Review

Effects of Sialic Acid Modifications on Virus Binding and Infection

Brian R. Wasik,¹,* Karen N. Barnard,¹ and Colin R. Parrish¹

Sialic acids (Sias) are abundantly displayed on the surfaces of vertebrate cells, and particularly on all mucosal surfaces. Sias interact with microbes of many types, and are the targets of specific recognition by many different viruses. They may mediate virus binding and infection of cells, or alternatively can act as decoy receptors that bind virions and block virus infection. These nine-carbon backbone monosaccharides naturally occur in many different modified forms, and are attached to underlying glycans through varied linkages, creating significant diversity in the pathogen receptor forms. Here we review the current knowledge regarding the distribution of modified Sias in different vertebrate hosts, tissues, and cells, their effects on viral pathogens where those have been examined, and outline unresolved questions.

Interaction between Viruses and Sias

Sialic acids (see Glossary) generally occur as terminal monosaccharides on many different cell surfaces and secreted glycoconjugates and are therefore involved in key interactions between cells and many different viruses (as well as other pathogens) at various points in their infection and transmission cycles. Many viruses specifically bind to host Sias and use them as primary receptors for cell infection, or as components in a series of interactions that lead to infection. In vertebrates, Sias may be present as constituents of a variety of different complexes on the cell surface, including dense layers of glycoconjugates referred to as the glyccalyx. Mucosal surfaces are further protected by a viscous secretion, termed mucus, which varies in structure and composition depending on the location in the body [1]. However, mucus is formed in complex layers of cell-associated and secretory forms [2,3]. Sias in mucus and at cell surfaces can bind and trap viruses and prevent them from accessing their target tissues, and can also remove the bound virions in an active process mediated by mucociliary transport [4]. Our understanding of virus–Sia interactions continues to be illuminated by new structural and biochemical analysis of viral proteins and bound glycans [5,6]. Currently we still do not clearly know which specific Sia structures are present on different cells or tissues of many viral hosts, including humans. It is clear that Sias exist in many diverse forms, and that their synthesis is tightly regulated yet highly variable on the cells and tissues of individual animals. The displayed glycan combinations vary under differing physiological conditions, stages of development, and also between different animal species. Many viruses also encode Sia-specific enzymes that alter the host Sia and therefore their specific interactions, including sialidases that remove the terminal Sia from glycans or esterases that remove ester-bonded acetyl modifications from the Sias where those are present.

Here we review examples of the various interactions between Sias and viruses in vertebrate animals, and examine how those can be modulated by variations in Sia structures, leading to
differences in recognition and/or viral enzyme activity. We further summarize what is known about the diversity and distribution of modified Sias, the known and potential impacts of these specific glycan variants on different viruses, and define some questions that remain to be resolved.

Sias: Modifications and Variation

Sias are a highly diverse family of monosaccharides that serve as terminal residues of N- and O-linked glycoproteins and glycolipids, and they are also components of various polysaccharides. The core Sia structure contains nine carbons, and modifications are most commonly found at the 4, 5, 7, 8, and 9 positions. Modifications at the 5-carbon position determine the primary Sia forms: N-acetyllneuraminic acid (Neu5Ac), N-glycoly neuraminic acid (Neu5Gc), and 2-keto-3-deoxyoctononic acid (KDN) [7]. These primary Sia forms serve as core structures, and additional modifications can be generated enzymatically to form more than 30 variant types of Sia [8]. Those include acetyl, sulfate, lactoyl, and methyl group additions to various positions. Display of the different modified forms clearly varies between different organisms, tissues, and cell types, and as the modifications may be present in myriad combinations, there may be a wide diversity of glycan forms (Figure 1). In addition, the linkage of the Sia to the underlying glycan structure may also vary, with α2,3 or α2,6 linkages being most common, as well as α2,8 linkages in the formation of poly-Sia structures which are primarily found in the brain and on certain cells of the immune system [9].

Our understanding of the expression and biological roles of this wide diversity of variant Sias is rather limited for various practical and technical reasons. However, the modifications are widespread and clearly allow the Sia to serve a variety of physiologically important roles in normal cell and tissue development and in controlling cell-cell interactions. For example, the constitutive expression of influenza C virus 9-O-acetylatedester in transgenic mice resulted in complete developmental arrest of embryos as early as the two-cell stage [10]. Polysialic acids play an important role in neural development and are involved in neuron growth, axon guidance, and synapse formation, and their synthesis is tightly regulated [11]. Transgenic mice lacking expression of synthases and transferases necessary for oligosaccharic acid expression in the central nervous system showed atypical neuronal development and a ‘sudden death’ phenotype [12]. In addition, Sias are involved in immune cell maturation and activation. Regulated 9-O-acetylation of ganglioside GD3 (also called CD60) on T cells correlates with cell differentiation and blocking of pro-apoptotic pathways in proliferating T cells [13]. Similarly, regulated O-acetylation at the C9-position of Sia on B cells is required for the correct development and activation of those cells, as 9-O-acetylation blocks recognition by the Sia-binding immunoglobulin-type lectin CD22 (SIGLEC-2), which acts as an inhibitor of B cell receptor signaling and activation [14]. Given the regulatory role of Sias in so many different cell processes and pathways, it is not surprising that abnormal expression of modified Sias has been identified as a hallmark of many cancers, including some types of leukemia [15,16] and colorectal cancers [17].

Formation of specific glycans is regulated by the expression of a variety of enzymes, many of which are localized to the endoplasmic reticulum and Golgi apparatus, where they direct co- and post-translational modification of the glycans on glycoproteins or glycolipids. Enzymes specific for Sia addition and modification are expressed in the Golgi apparatus while others may be expressed in the cytoplasm and modify the Sia precursors. The varying expression of sialyltransferases, sialidases, esterases, and other Sia-modifying enzymes makes Sia synthesis and modification highly variable and dynamic, and as a result their display is influenced by both external and internal cellular stimuli. Several of the genes encoding Sia-modifying enzymes have been identified and analyzed in detail (Figure 1). Those include cytidine monophosphate- N-acetyllneuraminic acid hydroxylase (CMAH), the enzyme responsible for formation of Neu5Gc from the CMP-Neu5Ac precursor [18]. The CASD1 gene (capsule structure 1 domain

---

**Glossary**

**Esterase (acyl esterase):** In this context, an enzyme that removes O-acetyl groups from Sia sugars by cleaving the ester bond. Esterases have specificity for the position of the O-acetyl group in the Sia, typically 4-O-Ac or 9-O-Ac. Esterases are encoded by influenza C and nidovirus family viruses in the same glycoprotein (related-origin) encoding a hemagglutinin domain that has specificity for the same modified Sia.

**Hemagglutinin (HA):** A viral lectin that binds to Sia (sialolectin), historically characterized by the ability of a virus to agglutinate erythrocytes in vitro. Orthomyxov-, paramyxov-, and nidoviruses encode hemagglutinating glycoproteins either alone or fused to sialidase or sialidase domains. Many surface capsid proteins of nonenveloped viruses also have hemagglutination sialolectin function.

**Sialic acid (Sia):** Nine-carbon α-keto acidic sugar whose name is derived from its presence in saliva. It was discovered from human samples as primarily N-acetyllneuraminic acid (Neu5Ac). It can be enzymatically modified at multiple positions. Sia is typically found at the termini of N- and O-glycans on glycoproteins, as well as components of glycolipids, where they perform important recognition reactions for other cellular factors and/or environmental agents such as pathogens.

**Sialidase:** An enzyme that removes the Sia sugar from glycosaccharide structures. Endo-sialidases cleave poly-Sia linkages. Sialidases of viruses have been traditionally defined as neuraminidases (NA) for their cleavage of Neu5Ac. Historically also identified as a virus-encoded receptor-destructing enzyme (RDE).

**Tropism:** The preferential viral infection of target hosts, tissues, or cells. Tropism is often related to permissiveness resulting from the presence of a specific receptor. Sia-binding virus tropisms may be greatly influenced by the relative distribution of Sia species on a given tissue. The degree of Sia diversity (including modified inhibitors or decoys) in alternative hosts or tissues may influence potential viral emergence.
containing 1) encodes a 9-O-acetyltransferase enzyme that mediates the formation of 9-O-acetyl Sia (9-O-Ac) from a CMP-Neu5Ac precursor [13,19,20]. However, while enzymatic activity for a 4-O-acetyltransferase has been demonstrated, a candidate gene has not yet been identified [21]. It is also still not clear what activities control 7-O-acetyl Sia (7-O-Ac), as in vitro studies have shown acetyl groups added at the C7-position migrate to the C9-position by way of C8-position [22,23]. Whether CASD1 adds acetyl groups at the C7-position or directly to the C9-position is not clear, and it is possible that a migrase enzyme is involved in that modification. A variety of sialyltransferases (likely 20 in humans) regulate the specific addition and linkage of the Sia to the underlying N- or O-linked glycan on glycoproteins or glycolipids [24,25].

Many sialylated glycans are synthesized in and presented on vertebrate cells, but detailed information about how those are produced and displayed is often lacking; however, it is clear that the relative display patterns of each modified Sia form can vary widely between even closely related animal species, as well as between and within different tissue types [26,27]. For example, 4-O-acetyl Sia (4-O-Ac) appears to be restricted to only a few tissue types in mice, including the colon, spleen, stomach, thymus, and uterus [27,28]. One factor determining the species-specific display of modified Sia is the presence and expression of the necessary enzymes – as some vertebrate lineages may lack the intact genes encoding certain Sia-modifying enzymes. The best documented case is a lack of Neu5Gc due to independent loss of the functional CMAH gene in humans and in some new world primates, as well as a lack of CMAH expression or activity in birds and reptiles (sauropterygia), seals (pinnipedia), ferrets (mustelids), and possibly in western breeds of dogs [18,29–32]. In many other mammalian species CMAH is actively expressed and Neu5Gc is generally abundant, but with observed variation between tissues and cell types. Further, there appears to be no enzyme that causes reversion of the glycolyl modification [33]. It is still unclear why such wide differences in Neu5Gc expression exist between different animals. It has been suggested that the loss of Neu5Gc in humans following divergence from our common ancestor with chimpanzees may have been beneficial by allowing resistance to Neu5Gc-specific pathogens, such as ancestral strains of Plasmodium malaria [34]. Another interesting case is the high variability of 4-O-Ac in many mammals, including apparent low levels or loss of that Sia form in cows, pigs, sheep, goats, and humans [26,27]. Again, the biological reasons for the highly variable synthesis and distribution of modified Sias are not known.
**Viral Interactions: An Overview**

Many different viruses bind to Sia-terminated glycans either as primary receptors for infection or by using them as accessory receptors in host cell binding and uptake (Table 1), and the evolution of the viruses has selected for specific interactions with particular Sia forms and linkages on different hosts and tissues. Where these have been defined, the evolved specificities for the host receptor variants and alternative linkages often play important roles in the tropism of the virus at the level of host, tissue, and specific differentiated cell types. In addition, the relative distribution of modified forms of Sia in given hosts and tissues may shape the evolution of Sia-interacting viral proteins, including hemagglutinins. Determining the exact nature of virus–Sia interactions, as well as the impact of varying Sia chemistry on those interactions, is therefore critical to

<table>
<thead>
<tr>
<th>Family</th>
<th>Virus Genus</th>
<th>Viral Sialolectin</th>
<th>Sia Enzyme</th>
<th>Sia Specificity</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reoviridae dsRNA</td>
<td>Reovirus</td>
<td>Sigma-1</td>
<td></td>
<td>Neu5Acx2-3 Gal (Type 1)</td>
<td>[72]</td>
</tr>
<tr>
<td></td>
<td>Rotavirus</td>
<td>VP4</td>
<td></td>
<td>Neu5Ac (RRV, UK) Neu5Gc (NCDV, OSU, SA11)</td>
<td>[49-51,73]</td>
</tr>
<tr>
<td></td>
<td>Orbivirus</td>
<td>VP2</td>
<td></td>
<td>Not known</td>
<td>[74]</td>
</tr>
<tr>
<td>Coronavirus +ssRNA</td>
<td>Betacoronavirus</td>
<td>Spike, HE</td>
<td>Esterase (HE)</td>
<td>Neu5,7,9Ac₃ (BCoV-Mebus) Neu4,5Ac₂ (MHV-S)</td>
<td>[27,38,41]</td>
</tr>
<tr>
<td></td>
<td>Toroviruses</td>
<td>Spike, HE</td>
<td>Esterase (HE)</td>
<td>Neu5,9Ac₂ (PToV-P4)</td>
<td>[27]</td>
</tr>
<tr>
<td>Picornaviridae +ssRNA</td>
<td>Enterovirus</td>
<td>VP1, VP3; VP1 only (CVA24v)</td>
<td>Neu5Acx2-3Gal Neu5Acx2-6 Gal (EV-D88); Neu5Acx2-3 Gal (EV-70, CVA24v)</td>
<td>[75-77]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carbovirus</td>
<td>VP1, VP3</td>
<td></td>
<td>Neu5Acx2-3Gal</td>
<td>[78]</td>
</tr>
<tr>
<td>Orthomyxoviridae -ssRNA</td>
<td>Influenza A</td>
<td>HA, NA</td>
<td>Sialidase (NA)</td>
<td>Neu5Acx2-3 Gal (Avian, Porcine) Neu5Acx2-6 Gal (Human, Porcine)</td>
<td>[79]</td>
</tr>
<tr>
<td></td>
<td>Influenza B</td>
<td>HA, NA</td>
<td>Sialidase (NA)</td>
<td>Neu5Acx2-3Gal Neu5Acx2-6 Gal</td>
<td>[80]</td>
</tr>
<tr>
<td></td>
<td>Influenza C</td>
<td>HEF</td>
<td>Esterase (HE)</td>
<td>Neu5,9Ac₂</td>
<td>[81]</td>
</tr>
<tr>
<td>Paramyxoviridae -ssRNA</td>
<td>Rubulavirus</td>
<td>HN</td>
<td>Sialidase (HN)</td>
<td>Neu5Acx2-3 Gal Neu5Acx2-6Gal Neu5Acx2-8polySia (Mumps)</td>
<td>[82]</td>
</tr>
<tr>
<td></td>
<td>Respirovirus</td>
<td>HN</td>
<td>Sialidase (HN)</td>
<td>Neu5Acx2-3 Gal (hPIV-1); Neu5Acx2-3 Gal (hPIV-3)</td>
<td>[83]</td>
</tr>
<tr>
<td></td>
<td>Avulavirus</td>
<td>HN</td>
<td>Sialidase (HN)</td>
<td>Neu5Ac (NDV)</td>
<td>[84]</td>
</tr>
<tr>
<td>Paroviridae ssDNA</td>
<td>Bocavirus</td>
<td>VP₁</td>
<td>2,3,6- and 2,3,8-Linked Sia (BPV)</td>
<td>[85]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protoparovirus</td>
<td>VP₁</td>
<td></td>
<td>Neu5Acx2-3 Gal (MVM)</td>
<td>[86]</td>
</tr>
<tr>
<td>Polyomaviridae dsDNA</td>
<td>Polyomavirus</td>
<td>VP₁</td>
<td></td>
<td>Neu5Acx2-3 Gal (BKV, SV-40), Neu5Acx2-6 Gal (JCV), Neu5Gcx2-3 Gal (SV-40)</td>
<td>[46-48]</td>
</tr>
</tbody>
</table>

*Abbreviations: dsRNA, double-stranded RNA; +ssRNA, positive-sense single-stranded RNA; −ssRNA, negative-sense single-stranded RNA; ssDNA, single-stranded DNA; dsDNA, double-stranded DNA.*
understanding the complex impacts of host glycan variation on virus systems. Here we review some of the Sia modifications and their known interactions with viruses, while also identifying areas where more information is required. In addition to the Sia modifications discussed, other modifications may also be present on Sias that potentially interact with viruses (lactoyl, phosphate, sulfate, or methyl additions), yet little is known about their relative synthesis and distribution [7].

4-O-Acetylated Sia
The distribution and roles of 4-O-Ac Sias are not well understood as, until recently, there were few means to detect their presence in cells and tissues. Mass spectrometry to identify glycan structures may use methods (permethylation) that remove the O-acetyl modifications, and many approaches that can detect the 4-O-Ac modification do not characterize its distribution at the cellular or subcellular level. Recently the hemagglutinin-esterases (HEs) of the mouse hepatitis virus (MHV) S strain and of infectious salmon anemia virus (ISAV) were found to specifically recognize 4-O-Ac Sias, and esterase-inactive forms of the MHV-S protein can be used to detect the presence of that modified Sia in cells or tissue sections (Box 1) [26,27]. These probes showed that 4-O-Ac-modified Sias are displayed at high levels on the erythrocytes and other tissues of some mammals, including guinea pigs and horses, and in many ‘lower’ vertebrates, particularly fish.

The physiological role of 4-O-Ac Sias and the reasons for the apparent loss or altered modifying enzyme expression in so many animals is unclear, but their ability to act as an inhibitor for many sialidases may be involved. 4-O-Ac Sias have been known since the 1950s to influence influenza tropisms due to their role as gamma inhibitors in horse and guinea pig sera, which interfere with infection and replication of some human influenza A H2 and H3 strains by inhibiting the activity of the viral neuraminidase (NA). Those inhibited viruses can escape by selecting variants in the viral hemagglutinin (HA) that do not bind the modified Sias [35–37]. In contrast to their inhibitory effects on influenza viruses, some other viruses use 4-O-Ac Sias as specific receptors. The presence of 4-O-Ac Sias in the gastrointestinal tract,

**Box 1. Detection of Modified Sialic Acid (Sia)**

Traditional characterization of Sia composition of biological samples has involved chemical release, purification, and chemical derivatization (DMB labeling) for analysis by high-performance liquid chromatography (HPLC), mass spectrometry, thin-layer chromatography, or nuclear magnetic resonance. However, many of the methods for sample preparation result in the loss of labile modifications of Sia. Further, these methods do not show the relative distribution of the modified Sia forms in the context of an intact biological sample, such as tissue, its various cell populations, or mucosal surfaces. Histochemical analyses have been previously undertaken with mild periodate oxidation and saponification to generally visualize Sia and 9-O-acetylated forms [7].

The accurate identification of modified Sia in biological samples requires new and robust molecular methodologies. The last two decades have seen steady progress in the field of glyobiology, including the use of polyclonal and monoclonal antibodies, naturally occurring lectins, sialidases, esterases, as well as Sia lysases to interrogate and manipulate Sia in their native state on cells and tissues. Monoclonal antibody generation against specific carbohydrate forms is difficult and often results in low affinity antibody forms with specificities for particular linkages or other specific variants, and is hampered by the need for a naive animal for immunization that does not present that particular Sia form (so that it is not a self-antigen). The use of plant-derived lectins (SNA, MAL, and MAH), as well as others from various sources, has greatly contributed to the analysis of linkage-variant Sia species, and has revealed the variable Sia linkages that control the specificities and tropisms of many viruses. More recently, the development of recombinant expressed sialolectins with various fusion and post-translational labels have expanded the molecular toolkit for probing Sia diversity. These expressed sialolectins include those of vertebrate (siglecs) and microbial origin, as well as a number of virus-derived virolectins. This approach allows for the harnessing of naturally coevolved specificities, under a recombinant molecular platform that would allow for rational or artificial-selection engineering. Recent publications have verified the use of nidovirus hemagglutinin-esterase (HE) probes for analysis of O-acetyl Sia forms in the tissues of humans and mice [27]. A robust survey and cataloging of virus-encoded sialolectins would contribute to our understanding of virus–Sia interactions while aiding in the development of key tools for interrogating the ‘sialome’ of cellular systems.
particularly the colon, and brain of mice likely contributes to the tropism seen in some MHV-S infections which target those organs [38]. As 4-O-Ac Sia is highly enriched in the mucus present on fish skin, some piscine pathogens (including paramyxoviruses and ISAV) may bind to or interact with this modified form [26,39]. Apart from ISAV, it is unclear how 4-O-Ac interacts with fish viruses or other pathogens, in part due to the difficulty of studying this Sia variant. Overall, very little is currently known about the synthesis and distribution of this modified Sia in most vertebrates, or how its presence shapes tropism of other Sia-binding viruses (or of other pathogens which bind or cleave Sia).

9-O-Acetylated (and 7-O-Acetylated) Sia

The addition of acetyl or other groups to the C9 position of Sia is widely seen in cells and tissues of a variety of animals, including humans, and 9-O-acetylated (9-O-Ac) Sias have previously been detected using the HE protein of human influenza C virus as a probe [40]. As discussed previously, acetyl modifications of the 7-position also occur, creating 7,9-O-Ac2 modified Sias in addition to 9-O-Ac modifications. 7-O-Ac modifications alone are rarely observed, likely due to the natural tendency for migration of the acetyl group to the C9-position. Besides influenza C virus, 9-O-Ac Sia can be used as a specific receptor for many coronaviruses (including in the 7,9-O-Ac2 form) and toroviruses [27,41]. It appears to be the receptor for the recently discovered influenza D virus (sometimes referred to as bovine influenza virus) which is 50% homologous with human influenza C virus and which also has a HE receptor-binding and esterase protein [42,43]. The presence of the 9-O-Ac Sia modification has been shown to affect the activities of the NA of influenza viruses, reducing its activity by 3- to 4-fold compared to the nonmodified forms [44], although the significance of that activity for natural viral fitness is not yet known.

Neu5Gc

Variations in modified Sia between related species may underlie major evolutionary host variability and in pathogen specificity. As mentioned above, humans differ from most nonhuman primates by the loss of the enzyme CMAH, which produces Neu5Gc from CMP-Neu5Ac precursors. Among viruses that infect both human and nonhuman primate species (as well as bacteria and macroparasites), divergence of strains and differential tropisms for Neu5Ac versus Neu5Gc have been observed [34,45]. The polyomaviruses are small double-stranded DNA (dsDNA) viruses that infect a wide range of hosts, and they bind Sia via the major capsid protein, VP1, which is arranged within the particle as 72 pentamers. The polyomaviruses from humans (JC virus, BK virus, and Merkel Cell polyomavirus) all have a strong preference for Neu5Ac Sia in cell attachment, while viruses from nonhuman primates and other mammalian hosts (simian virus 40) show increased flexibility in Sia recognition as they have a larger, more polar pocket allowing binding to Neu5Gc Sia [46-48]. Similarly, rotaviruses, double-stranded RNA (dsRNA) viruses in the Reovirus family, also show a strain-dependent specificity for Neu5Ac or Neu5Gc binding – bovine strain NCDV and porcine strains OSU and CRW-8 have specificity for Neu5Gc while rhesus macaque strain RRV prefers Neu5Ac [49]. Whether human rotaviruses interact with Sias has been an area of debate as they are classically termed sialidase-insensitive, meaning removal of terminal Sia via sialidases does not inhibit infection. However, some Sia forms are resistant to sialidase (including complex branching, internally linked Sia, and further modifications) and it appears likely that at least some human rotaviruses interact with Sias during cell entry [50,51]. It is also possible that the sialidase insensitivity that distinguishes human rotaviruses from other animal strains is due to the loss of Neu5Gc in humans and subsequent adaptation of the rotavirus to a different modified Sia form [50]. However, even within a single host there may be different specificities, such as bovine strains that include the Neu5Ac-binding UK or Neu5Gc-binding NCDV [49]. It is unknown whether the loss of Neu5Gc in humans has impacted the relative emergence potential of Sia-binding viruses from Neu5Gc-positive zoonotic reservoirs.
Viral Proteins Altering Sia: Sialidases and Esterases

Viruses can encode one of two varieties of Sia-modifying enzymes: sialidases or esterases. Viral sialidases (often referred to as neuraminidases) complement the Sia-binding glycoproteins to coordinate receptor interaction activity by removing the Sia from glycans. Viral esterases remove O-acetyl groups from modified Sia and are often present as HEs where both activities are expressed as separate functions of the same protein. Both sialidase and esterase activities appear to provide similar benefits to the viruses, as they alter the host environment by removing receptors and improving viral penetration of mucus, favoring cell infection, and enhancing virion release during the final stages of the virus lifecycle [52,53]. While some viral sialidases may be blocked by modified Sia (as discussed previously), viral esterases have evolved to remove specific modifications. We discuss two examples of how modified Sias influence viral infection.

Sialidases

Influenza A Viruses. The HA and NA are encoded in distinct genomic segments, so that reassortment of HA and NA variants can occur readily. However, the resulting viruses appear to need a balance between the HA Sia-binding and NA Sia-cleaving activities, which may be accomplished by coevolution of either or both genes [54]. It has been shown that sialidases from many pathogens (bacterial and influenza) have optimal enzymatic activities for α2,3-linked Neu5Ac Sia, and the shift to the α2,6 linkage, hydroxylation of Neu5Ac to Neu5Gc, or addition of O-acetyl groups to C4, 7, or 9 reduce the capacity of sialidases to cleave in vitro [44,55,56]. The natural in vivo activity of human or avian influenza NA towards modified Sia appears to be consistently reduced, while ‘drift’ in linkage specificity has been observed over large influenza evolutionary timescales [57]. It appears that, for influenza viruses, HA compensatory mutations are commonly used to avoid inhibitory or ‘decoy’ Sias, rather than changes in the sialidase specificity of NA, and also that sialidases cannot readily gain the ability to avoid inhibition by at least the 4-O-Ac Sia [58]. Similarly, NA resistance alleles to sialidase inhibitors (e.g., Oseltamivir-resistant influenza) can arise, but may require favorable compensatory epistatic alleles in HA to allow fixation [59,60]. For example, mutations that allow human influenza viruses to replicate in the presence of high levels of 4-O-Ac Sia occurred in the Sia binding site of the HA and prevented HA binding to the 4-O-Ac Sia. Those mutations were seen in the H3N8 and H7N7 equine influenza viruses, as well as in human influenza viruses experimentally adapted to grow in the presence of horse serum proteins [58]. It is not known whether selection of mutations in HA to avoid these inhibitory modified Sias also alters HA receptor binding specificities for other Sias, or alters antigenic epitopes, and thereby shapes the adaptive potential of the viruses in different hosts. However, it is likely that the relative abundance of modified Sia forms in different hosts contributes to host-specific HA and NA selection, and may also constrain the relative transmission potential in novel emergent viruses.

Paramyxoviruses. Many paramyxoviruses encode a hemagglutinin-neuraminidase (HN) glycoprotein, which carries both receptor-binding and sialidase functions [61]. The HN of paramyxoviruses is therefore analogous to both the HA and NA in influenza A viruses. The NA domain likely plays a part in the release of budding virions and in the clearing of Sia decoys, similar to influenza NA. However, unlike influenza HA, HN does not directly cause fusion of the virus with its target cell; instead, the HN binding of Sia activates fusion through an additional protein, the fusion (F) protein [62]. Structural models show that paramyxovirus HN sialidase domains share similarities with influenza NA despite having divergent polypeptide sequences; whether or not they share a common origin is unknown [63]. Recent studies have shown that the NA-active site also has functional receptor-binding activity. These secondary Sia binding sites aid in balancing the receptor-binding–receptor-destroying–F activation activities of HN [64,65]. Whether these secondary binding sites also avoid binding to modified forms requires further research, particularly as the presence and location of secondary sites varies between
paramyxovirus species [65]. Overall, little is known about how modified Sias interact with paramyxoviruses or how their presence shapes viral evolution in this complex and highly diverse family [61].

**Esterases**  
*Nidovirales.* This Order contains the evolutionarily related families of coronaviruses and toroviruses, which are both enveloped positive-strand RNA viruses that commonly infect the respiratory and/or gastrointestinal tracts of many vertebrates. Both toroviruses and group II coronaviruses often express an HE glycoprotein on the surface of their envelope in addition to the large binding and fusion active spike (S) protein [66]. Coronavirus and torovirus HEs appear to have originated through multiple horizontal gene transfer and recombination events between strains [67,68]. Interestingly, these HEs also appear homologous to the Orthomyxovirus family member influenza C HE glycoprotein, suggesting a common evolutionary ancestor for the protein resulting from gene exchange between these different viral taxa [69].

The modified Sia specificities of S and HE are determinants of tissue and cell tropism for both type II coronaviruses and for toroviruses. For example, the coronavirus MHV strains S and JHM preferentially bind to, as well as de-acetylate, 4-O-Ac modified Sia, while ancestral strains of MHV prefer 9-O-Ac Sia similar to other nidovirus members [23,38]. Both 9-O-Ac and 4-O-Ac are displayed in the colon of mice [27], and 4-O-Ac-specific MHV strains also exhibit a novel neurotropic phenotype likely due to the presence of this modified Sia in the brain [38]. The matching of S and HE binding specificities is further seen in bovine coronaviruses and toroviruses, porcine toroviruses, and many others, with Sia specificity matching that of their target tissue [27,68]. The HE is not an essential protein for all coronaviruses [70], but appears to increase the efficiency of cell infection, either by improving viral binding initially or by enhancing virion release [71]. However, the specific roles of HE in penetrating or altering host mucus, facilitating membrane fusion, or aiding in release of budding virions from infected cells are currently poorly understood. It also remains to be seen what role changes in HE specificity may play in viral adaptation during infection of new host species; this is particularly relevant considering the zoonotic potential of coronaviruses.

**Concluding Remarks**  
Sias are one of the first points of contact between many different pathogens and the host organism due to their presence on the outer surfaces of cells and mucosal tissues. The diversity of Sia forms results from modifications that can be added singly or in combinations (Figure 1), leading to complex barriers that pathogens must navigate to reach their target cells or tissues. The influence of Sia variation on viral adaptation is seen in the diverse viral families that use Sia as a receptor or coreceptor, often showing specificity for some modified forms over others. This is also shown in the number of viruses that carry Sia-modifying enzymes such as sialidases or esterases, which allow them to control their binding and release with high specificity (Table 1). Sia interactions and modifications also extend to bacterial and eukaryotic pathogens, as well as commensal colonizers (Box 2). While Sias have been known as pathogen receptors for many decades, much remains to be learned about the many modifications and how they are synthesized and distributed in different organisms and tissues. The recent development of new tools (Box 1) that allow for in situ analysis or removal of modified Sias has expanded our ability to study more complex forms and to answer questions about the roles of these modifications in shaping viral evolution, tissue tropism, and host range. These tools will also allow us to address how Sias with different modifications, including 4-O-Ac, 9-O-Ac, and Neu5Gc (and combinations), shape organismal development and cell-cell communication (see Outstanding Questions). While Sias have a great variety of forms, it is important to remember that they are part of an existing mosaic of glycans that vary from organism to organism, tissue to tissue, and even within a single cell over time. The roles played by modified...
Many bacteria and eukaryotic microbes that must bind to or pass through mucosal surfaces also interact with Sias. The gut microbiome contains many groups that utilize mucin glycans as carbohydrate sources, including Sia [87,88]. Since modified Sias may pose as difficult metabolic substrates, bacterial pathogens with Sia metabolic operons often express O-acetylerases along with sialidases to enhance their function by removing modifications [89]. Sias can also be used as an adhesion point for bacterial colonization or as a method of decorating the bacterial surface and avoiding immune detection [7,90,91]. For example, Neisseria gonorrhoeae displays synthesized or acquired Sias to mask their presence from innate immune responses [92]. For toxins produced by bacterial pathogens, such as Salmonella enterica serovar Typhi, the Neu5Ac Sia can act as a toxin receptor and determine susceptibility to certain bacterial infections [93]. Modifications of the target Sia could lead to resistance to that toxin. The activities of different pathogen sialidases and esterases can result in complex interactions during mixed infections of various infectious agents, as seen in viral–bacterial superinfections [94–96]. In the case of influenza, the activity of NA is known to contribute to secondary bacterial pneumatic infections – the removal of Sia molecules appears to enhance the binding of Streptococcus pneumoniaeiae via alveolar galectins as well as promoting bacterial growth by enhancing free Sia carbon substrates [97,98]. In addition to bacterial species, many eukaryotic parasites also use Sia as a receptor or adhesion point. Both Plasmodium falciparum and Trypanosoma cruzi use Sias expressed on erythrocytes to facilitate infection and express multiple Sia-binding proteins [7,99]. The relative effects of modified Sias on macroparasites are unexplored.

Sias in these complex environments and how they vary between organisms are key questions in glycobiology and in host–pathogen interaction studies.

Acknowledgments

We thank Aij Varki (UC-San Diego), members of the Parrish laboratory, and two anonymous reviewers for feedback and critical analysis of the manuscript. Support for this work was received from the NIAID Centers of Excellence in Influenza Research and Surveillance (HHSN272201400008C), NIH grant R01 GM080533-05 to CRP, and NIH Common Grant (U01CA199793).

References


57. Xu, G. et al. (1998) Sialidase of swine influenza A viruses: variation of the recognition specificities for sialy linkages and for the molec- ular species of sialic acid with the year of isolation. Glycoconjugate J. 12, 156–161


72. Helander, A. et al. (2003) The viral sigma1 protein and glycoconjugates containing alpha2,3-linked sialic acid are involved in type 1 reovirus adherence to M cell apical surfaces. J. Virol. 77, 7964–7977
75. Liu, Y. et al. (2015) Sialic acid-dependent cell entry of human enterovirus D68. Nature Commun. 6, 8865
76. Nickbbei, M.R. et al. (2003) Enterovirus 70 binds to different glycoconjugates containing alpha2,3-linked sialic acid on different cell lines. J. Virol. 79, 7097–7094
98. Siegel, S.J. et al. (2014) Influenza promotes pneumococcal growth during coinfection by providing host sialylated substrates as a nutrient source. Cell Host Microbe 16, 55–67