses mainly on the sialic acids (Sias), providing examples of ongoing studies in the area of Sia-dependent host–pathogen interactions. To place these interactions in a broader context, we consider the diversity, distribution, biosynthesis, and evolution of Sias. We also review common glycan analysis techniques that can result in loss of Sias or Sia modifications and finally suggest a "Sialome"[2] project for archiving information about Sias in nature.

4.2   Pathogens that Target or Mimic Host Sialic Acids

Sialic acids are a diverse family of sugars that often occupy the non-reducing outermost ends of glycan chains in many animals [13–17]. Due in part to this terminal location, Sias are the glycan receptors most frequently targeted for recognition by pathogens. There are numerous documented examples of pathogens that use their Sia-binding proteins (called agglutinins in viruses, adhesins in bacteria, and lectins in protozoa and fungi) to adhere to, or gain entry into, host cells. There are also a number of bacterial toxins that bind Sias to effect their toxicity on target cells. Figure 4.1 provides an incomplete listing of some common pathogens that target Sia residues [see ref 17 for a listing of ~100 Sia-binding pathogens/toxins in nature].

Many common and sometimes fatal illnesses are caused by Sia-binding pathogens. For example, *Plasmodium falciparum* is a common agent of malaria, the leading cause of illness (300–500 million per year) and death (>1 million per year) worldwide (World Health Organization estimates). *Plasmodium falciparum* merozoites typically infect red blood cells in a Sia-dependent manner, leading to fever, chills, and flu-like symptoms, and if not treated, kidney failure, seizures, coma, and death. Influenza virus A (the causative agent of the "flu") is also a Sia-binding pathogen responsible for >100 000 hospitalizations, >30 000 deaths, and ~\$15 billion costs in the United States each year. Yet another is *Helicobacter pylori*, commonly involved in the formation of gastric and duodenal ulcers, a condition experienced by as many as 5 million persons at any one time in the United States (estimate from Digestive Diseases in the United States: Epidemiology and Impact, NIH Publication No. 94-1447, 1994).

Among the Sia-expressing pathogens, *Escherichia coli* (K1 and K92), *Neisseria meningitidis* (Groups B, C, Y, and W135), and Group B Streptococci all express capsular Sias and cause sepsis and meningitis, particularly in young children [10]. Other important Sia-expressing pathogens include *Haemophilus influenzae*, the most common cause of childhood ear infections [18], and some strains of *Escherichia coli* that cause gastrointestinal and urinary tract infections [10]. Indeed, the list of Sia-binding and Sia-expressing pathogens continues to expand, and recently available genomic data suggest that this may be even more common than previously realized [19].

---

[2] The term "Sialome" was recently coined [5] to denote "the total complement of Sia types and linkages and their modes of presentation on a particular organelle, cell, tissue, organ or organism – as found at a particular time and under specific conditions."
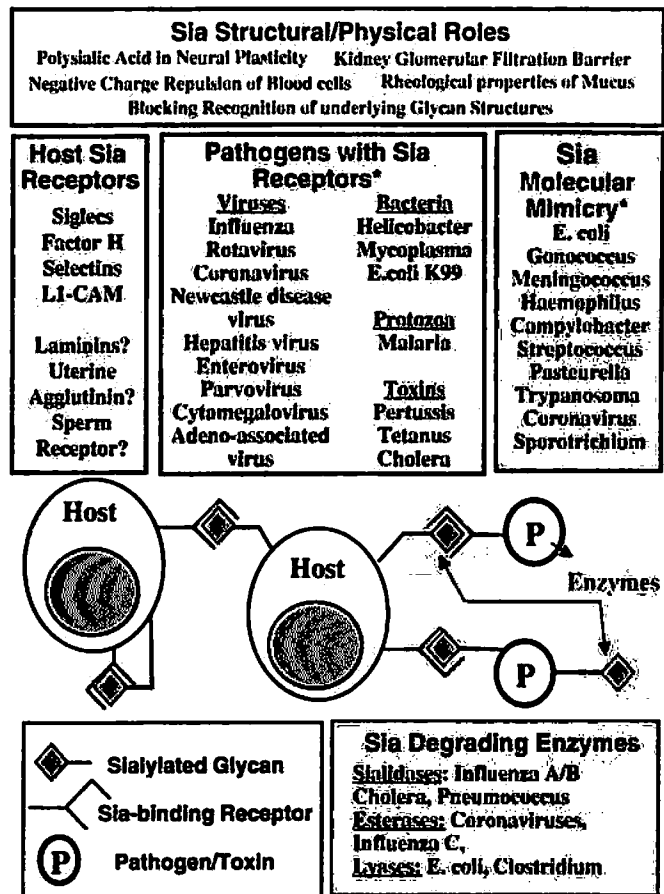
Sia Structural/Physical Roles

Polysialic Acid in Neural Plasticity   Kidney Glomerular Filtration Barrier

Negative Charge Repulsion of Blood cells   Rheological properties of Mucus

Blocking Recognition of underlying Glycan Structures

Host Sia Receptors	Pathogens with Sia Receptors*		Sia Molecular Mimicry*
Siglecs	**Viruses**	**Bacteria**	E. coli
Factor H	Influenza	Helicobacter	Gonococcus
Selectins	Rotavirus	Mycoplasma	Meningococcus
L1-CAM	Coronavirus	E.coli K99	Haemophilus
	Newcastle disease virus		Campylobacter
Laminins?	Hepatitis virus	**Protozoa**	Streptococcus
Uterine	Enterovirus	Malaria	Pasteurella
Agglutinin?	Parvovirus	**Toxins**	Trypanosoma
Sperm	Cytomegalovirus	Pertussis	Coronavirus
Receptor?	Adeno-associated virus	Tetanus	Sporotrichium
		Cholera	

◆— Sialylated Glycan

⌐< Sia-binding Receptor

Ⓟ Pathogen/Toxin

Sia Degrading Enzymes

Sialidases: Influenza A/B Cholera, Pneumococcus

Esterases: Coronaviruses, Influenza C,

Lyases: E. coli, Clostridium

Plate 4.1   Biological roles of sialic acids. The diverse biological roles of Sias include structural/physical functions (shown in gray), or functions that require Sia-recognizing proteins. In the latter category, intrinsic (host) Sia-recognizing proteins are typically involved in endogenous functions (yellow), while extrinsic (pathogen) Sia-recognizing proteins mediate mechanisms of host cell adherence or entry (green). Microbial pathogens can also decorate themselves with Sia to hijack host processes (blue). In addition, some microbes express enzymes that degrade Sias (pink). A schematic representation of these relationships is color-coded as above. See the text and the cited literature for details about the biological functions of Sias depicted in this figure. Note: most pathogens express Sias or Sia-binding proteins in a strain-specific manner; hence these properties do not apply to all members of the indicated pathogen classes. Also, this figure represents an incomplete listing of most of the categories above. See [17] for details.
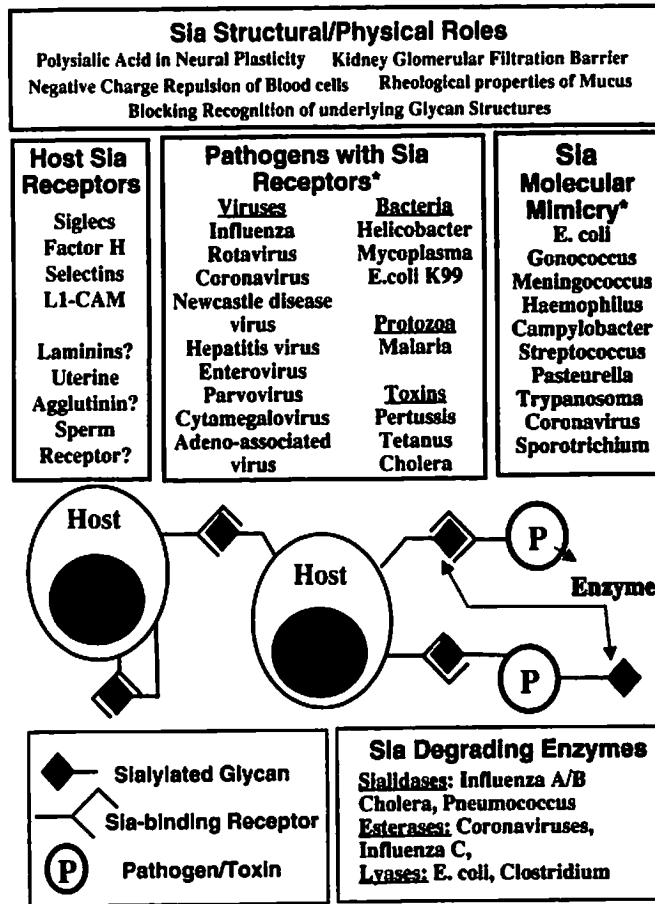
Figure 4.1 Biological roles of sialic acids. The diverse biological roles of Sias include structural/physical functions, or functions that require Sia-recognizing proteins. In the latter category, intrinsic (host) Sia-recognizing proteins are typically involved in endogenous functions, while extrinsic (pathogen) Sia-recognizing proteins mediate mechanisms of host cell adherence or entry. Microbial pathogens can also decorate themselves with Sia to hijack host processes. In addition, some microbes express enzymes that degrade Sias. See the text and the cited literature for details about the biological functions of Sias depicted in this figure. Note: most pathogens express Sias or Sia-binding proteins in a strain-specific manner; hence these properties do not apply to all members of the indicated pathogen classes. The asterisks indicate that, this figure represents an incomplete listing of most of the categories above. See [17] for details. A full-color version of this figure is included in the Plate section of this book.

In the light of the importance of Sias in host–pathogen interactions, the biology and evolution of this family of sugars are especially relevant. Why have so many pathogens come to rely on Sia binding or mimicry for colonization and/or invasion of host tissues? Has the host immune system adapted to meet these widespread pathogenic strategies? Can we use a bioinformatic approach to studying the sialome, and to orchestrate scientific research better in this important area of public health?

4.3 Diversity and Biology of Sialic Acids in Nature

4.3.1 Biological Importance of Sialic Acids

N-Acetylneuraminic acid (Neu5Ac) (Figure 4.2) is the most common Sia in mammals, and serves as a "core" structure that can be modified and presented as per the enzymatic repertoire of a particular cell (see Section 4.3.2). In mammals, the physical properties of Sias are known to be involved in processes such as nervous system plasticity and learning,
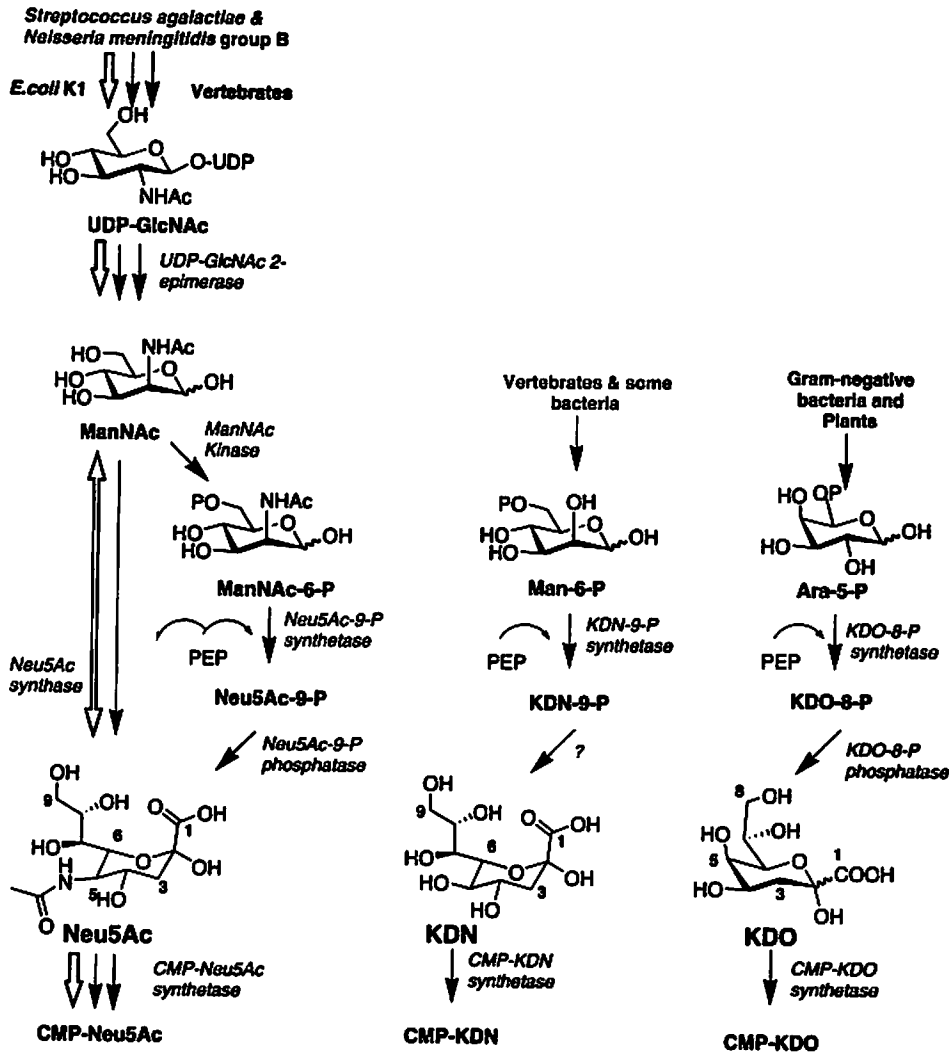
Figure 4.2 Biosynthetic pathways of sialic acids and sialic acid-like molecules. Shown on the left is the Neu5Ac biosynthesis pathway in vertebrates, which differs slightly from the pathways employed by bacteria. Biosynthetic pathways for Kdn and Kdo are shown for comparison. Refer to text for known evolutionary relationships.

kidney glomerular filtration, and repulsion between circulating blood cells [20–22]. Well-established immunological functions mediated by Sia-binding proteins include leukocyte trafficking by the selectins [23] and regulation of the alternative complement pathway activation by Factor H [24]. The Siglecs are a more recently discovered family of Sia-binding proteins that appear to mediate regulatory and/or phagocytic immune functions. [5, 12, 25]. Sialic acids are crucial for mammalian development, as evidenced by embryonic lethality of mice that lack a key biosynthetic enzyme [26]. Sialyltransferases can be regulated in a tissue-specific manner [27], and by stress [28, 29], infection [30], or malignancy [31]. Mammalian sialidases (enzymes that remove Sias) are also carefully regulated in normal and malignant mammalian cells [32, 33]. Sias are found in diverse structural contexts in Nature. This diversity exists on multiple levels, including modifications of core Sia structures, the linkage of Sia to an underlying sugar, the identity and arrangement of underlying sugars, structural attributes of the glycosylated molecule, and the higher level cellular and organismal milieu. Changes in Sia structural context can regulate Sia-dependent events by altering or preventing recognition by Sia-binding lectins and sialidases. The following sections briefly describe Sia diversity and distribution as they relate to Sia modifications/types and their representation among different species.

4.3.2 Sialic Acid Diversity in Nature

Sialic acids comprise a naturally occurring family of over 50 structures, many of unknown biological importance [16, 17]. *N*-Acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc) are the most common Sias in mammals, and serve as "core" structures that can be extensively modified. The other core Sia is Kdn (3-keto-2-deoxy-D-*manno*-nonulosonic acid). Although Kdo (2-keto-3-deoxy-D-*manno*-octulosonic acid) is not defined as a sialic acid, we have recently emphasized [17] that it is closely related both in structure and biosynthetic mechanism to Neu5Ac (discussed further in Section 4.6). CMP-Neu5Gc is derived from CMP-Neu5Ac via the enzymatic addition of an oxygen atom to the *N*-acetyl group [34]. Humans do not express Neu5Gc due to an inactivating mutation in the hydroxylase gene that mediates this step [35]. Neuraminic acid (Neu) is assumed (although not proven) to be derived from glycosidically-bound Neu5Ac by deacetylation of the *N*-acetyl group [36, 37]. Enzymatic alteration of Neu5Ac, Neu5Gc, Kdn or Neu can occur at various carbon positions [e.g. *O*-methylation (at C8), or esterification with acetyl (C4, C7, C8,or C9), lactyl (C9), sulfate (C8), or phosphate (C9) groups]. Further variability is derived by lactonization or lactamization of Sia structures [see [16, 17] for attempts at a comprehensive listing of chemical names, abbreviations, and reference publications of known Sia structures].

Sialic acid *O*-acetylation is a modification described in many biological contexts. While the biosynthetic mechanisms are still being elucidated, Sia *O*-acetylation is implicated in processes such as apoptotic regulation, colonic development and cancer, alternative complement pathway regulation, bacterial polysaccharide immunogenicity, and visceral leishmaniasis [38–43]. A mechanistic understanding of these and other *O*-acetylation-dependent processes has been hampered by the many failed attempts to clone the mammalian enzyme(s) responsible for this modification. Even in bacteria, definitive biochemical and genetic identification of Sia *O*-acetyltransferases was achieved only recently [19, 44]. Unfortunately for those trying to clone the mammalian enzymes, vertebrate genomes do not appear to

have any obvious homologs of these bacterial Sia *O*-acetyltransferases. Pre-existing work [16, 17] and unpublished data from our laboratory indicate extensive inter-species diversity in the expression profiles of vertebrate Sia *O*-acetylation. Thus, the biological roles of this modification are probably species specific in some instances.

4.3.3 Species Distribution of Sialic Acids

Understanding the distribution of Sias in nature requires two types of descriptions for different organisms, tissues or cells: (a) the presence of "core" Sia biosynthetic ability and (b) the presence of Sia-modifying enzymes. Echinoderms such as sea urchins and starfish were once thought to express the broadest diversity of Sias (i.e. the most Sia-modifying enzymes), with humans being the simplest [13]. With the advent of more sensitive techniques, it is now known that even human cells express a large variety of modified Sias, albeit in smaller quantities [45, 46]. Warren's pioneering work in developing and using the thiobarbituric acid assay to detect and quantify Sias suggested that the presence of "core" Sia biosynthetic ability was limited to certain species [47]. Specifically, Warren detected Sias in the deuterostome lineage of animals (vertebrates, and certain "higher" invertebrates such as echinoderms that literally have "two mouths" during development), but not in the protostomes (insects, mollusks, etc.). With the advent of very sensitive detection techniques and the ability to search genomes for genes involved in Sia synthesis, we now know that Sias have a much broader distribution than was originally thought [17]. They are expressed in octopus and squid (mollusks) [48], some insects (arthropods) [49, 50], and possibly even in some plants [51] – although the last finding has not been replicated by others [52, 53]. What is clear is that Sias are not ubiquitous among living organisms and certainly do not have the same biological roles in all contexts. This is particularly evident in the case of Sia-expressing pathogens, which apparently use this decoration to "hijack" a ride through host tissues.

4.4 Host Mechanisms for Evading Glycan-binding Pathogens

4.4.1 Is Sialic Acid Diversity a Reflection of Host Defense Against Past Invaders?

When pathogens exploit critical structures such as Sias, there are serious difficulties for the host. Does the structural diversity of Sia presentation reflect past host attempts at evading pathogens that bind to these structures? Consider the case of the Influenza virus hemagglutinin, which interacts with host cells in a Sia-dependent manner [54, 55]. The context of Sias varies widely between different animals, and dictates the specificity of Influenza virus binding. The host range of a particular Influenza virus is partly determined by Sia linkages (α2–3 or α2–6) and also by modifications (Neu5Ac, Neu5Gc, Neu5,9Ac$_2$) recognized by the virus hemagglutinin [55]. The critical importance of specific Sia recognition for viral evolution suggests that mammalian Sia structures have diverged (at least in part) to avoid host infection by Influenza viruses, harbored by other animals living in close proximity.

Another potential example of host evasion from pathogens by altering Sia expression involves Neu5Gc, a common Sia in nature that is not expressed by humans (see section on 4.3.2). Pathogens such as *E. coli* K99 can cause potentially lethal diarrheal infections by targeting host intestinal Neu5Gc residues [56]. One possibility in human evolution is that

a Neu5Gc-binding pathogen such as K99 eliminated all but a few individuals (who did not express Neu5Gc) in an early hominid population. Unfortunately, as in many such examples, it is nearly impossible to prove that a host species altered Sia linkages or modifications in order to escape viral or bacterial infection or limit cross-species transmission. Nevertheless, such discussion arouses interest about the effects of Sia diversity on pathogenic tropism for various species. Indeed, this tropism affects the "fitness" of both pathogens and hosts, and is broadly relevant to public health.

4.4.2 Mucins and Milk Oligosaccharides Can Prevent Infections

Mammals have evolved mucosal defense mechanisms against glycan-binding pathogens. Mucins and free oligosaccharides (OS) produced by mucosal tissues have long been hypothesized to have antibacterial properties. All mucosal surfaces exposed to microbes secret copious amounts of mucus, the main component of which is a family of heterogeneously and heavily glycosylated mucin molecules [57]. By providing multivalent arrays of glycan binding sites for pathogens, these act as "decoys", to prevent pathogens and toxins from reaching the cell surface, where they can initiate invasion. Such bound and trapped pathogens can also be eliminated mechanically, by expelling the mucus. Also, OS-based mucosal immunity transferred from mother to infant via breast milk appears to be important for the prevention of infant gastrointestinal infections. Human milk OS concentrations are estimated at $\sim 10\,\mathrm{g\,l^{-1}}$ [58] and mass spectrometric studies indicate >900 different OS species [59]. Compositional analyses of human breast milk and infant stool indicate that (uncharged) milk OS not only survive the infant digestive tract, but are actually present in higher concentrations than in mothers' milk [60]. This finding suggests that (neutral) milk OS may not have a primarily nutritive role; rather, they are in the right place at the right time to inhibit binding of potential pathogens. It has been argued that the considerable maternal resources expended on milk OS production must have an important biological role [61]. Certainly, the benefits of breast-feeding for protection from infant infection are well documented, but how much of the effect can we attribute to oligosaccharides? Is the expression of milk glycans an evolutionary response to the common pathogen strategy of glycan-binding? In this regard, it is interesting that the most complex mixtures of milk OS are typically found in large-brained social animals such as humans and elephants, which have relatively immature young, needing prolonged care before they become independent [62].

Interestingly, most of the OS in human milk have terminal fucose and/or Sia residues, both of which are commonly used by various pathogens for targeted entry. Studies indicate that the levels of $\alpha1$–2-linked fucosylated OS in milk correlate inversely with the risk of diarrheal infection caused by fucose-binding pathogens such as *Campylobacter* species, *Vibrio cholerae*, *E. coli* stable toxin, and calciviruses to target cells [63]. Similarly to fucosylated milk OS, some studies suggest that sialylated milk OS and mucins may inhibit Sia-dependent interactions between mammalian cells and pathogens. For example, sialidase treatment of mucins dramatically reduced their inhibitory activity against *Helicobacter pylori* [64] and *Haemophilus influenzae* [65]. Another study describes a Sia-dependent resistance to adenoviral gene transfer mediated by the mucin MUC1 [66]. Indeed, mucins are upregulated during infectious challenges of both murine [67] and human cells [68]. Furthermore, binding and replication of rotavirus, the major cause of severe

dehydrating diarrhea in young children, is inhibited in a Sia-dependent manner by a particular mucin in human milk [69]. Sialylated gangliosides and OS in milk also inhibit binding of several *E. coli* strains (or enterotoxins) that cause diarrhea and urinary tract infections [70, 71]. *Vibrio cholerae* and *E. coli* Sia-binding toxins are also inhibited by a Sia-rich ganglioside fraction of human milk [71]. The prevalence of Sia-binding pathogens and the abundance of sialylated glycans in breast-milk and mucins suggest that these glycans may have an even broader protective role against Sia-binding pathogens than is currently recognized.

4.5 Pathogens Exchange Glycan Biosynthesis Genes, Allowing Molecular Mimicry and Hijacking of Host Lectins

4.5.1 The Diversity of Bacterial Polysaccharides

Bacterial surface glycans exhibit remarkable diversity and in many cases influence pathogenicity. Bacterial glycans come in the form of cell-wall peptidoglycans, polysaccharide capsules, glycoproteins, or glycolipids. Capsules (sometimes referred to as "K-antigens") can be elaborated by both Gram-positive and Gram-negative bacteria and are comprised of single-monosaccharide homopolymers or repeating units of many monosaccharide types. In contrast, bacterial glycolipids are only found in the outer membranes of Gram-negative bacteria and are either referred to as lipopolysaccharides (LPS) or lipooligosaccharides (LOS). The core structure of bacterial glycolipids is often modified with outer "*O*-antigens," which can vary widely. As an example, *Escherichia coli* has well over 70 capsular polysaccharides and over 170 known *O*-antigen immunotypes. Likewise, *Streptococcus pneumoniae* can express over 90 different strain-specific capsular polysaccharides [72]. Pathogenic bacteria are thus adept at evolving or acquiring the machinery for synthesizing host-like structures such as sialic acids. Below we compare Sia biosynthesis pathways in vertebrates and bacteria and discuss known and potential mechanisms that make Sia decoration a successful pathogenic mechanism.

4.5.2 Vertebrates and Bacteria Use Phylogenetically-related Sialic Acid Biosynthesis Pathways

The biosynthesis of free Neu5Ac varies slightly depending on whether it is happening in a vertebrate or bacterial cell [17] (Figure 4.2). In vertebrates, UDP-GlcNAc is converted to ManNAc by UDP-GlcNAc 2-epimerase/ManNAc kinase, a bifunctional enzyme, which also phosphorylates ManNAc to give ManNAc-6-phosphate [73, 74], which is then converted into Neu5Ac-9-phosphate by Neu5Ac-9-phosphate synthetase [75, 76]. Bacteria obtain ManNAc in the same way as vertebrates, by epimerization of UDP-GlcNAc [77, 78]. *Neisseria meningitidis* was postulated in one study to obtain ManNAc by epimerization of GlcNAc-6-phosphate (via gene product SiaA), followed by dephosphorylation [79]. A subsequent study, however, demonstrated that *N. meningitidis* SiaA catalyzes the epimerization of UDP-GlcNAc to ManNAc [78]. In contrast to animals, bacteria synthesize Neu5Ac via a Neu5Ac synthetase directly from ManNAc, rather than ManNAc-6-phosphate [80, 81]. In both bacteria and animals, activation of Neu5Ac is accomplished by converting Neu5Ac to CMP-Neu5Ac using CTP (cytidine 5'-triphosphate) and a CMP-Neu5Ac synthetase [82].

Despite the biochemical differences in Sia biosynthesis between vertebrates and bacteria, genes encoding the responsible enzymes are evolutionarily related. Moreover, phylogenetic analysis of Neu5Ac and CMP-Neu5Ac synthetases suggests that there have been multiple horizontal transfers of Sia biosynthesis genes among bacteria [17]. Unlike the enzymes involved in Sia biosynthesis, bacterial sialyltransferases do not appear to be related to those of animals and were likely "reinvented", on multiple occasions. Alternatively, pathogens without complete biosynthetic pathways have devised multiple ways to "steal" sialic acids from their hosts [10], employing truncated Sia pathways that allow scavenging of host Sias for incorporation into their polysaccharides [10].

4.5.3 Microbial Sialic Acid Hijacks Host Factor H

The presence of Sias in microbial polysaccharides is often associated with pathogenicity in humans. Microbial expression of Sia enhances serum-resistance due to inactivation of the alternative complement pathway [83, 84], a tightly regulated proteolytic cascade that mediates an immediate antibody-independent response to foreign entities, and constitutes one of the frontline defenses against many invading pathogens [24]. Factor H is a regulatory component of the alternative complement pathway, which binds to a number of "self" components including sialic acid, and down-regulates activation of the pathway [85]. In this way, Factor H serves a critical function by preventing self-reactivity (i.e. "friendly fire"). Unfortunately, the otherwise useful molecular assumption that "Sia = self" becomes a liability when sialylated pathogens invade the body. The common mammalian terminal trisaccharide Neu5Acα-2-3Galβ1-(3/4)GlcNAc appears to be very common among pathogenic bacterial polysaccharide capsules and lipooligosaccharides (LOS) and may reflect bacterial optimization for binding to Factor H or other host Sia-binding proteins, such as Siglecs (see below) [11].

4.5.4 Is Siglec Hijacking Responsible for the Rapid Evolution of This Protein Family?

Siglecs are a large family of Sia-binding proteins that are expressed, with some exceptions, on cells of the immune system [5, 86, 87]. The precise contexts in which many Siglecs function remain largely unknown, although some clearly have important endogenous functions such as regulation of B-cell stimulation [88] and the maintenance of myelin [89]. A group of closely related Siglecs (known as the CD33-related Siglecs) are capable of regulating immune cells *in vitro* [90–92], as well as *in vivo* [93].

The CD33-related Siglecs are undergoing rapid evolution that is particularly evident in their amino-terminal V-set Ig-like Sia-binding domains [94]. This evolution likely reflects multiple selective pressures, including the human-specific loss of Neu5Gc. Indeed, it appears that the Sia-binding preference of at least one human Siglec (Siglec-9) has shifted from Neu5Gc towards Neu5Ac, when compared with the chimpanzee ortholog [95]. The loss of Neu5Gc may (partly) explain accelerated evolution in the human lineage, but the fact that Siglecs are evolving rapidly in multiple lineages requires further explanation. We have suggested that Sia-binding and Sia-expressing pathogens increased the rate of evolution among Siglecs in distinct, but interdependent, ways [5, 94].

Sia-binding and Sia-expressing mechanisms of pathogenesis not only apply to humans, but appear to be fairly ancient. Indeed, Sia binding and Sia expression are used by pathogens that infect a wide range of vertebrate host species. Pathogenic Sia expression may result in direct interactions with Siglecs, possibly to hijack inhibitory roles of these proteins in the host [12]. Through many generations and/or epidemics, Siglecs may have diverged from recognizing particular contexts of pathogen-expressed Sias – a theoretical example of the "Red Queen effect." In the case of Sia-binding pathogens, particular host Sias may have diverged to avoid pathogen recognition (see Section 4.4.1). Such changes in host Sia structure(s) may have created the need for Siglec divergence in order to preserve endogenous function(s). This hypothetical situation embodies a concept recently referred to as a "Secondary Red Queen effect" [5].

In a variation on this theme, two porcine viruses [reproductive and respiratory syndrome virus (RRSR) and arterivirus], apparently exploit the host Sia-recognizing lectin (Siglec-1/sialoadhesin) for invasion of macrophages, via recognition of the Sias on the viral cell membranes [96, 97].

4.5.5 Enzymes That Release, Destroy, or Alter Sialic Acids

There are many sialidases encoded by both host and pathogen that remove Sia molecules. In addition, enzymes such as lyases and esterases destroy Sias or Sia modifications, respectively (Figure 4.1). Microbial removal of host Sia is often related to survival in the host. For example, *Streptococcus pneumonia* (pneumococcus*)* is a pneumonia-causing agent and a leading cause of fatal infections in the very young and the very old. Pneumococcus expresses a sialidase (neuraminidase) that exposes underlying Galβ1–4GlcNAc residues and increases bacterial adhesion [98, 99]. Another example of a neuraminidase-expressing pathogen is Influenza virus, which has both a Sia-binding hemagglutinin *and* a neuraminidase, both of which are critical for replication and spread of the virus. Some bacteria are also able to use Sias as an energy source, and may encode sialidases to release nutritional value that is 'locked up' on the host cell surface [100–102]. Although it appears that host and pathogen sialidases share a common evolutionary origin, the roles of endogenous (host) sialidases are not as well understood as those of their microbial counterparts [103, 104]. Just as has been shown for Sia-binding lectins, many sialidases prefer particular Sia structures; the most commonly described situation is to find that sialidase action is hindered by *O*-acetylation [105, 106].

4.6 Evolution of Sialic Acids

Evolutionary relationships between biosynthetic pathways are often inferred by phylogenetic, structural, and/or mechanistic studies of their component enzymes. Some such studies are lacking for particular Sia or Sia-like molecules. Nonetheless, preliminary analyses suggest the biosynthetic pathways of Neu5Ac, Kdn, Kdo, and other 'sialic acid-like' molecules are at least partly homologous [17].

4.6.1 Neu5Ac versus Kdo

A bacterial sugar called Kdo (2-keto-3-deoxy-D-*manno*-octulosonic acid) bears a close resemblance to Sias (structures and biosynthetic pathways are shown in Figure 4.2). Kdo

is found as part of the "core" structure of Gram-negative bacterial lipopolysaccharides (LPS) and also in plant cell walls. Unlike the differences in Neu5Ac biosynthesis between bacteria and vertebrates (see Figure 4.2 and text above), it appears that the Kdo pathway in Gram-negative bacteria and plants does not differ substantially. Indeed, a Gram-negative mutant deficient in Kdo-8-phosphate synthetase was complemented by the orthologous enzyme from a pea plant [107]. Phylogenetic examination of CMP-Kdo synthetases indicates that this gene underwent horizontal gene transfer from bacteria to plants prior to or shortly after the divergence of plants from other eukaryotes [108]. There is also significant sequence similarity between CMP-Neu5Ac and CMP-Kdo synthetases, confirming their homologous relationship. These synthetases likely diverged by ancient gene duplication, as evidenced by their separate clustering in a phylogenetic tree [17]. Unfortunately, the exact timing of the presumed gene duplication leading to Neu5Ac and Kdo CMP-synthetases is difficult to infer. Analysis of the biosynthetic step prior to CMP activation indicates that Neu5Ac-9-P and Kdo-8-P synthetases do *not* show much sequence similarity; however, both reactions proceed by a similar mechanism that employs a TIM barrel fold, and is also shared by DHAP synthase (3-deoxy-D-*arabino*-heptulosonate-7-phosphate [109]. Recently, the structure of a bacterial Neu5Ac synthetase was also shown to have the TIM barrel fold [110]. Of course, the common TIM barrel fold shared by these enzymes is a very stable fold that is used by a large number of enzymes that perform many different functions [111].

4.6.2 Biosynthetic Machinery for Kdn, Legionaminic Acid, and Pseudaminic Acid

Although the enzymes of Kdn biosynthesis have not been cloned and characterized, biochemical studies indicate that the pathway intermediates are analogous to those of Neu5Ac biosynthesis, but are mostly catalyzed by distinct enzymes [112]. The essential difference between Kdn and Neu5Ac biosynthesis is that the Kdn pathway begins with mannose-6-phosphate (Man-6-P) rather than ManNAc. Interestingly, both human and *Drosophila* Neu5Ac-9-P synthetase enzymes can accept Man-6-P to generate Kdn-9-P [113, 114]. Studies of rat Neu5Ac-9-P synthetase, however, indicate that it does not accept Man-6-P [115]. Discovery of the genetic basis for Kdn biosynthesis will allow further characterization of both the evolutionary history and biological importance of this sialic acid.

Other Sia-like molecules produced by microbes include pseudaminic acid (5,7-diamino-3,5,7,9 -tetradeoxy-L-*glycero*-L-*manno*-nonulosonic acid) and legionaminic acid (5,7-diamino-3,5,7,9-tetradeoxy-D-*glycero*-D-*galacto*-nonulosonic acid). Whereas Kdo is found in all Gram-negative bacteria and plants, pseudaminic and legionaminic acids have so far been chemically identified in only a relatively small number of bacteria [19]. Surprisingly, bacterial synthetases (whether Neu5Ac, legionaminic acid, or pseudaminic acid) share 30–35% identity at the amino acid level when compared with the human Neu5Ac synthetase. All-in-all, although the enzymes of Sia biosynthesis appear to be related, an understanding of their ancient history will likely require a phylogenetic approach based on structure or structural modeling, not just on sequence identity [116]. Recent advances in our understanding of pseudaminic acid biosynthesis will be of particular interest towards this end [117].

4.7   Current Status and Future Directions of Sialic Acid Bioinformatics

4.7.1   Existing Glycan Databases Often Misrepresent Sialic Acids

Sialic acid linkages and modifications are somewhat delicate (i.e. susceptible to hydrolysis) compared with other carbohydrates. Methods for the release and analysis of glycans often damage Sias or their modifications (see Table 4.1); thus, Sias are often inaccurately depicted in glycan databases. In practice, there has been an unfortunate tendency to assume that any Sia that has been removed by mild acid or by a sialidase, or detected by virtue of its negative charge, must be *N*-acetylneuraminic acid. In fact, in most instances where the literature or databases indicate the presence of Neu5Ac (or other acronyms such as NeuAc, NeuNAc, or NANA), the actual native Sia structure remains unknown. For example, Sia modifications such as *O*-acetylation, *O*-lactylation, and sulfation are sensitive to pH and are often inadvertently removed during glycan release, purification, or analysis. Also, Sia *O*-acetyl and *N*-glycolyl groups are sensitive to the conditions employed during hydrazinolysis and methylation analysis (Table 4.1). *O*-Acetyl esters at the C7 position are known to migrate along the Sia side-chain to C9, even under physiological conditions [118, 119]. Hence the detection of a 9-*O*-acetylated Sia should be cautiously interpreted as "9(7)-*O*-acetylated" until further characterization reveals the native location.

Overall, it is our recommendation that the term "Sia" (rather than Neu5Ac) be used whenever the type of Sia at a particular position on a glycan has not been definitively determined. All existing databases need to be "cleaned up" following the same convention. Table 4.1 lists several glycan analysis techniques that employ conditions likely to damage Sias or the Sia modifications mentioned above; it also provides some practical suggestions for retaining native Sia structures.

4.7.2   The Need for a "Sialome" Project

The myriad glycan structures found in nature represent a vast and expanding biological frontier. As elaborated throughout this chapter, Sias are an extreme example of glycan diversity, both in structure and in function. We have recently defined the term "Sialome" as "the total complement of Sia types and linkages and their modes of presentation on a particular organelle, cell, tissue, organ, or organism – as found at a particular time and under specific conditions" [5].

In other words, a Sialome can be thought of as reflecting at least six levels of complexity:

1. Different possible "core" Sia structural variations at the 5-position of Neu, Neu5Ac, Neu5Gc, or Kdn.
2. Various modifications of the above, sometimes in combinations (see Section 4.3.2).
3. Differing linkages of the Sia to the underlying glycan (mostly α2–3 or α2–6 to various sugars, or α2–8 or α2–9 to underlying Sias).
4. The precise identity and arrangement of glycans immediately below the Sia.
5. Structural attributes of the underlying glycan (e.g. *N*-linked or *O*-linked to protein, lipid- or GPI-anchored, cell-associated, or secreted).
6. The higher level cellular, organismal, and environmental milieu (each of these ideally deserve their own category).

Table 4.1   Ex

Method
Endoglycosida e.g. PNGase *N*-glycosidas of *N*-glycans enzyme exhi specificity an commonly us *N*-glycan ans
Hydrazinolysis of *N*- and *O*-l glycans. Mor used for *O*-gl there are no k broad-spectru *O*-glycanases
Use of ~0.05–0 (trifluoroacet) applications s dialysis or col
Base treatment to release *O*-g sometimes to GPI-anchored bacterial poly e.g. Group B streptococcal polysaccharid through phos cell-wall pept
Thiobarbituric assay for Sia
Fluorescent der of monosacch (DMB-sialic reducing term oligosaccharid and 2-AMAC HPLC–fluores resolution and of different Si
Various

*If there is any

Table 4.1   Examples of methods for studying glycans and their negative impact on sialic acids.

Method	Effect on Sia	Recommendation*
Endoglycosidase release, e.g. PNGase F (peptide: *N*-glycosidase F) release of *N*-glycans. This enzyme exhibits broad specificity and is very commonly used in *N*-glycan analysis	Sia *O*-acetylation susceptible to migration and/or hydrolysis at the pH optimum of this enzyme (8.6), especially during prolonged incubations	Use larger amounts of enzyme at pH 7.5 for short time periods to retain *O*-acetylation
Hydrazinolysis for release of *N*- and *O*-linked glycans. More commonly used for *O*-glycans, since there are no known broad-spectrum. Peptide: *O*-glycanases	Results in loss of both *N*- and *O*-acetyl groups (and likely some others)	Products of this reaction are often subject to re-*N*-acetylation without regard for the possibility that *N*-glycolyl could have been the native structure or that *O*-modifications are not replaced
Use of ~0.05–0.1% TFA (trifluoroacetic acid) in applications such as dialysis or column elution	Can result in desialylation and may damage some Sia modifications	If necessary for separation from contaminants, freeze-dry immediately and interpret cautiously
Base treatment often used to release *O*-glycans, and sometimes to release GPI-anchored glycans or bacterial polysaccharides, e.g. Group B streptococcal (GBS) polysaccharide, linked through phosphodiester to cell-wall peptidoglycan	Harsh basic conditions are certain to destroy some Sia modifications which go unreported, e.g. *O*-acetylation of Sias on GBS capsule went unreported for ~25 years	Use alternatives to base-treatment to retain native structure or interpret cautiously
Thiobarbituric acid (TBA) assay for Sia quantitation	Assay relies on periodate-oxidation of the Sia side-chain, a reaction that is impeded by the presence of *O*-acetylation and other modifications of Sia at C9	The TBA assay does not quantitate modified Sias unless modifications are first removed (i.e. remove *O*-acetylation by base treatment)
Fluorescent derivatization of monosaccharides (DMB-sialic acid) and reducing termini of oligosaccharides (2-AB and 2-AMAC) for HPLC–fluorescence resolution and detection of different Sia species	Derivatization often employed under acidic conditions, which could result in some loss or migration of Sia modifications. Not systematically studied	Keep reactions cold after derivatization and analyze as soon as possible
Various	The *O*-acetyl ester modification of Neu5Ac migrates from C7 to C9 under physiological and experimental conditions (e.g. below pH 3–4 and above pH 8, especially for prolonged periods and/or high temperatures)	Avoid conditions under which migration is accelerated. Dry samples and freeze for storage or use immediately for best results. When position unknown, use "9(7)-*O*-acetylation"

*If there is any concern of damage during release or analysis, indicate "Sia" rather than "Neu5Ac."

Indeed, there are many factors that influence the presentation of Sias on cell surfaces, including normal developmental processes and pathological states such as inflammation, infection, and cancer – and many of these are cell-type and/or species-specific. Carefully constructed Sia databases will likely shed light on mechanisms of pathogen tropism for various tissues or species and may provide further insight into pathogenic mechanisms of *in*vasion and *e*vasion mediated by Sias. Given the diverse and extensive involvement of Sias in biological processes and disease states, these resources would no doubt provide insights into many other areas of interest. Despite the clear benefits of doing so, there has not been a concerted effort to collect Sia information into a comprehensive user-friendly database. Certainly, building databases that contain detailed information about glycans is a tedious task that must reflect methodological limitations of the past and present. We suggest that the field of Sia research needs two databases of Sia information in relation to particular biological and species contexts: one which would allow cataloguing of Sia types (points 1 and 2 above) and a second that would provide an overall idea of Sia linkages and modes of presentation (points 3–5 above). Information could be input and searched based on methodology or biological context. Such a "Sialome project" could have great predictive value in biological studies aimed at unraveling the functional significance and evolutionary history of this curious group of molecules.

References

1. Varki A: Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 1993, 3:97–130.
2. Gagneux P, Varki A: Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 1999, 9:747–755.
3. Varki A: Nothing in glycobiology makes sense, except in the light of evolution. *Cell* 2006, 126:841–845.
4. Bishop JR, Gagneux P: Evolution of carbohydrate antigens – microbial forces shaping host glycomes? *Glycobiology* 2007, 17:23R–34R.
5. Varki A, Angata T: Siglecs – the major subfamily of I-type lectins. *Glycobiology* 2006, 16:1R–27R.
6. Mammen M, Choi S-K, Whitesides GM: Polyvalent Interactions in Biological Systems: Implications for Design and Use of Mutlivalent Ligands and Inhibitors. *Angew Chem Int Ed* 1998, 37:2754–2794.
7. Ilver D, Johansson P, Miller-Podraza H, Nyholm PG, Teneberg S, Karlsson KA: Bacterium–host protein–carbohydrate interactions. *Methods Enzymol* 2003, 363:134–157.
8. Wills C, Green DR: A genetic herd-immunity model for the maintenance of MHC polymorphism. *Immunol Rev* 1995, 143:263–292.
9. Lindesmith L, Moe C, Marionneau S, Ruvoen N, Jiang X, Lindblad L, Stewart P, LePendu J, Baric R: Human susceptibility and resistance to Norwalk virus infection. *Nat Med* 2003, 9:548–553.
10. Vimr E, Lichtensteiger C: To sialylate, or not to sialylate: that is the question. *Trends Microbiol* 2002, 10:254–257.
11. Carlin AF, Lewis AL, Varki A, Nizet V: Group B streptococcal capsular sialic acids interact with Siglecs (immunoglobulin-like lectins) on human leukocytes. *J Bacteriol* 2007, 89:1231–1237.
12. Varki A: Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* 2007, 446:1023–1029.

13. Schauer R: Chemistry, metabolism, and biological functions of sialic acids. *Adv Carbohydr Chem Biochem* 1982, **40**:131–234.
14. Varki A: Diversity in the sialic acids. *Glycobiology* 1992, **2**:25–40.
15. Troy FA: Polysialylation: from bacteria to brains. *Glycobiology* 1992, **2**:5–23.
16. Kelm S, Schauer R: Sialic acids in molecular and cellular interactions. *Int Rev Cytol* 1997, **175**:137–240.
17. Angata T, Varki A: Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev* 2002, **102**:439–469.
18. Bouchet V, Hood DW, Li J, Brisson JR, Randle GA, Martin A, Li Z, Goldstein R, Schweda EK, Pelton SI, *et al.*: Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proc Natl Acad Sci USA* 2003, **100**:8898–8903.
19. Lewis AL, Desa N, Hansen EE, Knirel Y, Gordon JI, Gagneux P, Nizet V, Varki A: Innovations in host and microbial sialic acid biosynthesis revealed by genomic and phylogenetic prediction of nonulosonic acid structure. *Proc Natl Acad Sci USA* 2009. (in press).
20. Kerjaschki D, Vernillo AT, Farquhar MG: Reduced sialylation of podocalyxin – the major sialoprotein of the rat kidney glomerulus – in aminonucleoside nephrosis. *Am J Pathol* 1985, **118**:343–349.
21. Fujimoto I, Bruses JL, Rutishauser U: Regulation of cell adhesion by polysialic acid. *J Biol Chem* 2001, **276**:31745–31751.
22. Ronn LC, Berezin V, Bock E: The neural cell adhesion molecule in synaptic plasticity and ageing. *Int J Dev Neurosci* 2000, **18**:193–199.
23. Ley K: The role of selectins in inflammation and disease. *Trends Mol Med* 2003, **9**:263–268.
24. Pangburn MK: Host recognition and target differentiation by factor H, a regulator of the alternative pathway of complement. *Immunopharmacology* 2000, **49**:149–157.
25. Crocker PR, Paulson JC, Varki A: Siglecs and their roles in the immune system. *Nat Rev Immunol* 2007, **7**:255–266.
26. Schwarzkopf M, Knobeloch KP, Rohde E, Hinderlich S, Wiechens N, Lucka L, Horak I, Reutter W, Horstkorte R: Sialylation is essential for early development in mice. *Proc Natl Acad Sci USA* 2002, **99**:5267–5270.
27. Martin LT, Marth JD, Varki A, Varki NM: Genetically altered mice with different sialyltransferase deficiencies show tissue-specific alterations in sialylation and sialic acid 9-*O*-acetylation. *J Biol Chem* 2002, **277**:32930–32938.
28. Jamieson JC, McCaffrey G, Harder PG: Sialyltransferase: a novel acute-phase reactant. *Comp Biochem Physiol B* 1993, **105**:29–33.
29. Dabelic S, Flogel M, Maravic G, Lauc G: Stress causes tissue-specific changes in the sialyltransferase activity. *Z Naturforsch, Teil C* 2004, **59**:276–280.
30. Knight PA, Pemberton AD, Robertson KA, Roy DJ, Wright SH, Miller HR: Expression profiling reveals novel innate and inflammatory responses in the jejunal epithelial compartment during infection with *Trichinella spiralis*. *Infect Immun* 2004, **72**:6076–6086.
31. Fukushima K, Hara-Kuge S, Seko A, Ikehara Y, Yamashita K: Elevation of alpha2→6 sialyltransferase and alpha1→2 fucosyltransferase activities in human choriocarcinoma. *Cancer Res* 1998, **58**:4301–4306.
32. Stamatos NM, Liang F, Nan X, Landry K, Cross AS, Wang LX, Pshezhetsky AV: Differential expression of endogenous sialidases of human monocytes during cellular differentiation into macrophages. *FEBS J* 2005, **272**:2545–2556.
33. Miyagi T, Wada T, Yamaguchi K, Hata K: Sialidase and malignancy: a minireview. *Glycoconj J* 2004, **20**:189–198.
34. Kawano T, Koyama S, Takematsu H, Kozutsumi Y, Kawasaki H, Kawashima S, Kawasaki T, Suzuki A: Molecular cloning of cytidine monophospho-*N*-acetylneuraminic acid hydroxylase.

Regulation of species- and tissue-specific expression of *N*-glycolylneuraminic acid. *J Biol Chem* 1995, **270**:16458–16463.

35. Hayakawa T, Satta Y, Gagneux P, Varki A, Takahata N: Alu-mediated inactivation of the human CMP-*N*-acetylneuraminic acid hydroxylase gene. *Proc Natl Acad Sci USA* 2001, **98**:11399–11404.

36. Manzi AE, Sjoberg ER, Diaz S, Varki A: Biosynthesis and turnover of *O*-acetyl and *N*-acetyl groups in the gangliosides of human melanoma cells. *J Biol Chem* 1990, **265**:13091–13103.

37. Sonnenburg JL, Van Halbeek H, Varki A: Characterization of the acid stability of glycosidically linked neuraminic acid – use in detecting de-*N*-acetyl-gangliosides in human melanoma. *J Biol Chem* 2002, **277**:17502–17510.

38. Chen HY, Varki A: *O*-Acetylation of GD3: an enigmatic modification regulating apoptosis? *J Exp Med* 2002, **196**: 1529–1533.

39. Shen Y, Tiralongo J, Kohla G, Schauer R: Regulation of sialic acid *O*-acetylation in human colon mucosa. *Biol Chem* 2004, **385**: 145–152.

40. Shen Y, Kohla G, Lrhorfi AL, Sipos B, Kalthoff H, Gerwig GJ, Kamerling JP, Schauer R, Tiralongo J: *O*-Acetylation and de-*O*-acetylation of sialic acids in human colorectal carcinoma. *Eur J Biochem* 2004, **271**:281–290.

41. Shi WX, Chammas R, Varki NM, Powell L, Varki A: Sialic acid 9-*O*-acetylation on murine erythroleukemia cells affects complement activation, binding to I-type lectins, and tissue homing. *J Biol Chem* 1996, **271**:31526–31532.

42. Berry DS, Lynn F, Lee CH, Frasch CE, Bash MC: Effect of *O*-acetylation of *Neisseria meningitidis* serogroup A capsular polysaccharide on development of functional immune responses. *Infect Immun* 2002, **70**:3707–3713.

43. Chava AK, Bandyopadhyay S, Chatterjee M, Mandal C: Sialoglycans in protozoal diseases: their detection, modes of acquisition and emerging biological roles. *Glycoconj J* 2004, **20**: 199–206.

44. Deszo EL, Steenbergen SM, Freedberg DI, Vimr ER: *Escherichia coli* K1 polysialic acid *O*-acetyltransferase gene, neuO, and the mechanism of capsule form variation involving a mobile contingency locus. *Proc Natl Acad Sci USA* 2005, **102**:5564–5569.

45. Zanetta JP, Pons A, Iwersen M, Mariller C, Leroy Y, Timmerman P, Schauer R: Diversity of sialic acids revealed using gas chromatography/mass spectrometry of heptafluorobutyrate derivatives. *Glycobiology* 2001, **11**:663–676.

46. Bulai T, Bratosin D, Pons A, Montreuil J, Zanetta JP: Diversity of the human erythrocyte membrane sialic acids in relation with blood groups. *FEBS Lett* 2003, **534**:185–189.

47. Warren L: The distribution of sialic acids in Nature. *Comp Biochem Physiol* 1963, **10**:153–171.

48. Saito M, Kitamura H, Sugiyama K: Occurrence of gangliosides in the common squid and Pacific octopus among protostomia. *Biochim Biophys Acta Bio-Membr* 2001, **1511**:271–280.

49. Roth J, Kempf A, Reuter G, Schauer R, Gehring WJ: Occurrence of sialic acids in *Drosophila melanogaster*. *Science* 1992, **256**:673–675.

50. Malykh YN, Krisch B, Gerardy-Schahn R, Lapina EB, Shaw L, Schauer R: The presence of *N*-acetylneuraminic acid in Malpighian tubules of larvae of the cicada *Philaenus spumarius*. *Glycoconj J* 1999, **16**:731–739.

51. Shah MM, Fujiyama K, Flynn CR, Joshi L: Sialylated endogenous glycoconjugates in plant cells. *Nat Biotechnol* 2003, **21**:1470–1471.

52. Seveno M, Bardor M, Paccalet T, Gomord V, Lerouge P, Faye L: Glycoprotein sialylation in plants? *Nat Biotechnol* 2004, **22**:1351–1352.

53. Zeleny R, Kolarich D, Strasser R, Altmann F: Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta* 2006, **224**:222–227.

54. Couceiro JNSS, Paulson JC, Baum LG: Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. *Virus Res* 1993, **29**:155–165.

55. Suzuki Y: Sialobiology of influenza: molecular mechanism of host range variation of influenza viruses. *Biol Pharm Bull* 2005, **28**:399–408.

56. Kyogashima M, Ginsburg V, Krivan HC: *Escherichia coli* K99 binds to *N*-glycolylsialoparagloboside and *N*-glycolyl-GM3 found in piglet small intestine. *Arch Biochem Biophys* 1989, **270**:391–397.

57. Hollingsworth MA, Swanson BJ: Mucins in cancer: protection and control of the cell surface. *Nat Rev Cancer* 2004, **4**:45–60.

58. Newburg DS, Ruiz-Palacios GM, Altaye M, Chaturvedi P, Meinzen-Derr J, Guerrero MM, Morrow AL: Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhea in breastfed infants. *Glycobiology* 2004, **14**:253–263.

59. Stahl B, Thurl S, Zeng J, Karas M, Hillenkamp F, Steup M, Sawatzki G: Oligosaccharides from human milk as revealed by matrix-assisted laser desorption/ionization mass spectrometry. *Anal Biochem* 1994, **223**:218–226.

60. Chaturvedi P, Warren CD, Buescher CR, Pickering LK, Newburg DS: Survival of human milk oligosaccharides in the intestine of infants. *Adv Exp Med Biol* 2001, **501**:315–323.

61. Newburg DS: Innate immunity and human milk. *J Nutr* 2005, **135**:1308–1312.

62. Urashima T, Saito T, Nakamura T, Messer M: Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconj J* 2001, **18**:357–371.

63. Morrow AL, Ruiz-Palacios GM, Jiang X, Newburg DS: Human-milk glycans that inhibit pathogen binding protect breast-feeding infants against infectious diarrhea. *J Nutr* 2005, **135**:1304–1307.

64. Hirmo S, Kelm S, Iwersen M, Hotta K, Goso Y, Ishihara K, Suguri T, Morita M, Wadström T, Schauer R: Inhibition of *Helicobacter pylori* sialic acid-specific haemagglutination by human gastrointestinal mucins and milk glycoproteins. *FEMS Immunol Med Microbiol* 1998, **20**:275–281.

65. Solzbacher D, Hanisch FG, van Alphen L, Gilsdorf JR, Schroten H: Mucin in middle ear effusions inhibits attachment of *Haemophilus influenzae* to mucosal epithelial cells. *Eur Arch Otorhinolaryngol* 2003, **260**:141–147.

66. Arcasoy SM, Latoche J, Gondor M, Watkins SC, Henderson RA, Hughey R, Finn OJ, Pilewski JM: MUC1 and other sialoglycoconjugates inhibit adenovirus-mediated gene transfer to epithelial cells. *Am J Respir Cell Mol Biol* 1997, **17**:422–435.

67. Holmen JM, Olson FJ, Karlsson H, Hansson GC: Two glycosylation alterations of mouse intestinal mucins due to infection caused by the parasite *Nippostrongylus brasiliensis*. *Glycoconj J* 2002, **19**:67–75.

68. McNamara N, Basbaum C: Signaling networks controlling mucin production in response to Gram-positive and Gram-negative bacteria. *Glycoconj J* 2001, **18**:715–722.

69. Yolken RH, Peterson JA, Vonderfecht SL, Fouts ET, Midthun K, Newburg DS: Human milk mucin inhibits rotavirus replication and prevents experimental gastroenteritis. *J Clin Invest* 1992, **90**:1984–1991.

70. Martín-Sosa S, Martín MJ, Hueso P: The sialylated fraction of milk oligosaccharides is partially responsible for binding to enterotoxigenic and uropathogenic *Escherichia coli* human strains. *J Nutr* 2002, **132**:3067–3072.

71. Otnaess AB, Laegreid A, Ertresvag K: Inhibition of enterotoxin from *Escherichia coli* and *Vibrio cholerae* by gangliosides from human milk. *Infect Immun* 1983, **40**:563–569.

72. Weintraub A: Immunology of bacterial polysaccharide antigens. *Carbohydr Res* 2003, **338**:2539–2547.

73. Hinderlich S, Stäsche R, Zeitler R, Reutter W: A bifunctional enzyme catalyzes the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver – purification and characterization of UDP-*N*-acetylglucosamine 2-epimerase/*N*-acetylmannosamine kinase. *J Biol Chem* 1997, **272**:24313–24318.

74. Stäsche R, Hinderlich S, Weise C, Effertz K, Lucka L, Moormann P, Reutter W: A bifunctional enzyme catalyzes the first two steps in *N*-acetylneuraminic acid biosynthesis of rat liver – molecular cloning and functional expression of UDP-*N*-acetyl-glucosamine 2-epimerase/ *N*-acetylmannosamine kinase. *J Biol Chem* 1997, **272**:24319–24324.

75. Lawrence SM, Huddleston KA, Pitts LR, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ: Cloning and expression of the human *N*-acetylneuraminic acid phosphate synthase gene with 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid biosynthetic ability. J Biol Chem 2000, **275**:17869–17877.

76. Nakata D, Close BE, Colley KJ, Matsuda T, Kitajima K: Molecular cloning and expression of the mouse *N*-acetylneuraminic acid 9-phosphate synthase which does not have deaminoneuraminic acid (KDN) 9-phosphate synthase activity. *Biochem Biophys Res Commun* 2000, **273**:642–648.

77. Vann WF, Daines DA, Murkin AS, Tanner ME, Chaffin DO, Rubens CE, Vionnet J, Silver RP: The NeuC protein of *Escherichia coli* K1 is a UDP *N*-acetylglucosamine 2-epimerase. *J Bacteriol* 2004, **186**:706–712.

78. Murkin AS, Chou WK, Wakarchuk WW, Tanner ME: Identification and mechanism of a bacterial hydrolyzing UDP-*N*-acetylglucosamine 2-epimerase. *Biochemistry* 2004, **43**:14290–14298.

79. Petersen M, Fessner W, Frosch M, Luneberg E: The siaA gene involved in capsule polysaccharide biosynthesis of *Neisseria meningitidis* B codes for *N*-acylglucosamine–6-phosphate 2-epimerase activity. FEMS Microbiol Lett 2000, **184**:161–164.

80. Vann WF, Tavarez JJ, Crowley J, Vimr E, Silver RP: Purification and characterization of the *Escherichia coli* K1 neuB gene product *N*-acetylneuraminic acid synthetase. *Glycobiology* 1997, **7**:697–701.

81. Ringenberg MA, Steenbergen SM, Vimr ER: The first committed step in the biosynthesis of sialic acid by *Escherichia coli* K1 does not involve a phosphorylated *N*-acetylmannosamine intermediate. *Mol Microbiol* 2003, **50**:961–975.

82. Rodríguez-Aparicio LB, Luengo JM, González-Clemente C, Reglero A: Purification and characterization of the nuclear cytidine 5′-monophosphate *N*-acetylneuraminic acid synthetase from rat liver. *J Biol Chem* 1992, **267**:9257–9263.

83. Jarvis GA, Vedros NA: Sialic acid of group B *Neisseria meningitidis* regulates alternative complement pathway activation. *Infect Immun* 1987, **55**:174–180.

84. Marques MB, Kasper DL, Pangburn MK, Wessels MR: Prevention of C3 deposition by capsular polysaccharide is a virulence mechanism of type III group B streptococci. *Infect Immun* 1992, **60**:3986–3993.

85. Meri S, Pangburn MK: Discrimination between activators and nonactivators of the alternative pathway of complement: regulation via a sialic acid /polyanion binding site on factor H. *Proc Natl Acad Sci USA* 1990, **87**:3982–3986.

86. Crocker PR, Varki A: Siglecs, sialic acids and innate immunity. *Trends Immunol* 2001, **22**:337–342.

87. Crocker PR: Siglecs in innate immunity. *Curr Opin Pharmacol* 2005, **5**:431–437.

88. Cornall RJ, Goodnow CC, Cyster JG: Regulation of B cell antigen receptor signaling by the Lyn/CD22/SHP1 pathway. *Curr Top Microbiol Immunol* 1999, **244**:57–68.

89. Pan B, Fromholt SE, Hess EJ, Crawford TO, Griffin JW, Sheikh KA, Schnaar RL: Myelin-associated glycoprotein and complementary axonal ligands, gangliosides, mediate axon stability in the CNS and PNS: neuropathology and behavioral deficits in single- and double-null mice. *Exp Neurol* 2005, **195**:208–217.

90. Avril T, Floyd H, Lopez F, Vivier E, Crocker PR: The membrane-proximal immunoreceptor tyrosine-based inhibitory motif is critical for the inhibitory signaling mediated by Siglecs-7 and -9, CD33-related Siglecs expressed on human monocytes and NK cells. *J Immunol* 2004, **173**:6841–6849.

91. Ikehara Y, Ikehara SK, Paulson JC: Negative regulation of T cell receptor signaling by Siglec-7 (p70/AIRM) and Siglec-9. *J Biol Chem* 2004, **279**:43117–43125.

92. Lajaunias F, Dayer JM, Chizzolini C: Constitutive repressor activity of CD33 on human monocytes requires sialic acid recognition and phosphoinositide 3-kinase-mediated intracellular signaling. *Eur J Immunol* 2004, **35**:243–251.

93. Zhang M, Angata T, Cho JY, Miller M, Broide DH, Varki A: Defining the in vivo function of Siglec-F, a CD33-related Siglec expressed on mouse eosinophils. *Blood* 2007, **109**:4280–4287.

94. Angata T, Margulies EH, Green ED, Varki A: Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms. *Proc Natl Acad Sci USA* 2004, **101**:13251–13256.

95. Sonnenburg JL, Altheide TK, Varki A: A uniquely human consequence of domain-specific functional adaptation in a sialic acid-binding receptor. *Glycobiology* 2004, **14**:339–346.

96. Delputte PL, Nauwynck HJ: Porcine arterivirus infection of alveolar macrophages is mediated by sialic acid on the virus. *J Virol* 2004, **78**:8094–8101.

97. Vanderheijden N, Delputte PL, Favoreel HW, Vandekerckhove J, Van Damme J, Van Woensel PA, Nauwynck HJ: Involvement of sialoadhesin in entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages. *J Virol* 2003, **77**:8207–8215.

98. Barthelson R, Mobasseri A, Zopf D, Simon P: Adherence of *Streptococcus pneumoniae* to respiratory epithelial cells is inhibited by sialylated oligosaccharides. *Infect Immun* 1998, **66**:1439–1444.

99. Tong HH, Liu X, Chen Y, James M, Demaria T: Effect of neuraminidase on receptor-mediated adherence of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Acta Otolaryngol* 2002, **122**:413–419.

100. Corfield AP, Wagner SA, O'Donnell LJD, Durdey P, Mountford RA, Clamp JR: The roles of enteric bacterial sialidase, sialate *O*-acetyl esterase and glycosulfatase in the degradation of human colonic mucin. *Glycoconj J* 1993, **10**:72–81.

101. Vimr ER, Kalivoda KA, Deszo EL, Steenbergen SM: Diversity of microbial sialic acid metabolism. *Microbiol Mol Biol Rev* 2004, **68**:132–153.

102. Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J, Buhler JD, Gordon JI: Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science* 2005, **307**:1955–1959.

103. Roggentin P, Schauer R, Hoyer LL, Vimr ER: The sialidase superfamily and its spread by horizontal gene transfer. *Mol Microbiol* 1993, **9**:915–921.

104. Taylor G: Sialidases: structures, biological significance and therapeutic potential. *Curr Opin Struct Biol* 1996, **6**:830–837.

105. Varki A, Diaz S: A neuraminidase from *Streptococcus sanguis* that can release *O*-acetylated sialic acids. *J Biol Chem* 1983, **258**:12465–12471.

106. Zenz KI, Roggentin P, Schauer R: Isolation and properties of the natural and the recombinant sialidase from *Clostridium septicum* NC 0054714. *Glycoconj J* 1993, **10**:50–56.

107. Brabetz W, Wolter FP, Brade H: A cDNA encoding 3-deoxy-D-manno-oct-2-ulosonate-8-phosphate synthase of *Pisum sativum* L. (pea) functionally complements a kdsA mutant of the Gram-negative bacterium *Salmonella enterica*. *Planta* 2000, **212**:136–143.

108. Royo J, Gimez E, Hueros G: CMP-KDO synthetase: a plant gene borrowed from Gram-negative eubacteria. *Trends Genet* 2000, **16**:432–433.

109. Tanner ME: The enzymes of sialic acid biosynthesis. *Bioorg Chem* 2005, **33**:216–228.

110. Gunawan J, Simard D, Gilbert M, Lovering AL, Wakarchuk WW, Tanner ME, Strynadka NC: Structural and mechanistic analysis of sialic acid synthase NeuB from *Neisseria meningitidis* in complex with Mn^{2+}, phosphoenolpyruvate, and *N*-acetylmannosaminitol. *J Biol Chem* 2004, **280**:3555–3563.

111. Wise EL, Rayment I: Understanding the importance of protein structure to Nature's routes for divergent evolution in TIM barrel enzymes. *Acc Chem Res* 2004, **37**:149–158.

112. Angata T, Nakata D, Matsuda T, Kitajima K, Troy FAII: Biosynthesis of KDN (2-keto-3-deoxy-D-glycero-D-galacto-nononic acid) – identification and characterization of a KDN-9-phosphate synthetase activity from trout testis. *J Biol Chem* 1999, **274**:22949–22956.

113. Lawrence SM, Huddleston KA, Tomiya N, Nguyen N, Lee YC, Vann WF, Coleman TA, Betenbaugh MJ: Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells. *Glycoconj J* 2001, **18**:205–213.

114. Kim K, Lawrence SM, Park J, Pitts L, Vann WF, Betenbaugh MJ, Palter KB: Expression of a functional *Drosophila melanogaster* N-acetylneuraminic acid (Neu5Ac) phosphate synthase gene: evidence for endogenous sialic acid biosynthetic ability in insects. *Glycobiology* 2002, **12**:73–83.

115. Chen H, Blume A, Zimmermann-Kordmann M, Reutter W, Hinderlich S: Purification and characterization of N-acetylneuraminic acid-9-phosphate synthase from rat liver. *Glycobiology* 2002, **12**:65–71.

116. Pirun M, Babnigg G, Stevens FJ: Template-based recognition of protein fold within the midnight and twilight zones of protein sequence similarity. *J Mol Recognit* 2005, **18**:203–212.

117. Schoenhofen IC, McNally DJ, Brisson JR, Logan SM: Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* 2006, **16**:8C–14C.

118. Varki A, Diaz S: The release and purification of sialic acids from glycoconjugates: methods to minimize the loss and migration of O-acetyl groups. *Anal Biochem* 1984, **137**:236–247.

119. Kamerling JP, Schauer R, Shukla AK, Stoll S, van HH, Vliegenthart JFG: Migration of O-acetyl groups in N,O-acetylneuraminic acids. *Eur J Biochem* 1987, **162**:601–607.