Synthetic O-Acetyl-N-glycolylneuraminic Acid Oligosaccharides Reveal Host-Associated Binding Patterns of Coronaviral Glycoproteins

Published as part of the ACS Infectious Diseases virtual special issue “Glycoscience in Infectious Diseases”.

Zeshi Li, Lin Liu, Luca Unione, Yifei Lang, Raoul J. de Groot, and Geert-Jan Boons*

Cite This: https://doi.org/10.1021/acsinfecdis.2c00046

ABSTRACT: A panel of O-acetylated N-glycolylneuraminic acid oligosaccharides has been prepared by diversification of common synthetic precursors by regioselective de-O-acetylation by coronaviral hemagglutinin-esterase (HE) combined with C7-to-C9 acetyl ester migration. The resulting compound library was printed on streptavidin-coated glass slides to give a microarray to investigate receptor binding specificities of viral envelope glycoproteins, including spike proteins and HEs from animal and human coronaviruses. It was found that the binding patterns of the viral proteins for N-glycolylated sialosides differ considerably from those of the previously synthesized N-acetylated counterparts. Generally, the spike proteins tolerate N-glycolyl modification, but selectivities differ among viruses targeting different hosts. On the other hand, the lectin domain of the corresponding HEs showed a substantial decrease or loss of binding of N-glycolylated sialosides. MD simulations indicate that glycolyl recognition by HE is mediated by polar residues in a loop region (109−119) that interacts with the 5′-N-glycolyl moiety. Collectively, the results indicate that coronaviruses have adjusted their receptor fine specificities to adapt to the sialoglycome of their host species.

KEYWORDS: coronaviral glycoproteins, receptor binding, acetylated N-glycolylneuraminic acid oligosaccharides, spike proteins, sialoglycomes

INTRODUCTION

Evolutionary pressure exerted by pathogenic microbes has caused glycan biosynthesis to diverge, and as a result, glycomes of even closely related species may differ considerably.1,2 This is well-documented in sialic acids, a diverse family of negatively charged, nine-carbon keto-sugars distributed across all domains of life. In vertebrates, N-acetylneuraminic acid (Neu5Ac) and N-glycolynleuraminic acid (Neu5Gc) are the most common forms of sialic acids.3 The latter modification is biosynthesized by hydroxylation of the N-acetyl moiety at C5 of cytidine-5′-monophosphoryl-Neu5Ac (CMP-Neu5Ac) by the enzyme CMP-N-acetylneuraminic acid hydroxylase (CMAH).4 Various sialyl transferases can utilize the resulting CMP-Neu5Gc to biosynthesize Neu5Gc-containing glycoconjugates. In humans, the gene encoding CMAH has been inactivated and, as a result, cannot biosynthesize CMP-Neu5Gc. Although Neu5Gc is considered a non-human sialic acid, it can be incorporated into human glycans by utilizing exogenous Neu5Gc from dietary sources.5

Neu5Ac and Neu5Gc commonly occur at the termini of complex glycans and can be attached to a penultimate sugar in several glycosidic forms: α2,3-linked to galactose (Gal), α2,6-linked to Gal, N-acetylgalactosamine (GalNAc) or N-acetylglucosamine (GlcNAc), or α2,8-linked to another sialic acid. Neu5Ac and Neu5Gc can also undergo hydroxyl modifications considered a non-human sialic acid, it can be incorporated into human glycans by utilizing exogenous Neu5Gc from dietary sources.6

Received: January 20, 2022
including mono-, di-, tri-, and even tetra-O-acetylation at C4, C7, C8, and/or C9 position(s), giving rise to over 10 molecular variants (Figure 1). The biological roles of specific variants have been difficult to dissect, partly due to their lability and a lack of well-defined compounds.

O-Acetylated variants of sialic acid serve as receptors for many viruses, including embecoviruses (family Coronaviridae), toroviruses (Tobaniviridae), influenza C and D viruses, and infectious salmon anemia virus (Orthomyxoviridae). Receptor binding preferences critically determine viral host range and tissue/organ tropism. Insight into the receptor preferences of these viruses will provide important knowledge of the factors driving and hampering cross-species transmission, thereby facilitating prediction and prevention of emerging zoonotic outbreaks.

Recently, we demonstrated that the spikes of OC43 and HKU1, which both are human coronaviruses of zoonotic origin that can cause upper respiratory tract infections, have convergently evolved to selectively engage with 9-O-acetylated α2,8-linked disialosides for cell entry. Furthermore, it was found that the spikes of related animal viruses exhibit distinctive selectivities for various O-acetylation and glycosidic linkages. It remained elusive whether S-N-glycolylation can modulate virus–sialic acid interactions. Given that Neu5Gc repertoires differ considerably among mammalian species, it is conceivable that viruses targeting specific hosts may exhibit differential preferences for acetylated sialosides having a S-N-glycolyl moiety.

Herein, we describe the chemoenzymatic synthesis of a library of N-glycolyl-O-acetylneuraminic acid oligosaccharides by hemagglutinin-esterase (HE)-mediated regioselective de-O-acetylation in combination with controlled base-mediated C7-to-C9 acetyl ester migration of common synthetic 4,7,9-tri-O-acetylated sialosides. It gave 14 O-acetylated sialosides from two common precursors in high overall yields. The resulting compounds and previously synthesized Neu5Ac counterparts were printed as a glycan microarray that was probed for binding of recombinant HEs and spike proteins from various beta-coronaviruses that target different host species. It was found that the lectin domain of HE of all examined viruses showed a reduced or no affinity for the glycolyl-modified structures. The spike of the corresponding viruses did bind...
Neu5Gc-modified glycans, but the binding patterns differed from those observed for Neu5Ac-containing glycans.

**RESULTS AND DISCUSSION**

Previously, we described a strategy for the preparation of a panel of O-acetylated Neu5Ac-containing oligosaccharides starting from α2,3- and 2,6-linked common precursors 1a and 2a, respectively, having acetyl esters at C4, C7, and C9 (Figure 2a). These compounds could be diversified by treatments with HEs from bovine and murine coronaviruses (BCoV and MHV-S, respectively), which cleave 9- and 4-O-acetyl esters with exquisite regioselectivity, respectively, in combination with controlled base-induced acetyl ester migration from C7 to C9 (Figure 2b). We envisaged that if HEs can also cleave glycolyl-modified O-acetate sialosides, the methodology can be extended to the preparation of an almost complete panel of Neu5Gc sialoglycans with O-acetylation patterns mirroring that of previously prepared Neu5Ac counterparts. The resulting glycan library would cover most of the commonly found O-acetylated sialic acids in humans and other mammals, allowing in-depth characterization of virus–receptor interactions. To implement the chemoenzymatic strategy for preparing O-acetylated Neu5Gc derivatives, we chemically synthesized common precursors 3a and 4a by adapting the late-stage N- and O-acetylation steps employed for the synthesized sialosides 1a and 2a, which were then subjected to HE treatments and controlled C7-to-C9 O-acetyl migration.

The preparation of O-acetylated sialo-glycans comes with a number of challenges. These included stereocontrol of alpha-selective chemical sialylation and the unstable nature of acetyl esters on sialic acid. In particular for N-glycolylated sialosides, an additional challenge lies in C5-amine derivatization, which requires orthogonal protecting groups for the secondary C2-amino of glucosamine and the C5-amine of neuraminic acid. Building blocks 5, 7, and 9 (Figure 3a) were employed for the chemical synthesis of trisaccharides 10 as previously reported,8 which through a number of manipulations was expected to provide common precursor 3a. The 5-N,4-O-carbonyl of sialyl donor 5 ensured high reactivity and facilitated α-glycosidic bond formation.11 The galactosyl donor 7 could be coupled with acceptor 9 to give a β-linked disaccharide that, after deacetylation, provided a highly reactive glycosyl acceptor for regioselective glycosylations with 5. After sialylation, the naphthyl-2-methyl (Nap) protecting group allowed protection of the C4 neighboring alcohol via oxidative ring closure, forming an acetal.15 The 2-amino moiety of glucosamine 9 is masked as an azide, which allows selective deprotection of the 5-N,4-O-carbonyl of sialic acid for selective installation of a glycolyl moiety.

Treatment of 10 with aqueous KOH resulted in the removal of base-sensitive protecting groups, including the 4,5-carbamate, thereby exposing the C5-amine of sialic acid (Figure 3b). After desalting by LH-20 size exclusion column chromatography, the free amine of the resulting compound was acylated with active ester to install a properly protected anomeric spacer using azide-modified activated ester 14 (13 → 15), and finally acetylation of the remaining amine and

---

**Figure 3.** Chemoenzymatic synthesis of Neu5Gc glycosides. (a) Glycosyl donors and acceptor for chemical assembly of trisaccharide common precursors. Building blocks 5 and 7 were used for the synthesis of α2,3-linked sialoside 3a, and 6 and 8 for α2,6-linked sialoside 4a. (b) Chemical assembly of α2,3-linked Neu5Gc glycoside common precursor 3a (4,7,9-tri-O-Ac). See ref 8 for synthesis of 10. See Supporting Information for synthesis of α2,6-linked common precursor 4a. (c) HE-catalyzed regioselective de-O-acetylation of 3a and 4a. Complete enzymatic conversion was observed for all steps. Yields were determined after HILIC purification, hence were not quantitative.
hydroxyls using acetic anhydride in pyridine provided trisaccharide 16 in an overall yield of 42% (five steps). Hydrogenation of 16 over Pd(OH)₂ led to the removal of the benzyl ethers and reduction of the azide of the anomeric linker to an amine. The resulting compound was subjected to controlled C8-to-C9 acetyl ester migration using cesium fluoride and acetic acid in DMSO as an aprotic, buffered system. This was followed by selective biotinylation of the amine of the anomeric linker using biotin succinimide to give, after purification by human Fc and detected with AlexaFluor647-conjugated goat anti-human IgG antibody. Solid fill indicates Neu5Ac binding, and pattern fill indicates Neu5Gc binding.

Figure 4. Microarray analysis of coronaviral spike proteins and HEs. (a–d, i) Column graphs for spike proteins (S1α domain) from RbCoV, BCoV, ECoV, CRCoV, and OC43. (e–h) Column graphs for HEs from RbCoV, BCoV, ECoV, and CRCoV. OC43 HE has a non-functional lectin domain and does not bind O-acetylated sialic acid. Protein concentrations: RbCoV S1α, 1 µg/mL; BCoV S1α, 0.3 µg/mL; ECoV S1α, 0.3 µg/mL; CRCoV S1α, 3 µg/mL; OC43 S1α, 0.3 µg/mL; RbCoV HE, 3 µg/mL; BCoV HE, 1 µg/mL; ECoV HE, 1 µg/mL; CRCoV HE, 1 µg/mL. All proteins are tagged with human Fc and detected with AlexaFluor647-conjugated goat anti-human IgG antibody. Solid fill indicates Neu5Ac binding, and pattern fill indicates Neu5Gc binding.  

BCoV and MHV-S HEs can hydrolyze with exquisite regioselectivity the 9- and 4-O-acetyl esters of glycosides of Neu5Ac, respectively. However, the activities of these two enzymes for O-acetylated Neu5Gc have not been systematically investigated. To test if O-acetylated Neu5Gc is a substrate for the two enzymes, the 4,7,9-tri-O-acetylated common precursors 3a and 4a were incubated with BCoV and MHV-S HEs (Figure S1). The former HE readily cleaved the 9-O-acetyl ester to give 3b and 4b, whereas the latter enzyme specifically removed 4-O-Ac to provide 3d and 4d. The enzymatic transformations were monitored by electrospray ionization mass spectrometry (ESI-MS), where loss of only one acetyl moiety (42 units) indicated regio-specific ester hydrolysis. The products were purified by high-performance hydrophilic interaction liquid chromatography (HILIC) and the structures confirmed by 1D and 2D NMR experiments. In particular, the shift of sialic acid protons at C9 (for 3b and 4b) and C4 (for 3d and 4d) to higher field confirmed deacetylation of the corresponding hydroxyls. 4,7-Di-O-acetylated sialosides 3b and 4b were treated with mild base (100 mM ammonium bicarbonate in D₂O), resulting in the migration of the C7 acetyl ester to C9, thereby providing Neu5Gc glycosides 3c and 4c. 1H NMR with water peak suppression showed the disappearance of the C7 proton above 5.0 ppm, confirming migration had occurred. Upon further treatment with BCoV HE, 3c and 4c were converted into 3e and 4e, respectively. The 7-mono-O-acetylated 3f and 4f were obtained by cleaving the 9-O-acetyl of 3d and 4d using BCoV HE. In 100 mM ammonium bicarbonate, the C7 acetyl ester of 3f and 4f readily migrated to C9 to afford 3g and 4g.  

Having prepared a panel of O-acetylate Neu5Gc glycosides, we focused our attention on examining the receptor specificities of Enbecovirus-derived envelope glycoproteins, spike, and HE. These viruses bind to O-acetylated sialosides via their spike (S) for virion uptake. Destruction of decoy receptors and virion release is mediated by HE, which is comprised of a lectin domain that recognizes O-acetylated sialosides and an esterase domain that hydrolyzes acetyl esters.
of sialosides, thereby destroying decoy receptors. The lectin domain has an important function: to enhance the catalytic efficiency of HE for multivalently presented substrates. Thus, the O-acetylated Neu5Gc glycosides (3a–g and 4a–g) were printed on streptavidin-coated slides. The Neu5Ac derivatives having the same O-acetylation pattern were incorporated in the array for comparison (Figure 4). Receptor binding domains (S1A) of the spike of human coronavirus OC43, bovine coronavirus (BCoV), rabbit coronavirus (RbCoV), equine coronavirus (ECoV), and canine respiratory coronavirus (CRCoV) were expressed as Fc-fusion proteins in HEK293T cells. In addition, enzymatically inactive forms of the ectodomains of the corresponding HEs were also expressed as Fc-fusion proteins. The viral proteins were exposed to the array, and detection of binding was accomplished using an anti-Fc antibody modified with AlexaFluor647. HEs of all examined viruses showed a decrease or loss of binding of the glycolyl-modified sialosides. In the case of BCoV, the additional hydroxyl group at the C5 substituent resulted in a moderate reduction in binding. On the other hand, the HEs of CRCoV and RbCoV showed a dramatic reduction in binding for all 5-N-glycolylated structures, and signals for the 7,9-di- and 4,7,9-triacetylated forms were detected slightly above baseline. In the case of ECoV HEs, no detectable binding of glycolyl-modified structures was observed. The HEs of human coronaviruses have lost the ability to bind sialosides and therefore were not examined.

Unlike the HEs, the spike-derived proteins displayed variable responses to S-N-glycolylation. In the cases of BCoV and CRCoV spikes, 5-N-glycolylation resulted in an increase in the binding of 4,9-di-O- and 4,7,9-tri-O-acetylated Neu5Gc sialosides. Conversely, the spikes of human and canine viruses, OC43 and CRCoV, respectively, gave considerably lower responses for S-N-glycolylated structures. The 4,9-di-O-acetylated form was an exception, displaying a striking increase in binding compared to the corresponding Neu5Ac glycosides. Interestingly, the effect of S-N-glycolylation on the ECoV spike was glycosidic linkage-selective. In the case of the α2,3-linked sialosides, 5-N-glycolylation caused a substantial reduction of binding, whereas the same modification was better tolerated for the corresponding α2,6-linked glycosides. The binding-decreasing effect of 5-N-glycolylation was particularly striking for 9-monoacetylated sialosides. No binding was observed when the viral proteins were subjected to another microarray populated with common glycan structures, including many sialosides (Figure S2), further supporting a requirement of O-acetylation of sialosides for binding.

Recently, we provided an atomic-level interpretation of the recognition of Neu5Ac by the HEs of several host species by using a combination of NMR experiments and MD simulation. It demonstrated the importance of complementarity of the acetyl moieties of the sialosides and hydrophobic pockets of HE for binding. The microarray data indicate that the lectin domain of the HEs of ECoV and RbCoV poorly recognize glycolyl-modified sialosides, whereas this modification is tolerated by the HE from BCoV. To offer a structural explanation of these observations, MD simulations (for details see the Supporting Information) were performed on the three...
HEs in complex with the 4,7,9-triacylated form of the 5-N-glycolylated sialoside 3a.

MD simulations of the complexes of compound 3a with the BCoV, ECoV, and RbCoV HEs were performed in explicit water. The BCoV HE:3a complex was built by superimposing 3a onto 4,9-di-O-acetylated sialic acid in the holo-structure of BCoV HE (PDB 3CLS). No crystal structures are available for ECoV and RbCoV, and therefore homology models were generated. The resulting geometries were energy-minimized using the AMBER program, and their structural models were generated. The resulting geometries were energy-minimized using the AMBER program, and then subjected to 100 ns molecular dynamics simulation using the AMBER program, exploiting the GLYCAM 06j-1 force fields for the carbohydrate and the proteins, respectively.

A comparison of the MD-derived structures with those previously obtained for the 5-N-Ac analogue showed that the 9-O-acetyl, the glycerol chain, and the 7-O-acetyl group are very similarly accommodated in hydrophobic pockets defined by residues 184, 266, and 267, and residues 161, 207, and 114, respectively. Similar binding modes were also observed for polar groups of the sialoside with hydrogen-bonding interactions of the carboxylic acid with residue 214, the NH of the amide at C5 with L212, and the hydroxyl at C8 with the S213.

The MD simulations indicate that the hydroxyl of 5-N-glycolyl moiety disrupts interactions with the hydrophobic residues F211 and F/L245 (Figure S3), which probably results in a reduction in binding affinity. It also indicates that residues in the loop region (aromatic acid residues 109–119) of the sialic acid binding domain are responsible for the observed differences in Neu5Gc binding of the different HEs, and in the case of BCoV, additional polar interactions are established to compensate for the loss of apolar contacts. In particular, residues 114–116 face the glycolyl amide of the ligand, and in the case of BCoV, these residues are polar, entailing a TTS sequon. The two threonine residues can establish interactions with the amide at C5 of the sialoside by engaging either directly with the carbonyl oxygen or through water molecules. In addition, S116 can make a hydrogen bond with the hydroxyl of glycolyl moiety (Figure S5a). These intermolecular interactions are possible because of appropriate positioning of ligand and loop (Figure 5d).

In the case of ECoV and RbCoV, one of the 114–116 residues is replaced by an apolar amino acid (TAS and TSF sequons, respectively), which reduces the potential for hydrogen bond formation. Furthermore, it caused a distancing of ligand and loop, further limiting the possibility for hydrogen bond formation with the remaining polar amino acids. In the case of ECoV, residue 115 is alanine, which cannot establish a hydrogen bond with the C5 amide. Furthermore, S116 is not sufficiently close to make a hydrogen bond with the hydroxyl of the glycolyl moiety (Figure 5b). In the case of RbCoV, residue 116 is a Phe, which cannot establish a hydrogen bond with the hydroxyl of glycolyl, and although residues 114 and 115 are Thr and Ser, respectively, they do not engage with amide of the glycolyl moiety because of displacement of the loop (Figure 5c,d). Collectively, the in silico studies indicate that the residues in the 109–119 loop modulate 5-N-glycolyl recognition.

The differences in receptor specificity of the spike and lectin domain of HE of the same virus may be due to the fact that they evolved to preferentially recognize different extracellular matrices and cell surface glycoconjugates. The spike has as a main function to bind to cell surface sialylated glycolipid receptors to promote virion uptake. The lectin domain of HE facilitates initial attachment by binding to O-acetylated glycans present in the mucus of the host. It regulates sialate-O-acetyl esterase activity by increasing the enzyme activity of clustered receptors. As a result, decoy receptors are destroyed, facilitating virion migration through mucus. The spike protein may also participate in attaching to multivalent decoy receptors, and for BCoV and OC43, it has been shown that HE and the spike function as a two-component system for dynamic receptor interactions.

As a result of adaptation to their host sialoglycomes, which can differ dramatically between various species, the binding selectivities of spike and HE are fine-tuned to achieve viral fitness. It is conceivable that, when Neu5Gc-modified receptors are present in smaller quantities than a Neu5Ac counterpart, the degradation of the former by HE should be reduced accordingly. This can be achieved via decreasing the avidity of the lectin domain for Neu5Gc, while the esterase domain remains catalytically active. In bovine, both 5-N-glycolylated sialylglycolipids (cell surface receptors) and mucin (decoy receptors) are highly abundant, which is in line with the observation that both the spike and HE of BCoV tolerate the 5-N-glycolyl modification. In rabbits, Neu5Gc gangliosides are highly expressed across multiple tissues. Consistently, the spike protein showed excellent binding to Neu5Gc structures. Mucins of rabbits appear to have a lower level of 5-N-glycolylation compared to 5-N-acetylation, which may explain the reduced tolerance of the HE lectin domain for O-acetylated Neu5Gc sialosides. Likewise, in equines, Neu5Gc has been found largely on cell surface glycoconjugates, and it appears that secreted mucus harbor a greatly reduced level of Neu5Gc. This is consistent with the observation that the spike moderately tolerated Neu5Gc sialosides, whereas this is not the case for HE. It is likely that RbCoV and ECoV have tuned-down the hydrolytic activities for 9-O-acetylated Neu5Gc glycans in response to the smaller quantities present in the mucus of these animals to facilitate initial attachment. An extreme example of decreasing catalytic efficiency by compromising the lectin domain binding is the HE of OC43, which harbors a non-functional lectin domain and, as a result, hydrolyzes multivalent 9-O-acetylated sialosides much slower than the HE of ancestral BCoV.

BCoV, OC43, and CRCoV spikes and HEs exhibit much lower affinities for Neu5Gc. Recently, it was observed that a 5-N-glycolyl moiety is not tolerated by human influenza C virus hemagglutinin–esterase-fusion protein but is recognized by the distantly related influenza D virus, which infects ruminants or swine. The loss or reduction of Neu5Gc binding is likely a result of viral adaptation to their host sialoglycomes, which were surprising because humans do not harbor 4-O-acetylated sialosides. It is possible that the recognition of this sialoside is a remnant from an ancestral virus infecting a different host species. The observations warrant structural studies of spike proteins in complex with O-acetylated sialic acid receptors other than the 9-monooacylated form. Furthermore, it calls for more detailed characterization of O-acetylated sialosides in the extracellular matrix and on the cell
surface to further rationalize receptor specificities of the spike and the lectin domain of HE of the various viruses.

**CONCLUSIONS**

We have found that the hemagglutinin-esterases (HEs) of BCoV and MHV-S can readily cleave O-acetylated sialosides, opening the possibility for HE-mediated chemoenzymatic preparation of O-acetylated Neu5Gc oligosaccharides. The resulting compounds were printed as glycan microarrays for in-depth analysis of the receptor binding preference of a panel of coronaviral S proteins and HEs. It was found that the spike proteins exhibit distinct but promiscuous binding patterns for O-acetylated, N-glycolylated sialosides. The lectin domain of the examined HE exhibited a moderate to large reduction in binding. The observations indicate the sialic acid populations, or sialome, across the corresponding hosts may vary considerably in composition, and accordingly, the receptor binding specificities of the viral glycoproteins have been fine-tuned for optimal interaction with specific sialoside forms for the virus to breach the species barrier and become adapted in the new hosts. We reveal by MD simulations that glycolyl recognition is modulated by polar residues in a loop region (109−119) close to the S-N-glycolyl moiety. In addition to host−guest interactions, Neu5Gc plays a significant role in chronic inflammation and cancer progression, likely by inducing anti-Neu5Gc antibodies. This repertoire remains to be characterized in more depth with a full spectrum of Neu5Gc variants, particularly as antibodies against O-acetylated sialic acid have been reported.57,38 We envision that the O-acetylated Neu5Gc sialoside library described here will facilitate fingerprinting of the anti-Neu5Gc response, allowing dissection of the pathological relevance of Neu5Gc sialoglycans.

**METHODS**

**General Methods.** Thin-layer chromatographic analysis was performed using precoated silica gel 60 F-254 plates (Merck). The plates were either visualized under a UV lamp or stained with ceric ammonium molybdate (5 g of Ce(SO4)2·25H2O, 50 mL of concentrated H2SO4, 450 mL of H2O). Synthetic intermediates were purified by column chromatography using flash silica gel (60 Å, from Silicycle, Canada). Deprotection compounds were purified by size exclusion chromatography using P-2 biogel (Biorad) and collected with a BioFrac fraction collector (Biorad). The resulting products were analyzed by HPLC using an analytical HILIC column (XBridge HILIC 5 μm, 4.6 × 250 mm) prior to preparative-scale HPLX purification (XBridge BEH Prep Amide 5 μm, 10 × 250 mm). 1H and 13C NMR spectroscopy was conducted on an Agilent 400-DR or Bruker 600 UltraShield instrument. Reaction monitoring using ESI-MS was performed on a Bruker microOTOF-Q II ESI mass spectrometer. HRMS data was collected on an Agilent 6560 ion mobility Q-ToF MS.

**General Procedure for Hydrogenation, Acetyl Ester Migration, and Biotinylation (Procedure I).** In a two-necked flask, the starting material (e.g., 16) was dissolved in a mixture of THF and water (1:1, v/v) containing acetic acid (5 equiv relative to starting materials). Degussa-type palladium hydroxide was added while the solution was kept under an atmosphere of hydrogen at atmospheric pressure. The progress of the reaction was monitored by ESI-MS (direct infusion, negative mode). Upon completion of the reaction, the mixture was filtered through a pad of Celite, and the filtrate was concentrated under reduced pressure in a water bath that was kept at ~20 °C. The residue was lyophilized and then dissolved in DMSO-d6 to which acetic acid-d4 (3 equiv) and cesium fluoride (6 equiv) were added. An aliquot (~500 μL) was taken to record a 1H NMR spectrum of the starting material. The mixture was heated (80 °C) for 14−20 h, after which an end-point 1H NMR was recorded. A shift of Sia H-8 from ~5 ppm to ~4 ppm indicated completion of the reaction. The mixture was lyophilized and the residue dissolved in DMSO to which a solution of biotin N-hydroxysuccinimide ester in DMSO was added. The progress of the reaction was monitored by ESI-MS (direct infusion, negative mode). At completion, the solution was freeze-dried and the residue desalted by biogel P2 size exclusion chromatography (eluents: 5% n-butanol, 4 °C) and then further purified by HPLC using a HILIC column (eluents: acetonitrile/water 95% to 65%, 0.1% formic acid, 2.4 mL/min, 50 min) to give the biotinylated common precursor sialosides 3a and 4a.

**General Procedure for HE-Mediated De-O-acetylation (Procedure II).** α2,3- or 2,6-linked common precursor (~2 mg, 3a or 4a, respectively) was dissolved in freshly prepared ammonium formate buffer (50 mM, pH 6.4, 5 mM substrate concentration). The solution of HE was added to achieve a working concentration of 10 μg/mL for both HEs. The progress of the reaction was monitored by ESI-MS (direct infusion, negative mode). At completion, the solution was freeze-dried and purified by HPLC using a HILIC column (eluents: acetonitrile/water 95% to 65%, 0.1% formic acid, 2.4 mL/min, 50 min).

**General Procedure for Base-Mediated Acetyl Ester Migration (Procedure III).** 7-O-Acetylated or 4,7-di-O-acetylated sialoside (2−4 mg) was dissolved in freshly prepared 100 mM ammonium bicarbonate in D2O to achieve a concentration of 2.5−5 mM, resulting in a pd(H) of ~8. The solution was transferred to an NMR tube, and a starting 1H NMR spectrum was recorded with water suppression. The reaction mixture was heated (37 °C), and the progress of the reaction was monitored by recording 1H NMR spectra at different time points. The disappearance of the signal at ~5.15 ppm, which corresponds to the acetylated H7 of Neu5Gc, indicated O7-to-O9 migration. The solution was then lyophilized 2−3 times to completely remove ammonium bicarbonate.

**Microarray Printing and Screening.** The compounds were immobilized on SuperStreptavidin microarray substrate slides (ArrayIt Inc.) using a sciFLEXARRAYER S3 non-contact microarray printer equipped with a PDC80 nozzle (Scienion Inc.). The sialosides were dissolved in Milli-Q water at 1 mM to make stock solutions which were stored at ~80 °C. The printing solutions were freshly prepared at 100 μM by diluting the stock solutions in ice-cold printing buffer (10 mM PBS, pH 7.0). The compounds were printed in replicates of six with print volume ~400 pL, at 20 °C and 50% humidity. The resulting slides were stored at 4 °C for up to 1 month. For binding studies, the slides were blocked with TSM binding buffer (20 mM Tris-HCl, pH 7, 150 mM NaCl, 2 mM CaCl2 and 2 mM MgCl2, 0.05% Tween-20, 1% BSA) for 1 h at 4 °C. Human-Fc fused viral proteins at the indicated concentrations were pre-mixed with goat anti-human IgG (AlexaFluor647 conjugated, 109-605-008, Jackson ImmunoResearch) in a 1:1 weight ratio in TSM binding buffer and incubated at 4 °C for 1

https://doi.org/10.1021/acsinfecdis.2c00046

ACS Infect. Dis. XXXX, XXX, XXX−XXX
h. The slides were then incubated with the pre-complexed virolectin–secondary antibody solutions at 4 °C for 1 h. The slides were washed with TSM washing buffer (20 mM Tris-HCl, pH 7, 150 mM NaCl, 2 mM CaCl₂ and 2 mM MgCl₂, 0.05% Tween-20), then with TSM buffer (20 mM Tris-HCl, pH 7, 150 mM NaCl, 2 mM CaCl₂ and 2 mM MgCl₂), and finally with water. The slides were spun dry, and the fluorescence intensity was measured using a GenePix 4000B microarray scanner (Molecular Devices). The scanning was performed at the appropriate gain to avoid saturation of the signal. GenePix Pro 7 software (Molecular Devices) was used to analyze the images, and the digitized raw data was processed using a macro run in Microsoft Excel (available at https://zenodo.org/record/5146251).

Molecular Dynamics Simulation. The initial geometry of the ligand was built in the Glycam web (http://glycam.org). The pdb coordinates for BCoV HE were derived from the crystal structure Protein Database (PDB) 3cl5, whereas those for RbCoV and ECoV HEs were generated as homology pdb using SWISS-MODEL. The Neu5Gc residue of the glycan was superimposed onto the corresponding sugar in the deposited 3cl5 structure. The resulting binding poses were used as starting points for MD simulations. The MD simulations were performed at the appropriate gain to avoid saturation of the signal. GenePix Pro 7 software (Molecular Devices) was used to analyze the images, and the digitized raw data was processed using a macro run in Microsoft Excel (available at https://zenodo.org/record/5146251).

Molecular Dynamics Simulation. The initial geometry of the ligand was built in the Glycam web (http://glycam.org). The pdb coordinates for BCoV HE were derived from the crystal structure Protein Database (PDB) 3cl5, whereas those for RbCoV and ECoV HEs were generated as homology pdb using SWISS-MODEL. The Neu5Gc residue of the glycan was superimposed onto the corresponding sugar in the deposited 3cl5 structure. The resulting binding poses were used as starting points for MD simulations. The MD simulations were performed using the Amber16 program with the ff14SB force field parameters for protein and GLYCAM_06hj-1 for the saccharides. The resulting 3D geometries were placed into an octahedral box of explicit TIP3P water molecules with a buffered distance between solute and box of 10 Å. Counters were added to maintain electroneutrality. Two consecutive minimization stages were performed involving only the water molecules and ions and then all-atoms minimization with a higher number of cycles, using the steepest-descent algorithm. The system was subjected to two rapid MD simulations (heating and equilibration). During the first simulation, the systems were heated from 0 to 300 K under a constant pressure of 1 atm at periodic boundary conditions. Next, the systems were equilibrated for 2 ns with a 2 fs time step at constant volume and temperature. The equilibrated structures were the starting points for the final MD simulations at constant temperature (300 K) and pressure (1 atm). MD simulations of 100 ns without constraints were recorded, using an NPT ensemble with periodic boundary conditions, a cutoff of 10 Å, and the particle mesh Ewald method. A total of 10,000,000 MD steps were run with a time step of 1 fs per step. Coordinates and energy values were recorded every 50,000 steps for a total of 200 MD models. Analysis of the intermolecular hydrogen bonds and hydrophobic interactions was performed along the MD trajectory using the cpptraj module included in the Amber-Tools 16 package.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsinfecdis.2c00046.

Figure S1, synthesis of α2,6-sialoside 4a; Figure S2, microarray results using an in-house library encompassing common glycan structures; Figure S3, structural comparison of HEs in complex with 4,7,9-tri-O-acetylated sialic acid; materials and methods; synthetic procedures; NMR spectra; and description of the common glycan library (PDF)

Source file microarray data (XLSX)

Author Information
Corresponding Author
Geert-Jan Boons — Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands; Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602, United States; Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CH Utrecht, The Netherlands; Chemistry Department, University of Georgia, Athens, Georgia 30602, United States; orcid.org/0000-0003-3111-5954; Email: gjboons@crc.uga.edu

Authors
Zeshi Li — Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands
Lin Liu — Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602, United States; orcid.org/0000-0002-0310-5946
Luca Unione — Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands
Yifei Lang — Virology Division, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, The Netherlands
Raoul J. de Groot — Virology Division, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, The Netherlands

Complete contact information is available at: https://pubs.acs.org/10.1021/acsinfecdis.2c00046

Notes
The authors declare no competing financial interest.

Acknowledgments
The research was supported by TOP-PUNT grant 718.015.003 of The Netherlands Organization for Scientific Research (G.J.B.); the Human Frontier Science Program Organization (HFSP) grant LTF00747/2018-C (L.U.); ECHO 711.011.006 of the Council for Chemical Sciences of The Netherlands Organization for Scientific Research (R.J.d.G.); and China Scholarship Council 2014-03250042 (Y.L.).

References