Human coronaviruses OC43 and HKU1 bind to 9-*O*-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A

Ruben J. G. Hulswit[a,1], Yifei Lang[a,1], Mark J. G. Bakkers[a,1,2], Wentao Li[a], Zeshi Li[b], Arie Schouten[c], Bram Ophorst[a], Frank J. M. van Kuppeveld[a], Geert-Jan Boons[b,d,e], Berend-Jan Bosch[a], Eric G. Huizinga[c], and Raoul J. de Groot[a,3]

[a]Virology Division, Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, 3584 CH Utrecht, The Netherlands; [b]Department of Chemical Biology and Drug Discovery and Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CG Utrecht, The Netherlands; [c]Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Faculty of Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands; [d]Department of Chemistry, University of Georgia, Athens, GA 30602; and [e]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602

Edited by Mary K. Estes, Baylor College of Medicine, Houston, TX, and approved December 19, 2018 (received for review June 6, 2018)

Human betacoronaviruses OC43 and HKU1 are endemic respiratory pathogens and, while related, originated from independent zoonotic introductions. OC43 is in fact a host-range variant of the species *Betacoronavirus-1*, and more closely related to bovine coronavirus (BCoV)—its presumptive ancestor—and porcine hemagglutinating encephalomyelitis virus (PHEV). The β1-coronaviruses (β1CoVs) and HKU1 employ glycan-based receptors carrying 9-*O*-acetylated sialic acid (9-*O*-Ac-Sia). Receptor binding is mediated by spike protein S, the main determinant of coronavirus host specificity. For BCoV, a crystal structure for the receptor-binding domain $S1^A$ is available and for HKU1 a cryoelectron microscopy structure of the complete S ectodomain. However, the location of the receptor-binding site (RBS), arguably the single-most important piece of information, is unknown. Here we solved the 3.0-Å crystal structure of PHEV $S1^A$. We then took a comparative structural analysis approach to map the β1CoV S RBS, using the general design of 9-*O*-Ac-Sia-binding sites as blueprint, backed-up by automated ligand docking, structure-guided mutagenesis of OC43, BCoV, and PHEV $S1^A$, and infectivity assays with BCoV-S–pseudotyped vesicular stomatitis viruses. The RBS is not exclusive to OC43 and related animal viruses, but is apparently conserved and functional also in HKU1 $S1^A$. The binding affinity of the HKU1 S RBS toward short sialoglycans is significantly lower than that of OC43, which we attribute to differences in local architecture and accessibility, and which may be indicative for differences between the two viruses in receptor fine-specificity. Our findings challenge reports that would map the OC43 RBS elsewhere in $S1^A$ and that of HKU1 in domain $S1^B$.

coronavirus | spike | 9-*O*-acetylated sialic acid | OC43 | HKU1

Coronaviruses (CoVs; order *Nidovirales*, family *Coronaviridae*) are enveloped positive-strand RNA viruses of mammals and birds. So far, four coronaviruses of zoonotic origin are known to have successfully breached the species barrier to become true human pathogens (1–6). These viruses—NL63, 229E, HKU1, and OC43—are persistently maintained in the human population through continuous circulation. Remarkably, the latter two both belong to a single minor clade, "lineage A," in the genus *Betacoronavirus*. Although generally associated with common colds, HKU1 and OC43 may cause severe and sometimes fatal pulmonary infections in the frail (7, 8), and in rare instances, OC43 may cause lethal encephalitis (9). OC43 and HKU1 are distinct viruses that entered the human population independently to seemingly follow convergent evolutionary trajectories in their adaptation to the novel host (10). OC43 is in fact more related to coronaviruses of ruminants, horses, dogs, rabbits, and swine, with which it has been united in a single species, *Betacoronavirus-1*.

Lineage A betacoronaviruses like HKU1 and OC43 differ from other CoVs in that their virions possess two types of surface projections, both of which are involved in attachment: large 20-nm peplomers or "spikes" that are very much a CoV hallmark and comprised of homotrimers of spike (S) protein, and 8-nm protrusions, unique to this clade, comprised of the homodimeric hemagglutinin-esterase (HE). S is central to viral entry and the key determinant of host and tissue tropism (11). It mediates binding to cell-surface receptors and, upon uptake of the virion by the host cell, fusion between the viral envelope and the limiting endosomal membrane (12). In the case of HKU1 and β1 coronaviruses (β1CoVs), OC43 included, S binds to sugar-based receptor-determinants, specifically to 9-*O*-acetylated sialic acids (9-*O*-Ac-Sias) attached

Significance

Human coronaviruses OC43 and HKU1 are related, yet distinct respiratory pathogens, associated with common colds, but also with severe disease in the frail. Both viruses employ sialoglycan-based receptors with 9-*O*-acetylated sialic acid (9-*O*-Ac-Sia) as key component. Here, we identify the 9-*O*-Ac-Sia–specific receptor-binding site of OC43 S and demonstrate it to be conserved and functional in HKU1. The considerable difference in receptor-binding affinity between OC43 and HKU1 S, attributable to differences in local architecture and receptor-binding site accessibility, is suggestive of differences between OC43 and HKU1 in their adaptation to the human sialome. The data will enable studies into the evolution and pathobiology of OC43 and HKU1 and open new avenues toward prophylactic and therapeutic intervention.

Author contributions: R.J.G.H. and M.J.G.B conceived the study; R.J.d.G. coordinated and supervised the study; R.J.G.H., Y.L., M.J.G.B., W.L., E.G.H., and R.J.d.G. designed research; R.J.G.H., Y.L., M.J.G.B., A.S., B.O., and E.G.H. performed research; W.L., Z.L., G.-J.B., and B.-J.B. contributed new reagents/analytic tools; R.J.G.H. and A.S. performed crystallization experiments; M.J.G.B., Y.L., and E.G.H. refined the PHEV $S1^A$ structure and performed automated ligand docking; Y.L. established assays to detect HKU1 S receptor binding and performed infection experiments with pseudotyped vesicular stomatitis virus; E.G.H. supervised crystal structure analysis; R.J.G.H., Y.L., M.J.G.B., W.L., Z.L., A.S., F.J.M.v.K., G.-J.B., B.-J.B., E.G.H., and R.J.d.G. analyzed data; M.J.G.B. performed in silico structural analysis to identify the RBS; and R.J.G.H., Y.L., M.J.G.B., E.G.H., and R.J.d.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID code 6QFY). The amino acid sequences of the $S1^A$–Fc fusion proteins were deposited in the GenBank database (accession no. MG999832-35).

[1]R.J.G.H., Y.L., and M.J.G.B. contributed equally to this work.

[2]Present address: Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115.

[3]To whom correspondence should be addressed. Email: r.j.degroot@uu.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1809667116/-/DCSupplemental.

Published online January 24, 2019.

as terminal residues to glycan chains on glycoproteins and lipids (13, 14). HE, a sialate-*O*-acetylesterase with an appended 9-*O*-Ac-Sia–specific lectin domain, acts as a receptor-destroying enzyme (15). During preattachment, the receptor-destroying enzyme activity of HE averts irreversible binding of virions to the decoy receptors that are omnipresent in the extracellular environment. Furthermore, at the conclusion of the replication cycle, HE-mediated destruction of intracellular and cell-surface receptors facilitates the release of viral progeny from the infected cell (16). Remarkably, the HEs of OC43 and HKU1 lost their ability to bind 9-*O*-Ac-Sias, because their lectin domain—functional in all other HEs studied so far—was rendered inactive (10). In consequence, the dynamics and extent of virion-mediated destruction of clustered receptor populations were altered presumably to match the sialome of the human respiratory tract and to optimize infection and/or transmission. Whether adaptation to the human host also entailed adaptations in S is not known and as yet cannot be assessed because for HKU1, the S receptor-binding site (RBS) has not been identified and the β1CoV S RBS has not been established with certainty (17).

Crystal structure analysis of orthomyxo-, toro-, and coronavirus HEs complexed to receptor/substrate analogs identified 9-*O*-Ac-Sia binding sites, yielding exquisite insight into the architecture of these sites and into the general principles of ligand/substrate recognition at the atomic level (18–24). Recently, cryoelectron microscopy (cryo-EM) structures were reported for several S proteins, including that of HKU1 (25–29). The findings have greatly increased our understanding of the overall quaternary structure and function of the S homotrimers. Among others, the structures revealed how in betacoronavirus spikes the N-terminal S1 subunit of each S monomer folds into four individual domains—designated A through D as numbered from the N terminus (Fig. 1*A*)—of which domains A and B may function in receptor binding (11). The spikes of β1CoVs OC43 and BCoV bind to *O*-acetylated sialic acid through domain A ($S1^A$), as determined by in vitro binding assays (30). There is limited structural information, however, on how the spikes of HKU1 and the β1CoVs bind their ligands. The apo-structure of the BCoV $S1^A$ lectin domain was solved, but attempts to also solve the holo-structure reportedly failed (17). Based on the galectin-like fold of the $S1^A$ domain and mutational analysis, the RBS was predicted (17). Although this model remains to be confirmed, it has been widely accepted by the field (27, 28, 31–33). Surprisingly, HKU1 S was recently reported to bind to its receptor via a domain other than $S1^A$ (31). Despite the similarity of HKU1 domain A to that of BCoV and OC43, binding of HKU1 $S1^A$ to 9-*O*-Ac-Sia was reportedly not detectable (30, 31). Prompted by these observations and the fact that the predicted RBS in β1CoV $S1^A$ in its design and architecture bears no resemblance to other 9-*O*-Ac-Sia binding sites, we sought to test the published model. The results led us to look for alternative binding sites in the β1CoV $S1^A$ domain through comparative structural analysis and in silico modeling using the general design of HE 9-*O*-Ac-Sia binding sites as a blueprint, backed up by structure-guided mutagenesis. Our findings show that the actual $S1^A$ RBS in β1CoV S maps elsewhere than currently believed. Moreover, we demonstrate that the newly proposed site is not exclusive to β1CoVs, but in fact conserved and functional also in the $S1^A$ domain of HKU1.

Results and Discussion

The $S1^A$ RBS Is Located Elsewhere than Currently Believed. To test the validity of the current model, we measured the effect of substitutions in the proposed RBS using $S1^A$–Fc fusion proteins. The binding properties of the mutated proteins were studied by hemagglutination assay (HAA) with rat erythrocytes and by solid-phase lectin-binding enzyme-linked immune assay (sp-LBA) with bovine submaxillary mucin (BSM) as ligand. These assays are complementary: HAA is the more sensitive of the two,

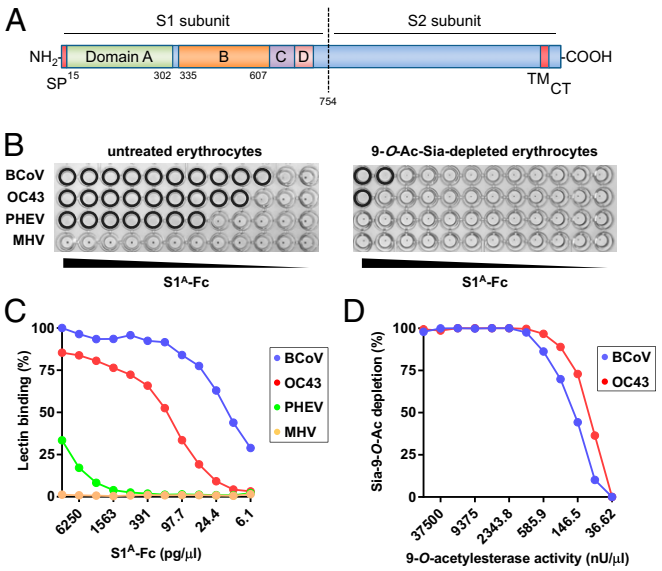

Fig. 1. $S1^A$ domains of β1CoV host range variants all bind 9-*O*-Ac-Sias, but with different affinities. (*A*) Schematic representation of β1CoV spike protein with subunits S1 and S2, and S1 domains A through D indicated. Residue numbering is based on the OC43 strain ATCC/VR-759 spike protein (GenBank: AAT84354.1), with domain boundaries based on the MHV S structure (26). CT, cytoplasmic tail; SP, signal peptide; TM, transmembrane domain. (*B*) $S1^A$ β1CoV variants differ in 9-*O*-Ac-Sia-binding affinity. (*Left*) Conventional HAA with rat erythrocytes (diluted to a final concentration of 0.25% in PBS, 0.1% BSA) and twofold serial dilutions of $S1^A$–Fc fusion proteins of β1CoV variants BCoV-Mebus, OC43-ATCC and PHEV-UU. $S1^A$–Fc of MHV-A59 (starting at 25 ng/μL) was included as a negative control (56). HA was assessed after 2-h incubation at 4 °C. Wells scored positive for HA are encircled. HAAs were repeated at least three times. Representative experiments are shown. (*Right*) $S1^A$–Fc-mediated HA is 9-*O*-Ac-Sia–dependent. HAAs were performed as above, but now with rat erythrocytes, depleted for 9-*O*-Ac-Sias by prior sialate-*O*-acetylesterase treatment, as in ref. 47. (*C*) Differences in $S1^A$-mediated 9-*O*-Ac-Sia–binding affinity among β1CoV variants $S1^A$ protein as demonstrated by sp-LBA. $S1^A$–Fc fusion proteins (twofold serial dilutions, starting at 12.5 ng/μL) were compared by sp-LBA for relative binding to BSM at 37 °C. MHV $S1^A$ was included as a negative control. Experiments were performed at least three times, each time in triplicate, with each data point representing the average of the independent mean values. Mean ± SDs were less than 10%; error bars omitted for esthetical reasons. (*D*) $S1^A$–Fc-binding to BSM is 9-*O*-Ac-Sia dependent. BSM was specifically depleted for sialate-9-*O*-acetyl moieties by on-the-plate sialate-*O*-acetylesterase treatment for 2 h with twofold serial dilutions of soluble hemagglutinin-esterase, as in refs. 10 and 47, starting at 75 μU/μL. Receptor destruction was assessed by sp-LBA with fixed concentrations of BCoV and OC43 $S1^A$-Fc (0.3 ng/μL and 1.2 ng/μL, respectively).

while the results of sp-LBA more precisely reflect differences in binding-site affinity (10). HAA and sp-LBA performed with $S1^A$ of BCoV strain Mebus confirmed the binding to 9-*O*-Ac-Sias (Fig. 1 *B*–*D*). For comparison, the $S1^A$ domains of related β1CoVs OC43 strain ATCC and porcine hemagglutinating encephalomyelitis virus (PHEV) strain UU were included. Again, 9-*O*-Ac-Sia–dependent binding was observed, but the affinity of the $S1^A$ domain of OC43 was ~32-fold lower than that of BCoV as measured by sp-LBA and that of PHEV was lower still (~1,450-fold). Substitution in PHEV and OC43 $S1^A$ of Ala for Tyr^{162}, Glu^{182}, Trp^{184}, and His^{185} (Fig. 2*A*), residues deemed critical for binding of BCoV $S1^A$ (17), indeed resulted in decreased binding (Fig. 2 *B* and *C*), but importantly and as in the original report, none of the mutations gave complete loss-of-function. The strongest effect was observed for Trp^{184}, but its substitution in the context of OC43 $S1^A$ merely reduced binding affinity to that of wild-type PHEV $S1^A$. In a reverse approach, we attempted to identify substitutions

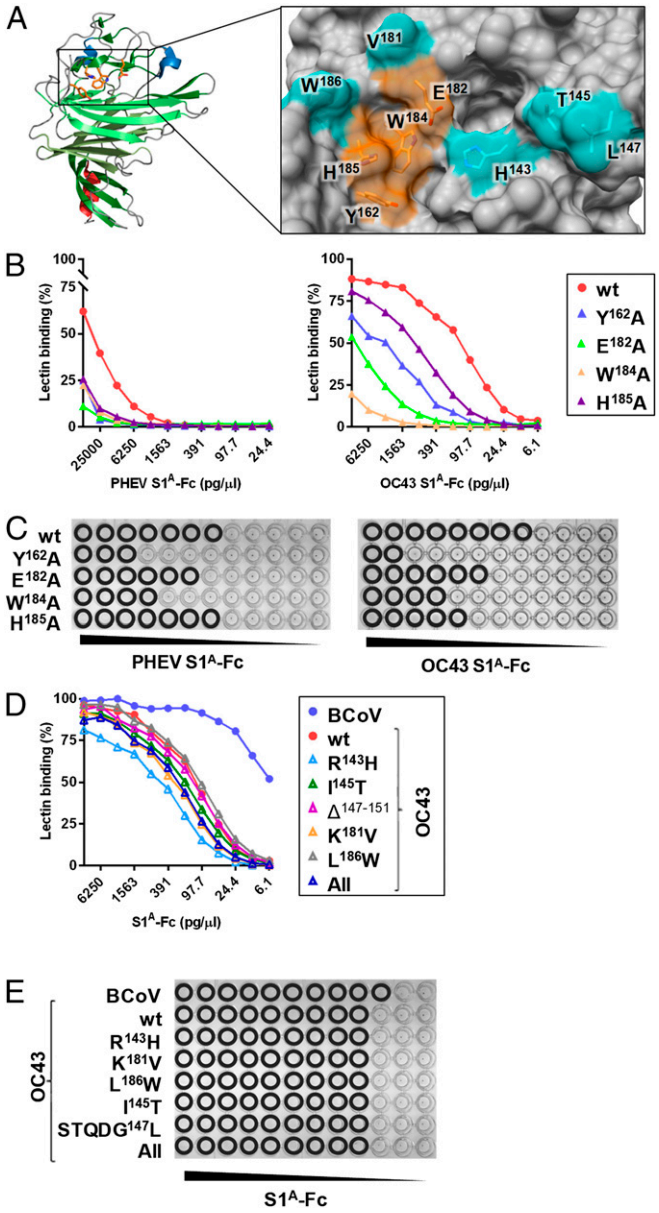

Fig. 2. Evidence from structure-guided mutagenesis against site A being the S1^A RBS. (*A*, *Left*) Cartoon representation of the crystal structure of the BCoV S1^A domain as determined by Peng et al. (17) (PDB ID code 4H14). β-Sheets colored green, the α-helix colored red, 3_{10}-helices colored blue, and side chains of site A residues, supposedly critical for 9-*O*-Ac Sia binding (17), indicated in sticks and colored orange. (*Right*) Site A close-up in surface representation with side chains of residues, supposedly critical for 9-*O*-Ac-Sia binding, indicated as orange sticks. Side chains of residues that differ between BCoV Mebus and OC43 ATCC S1^A are indicated in cyan. (*B*) Substitutions of site A residues in PHEV and OC43 S1^A–Fc result in partial, but not complete loss of RBS binding affinity. In PHEV and OC43 S1^A–Fc, orthologs of BCoV Tyr^{162}, Glu^{182}, Trp^{184}, and His^{185} were substituted by Ala (See *SI Appendix*, Table S1 for residue numbering). Mutant proteins were compared with parental (wild-type) S1^A with maximum binding of BCoV S1^A–Fc set at 100%. sp-LBA was as in Fig. 1*C*, but with data points representing mean averages of independent duplicate experiments. (*C*) Residual binding of PHEV and OC43 S1^A–Fc mutants with Ala substitutions of site A residues as detected by HAA, performed as in Fig. 1*B*. (*D*) Substitutions of OC43 site A residues by BCoV orthologs do not increase RBS binding affinity as tested by sp-LBA and (*E*) by HAA.

that result in gain-of-function of OC43 S1^A (i.e., mutations that would raise its binding to that of BCoV S1^A). To this end, we systematically replaced residues within or proximal to the proposed binding site in OC43 S1^A by their BCoV orthologs (Fig. 2*A*; see also *SI Appendix*, Fig. S1). Individual substitutions Arg^{143}His, Lys^{181}Val, Leu^{186}Trp, and Ile^{145}Thr (Fig. 2*D*), however, did not increase binding affinity, and neither did the replacement of residues 147–151 (Ser-Thr-Gln-Asp-Gly) by Leu, which resulted not only in a large deletion but also in the removal of an *N*-glycosylation site. In fact, an OC43 S1^A mutant, in which we combined all substitutions and thus essentially reconstructed the proposed BCoV S1^A RBS in the OC43 background, did not differ in its binding affinity from the parental wild-type OC43 protein as measured by HAA and sp-LBA (Fig. 2*D*). From the combined results, we conclude that the actual RBS is located elsewhere. In further support, the RBS proposed by Peng et al. (17)—henceforth referred to as "site A"—does not conform to the typical anatomy of *O*-Ac-Sia binding sites (18–24).

Crystallization Attempts to Identify the S1^A RBS. The most obvious and direct approach to identify the RBS would be by crystallographic analysis of an S1^A–ligand complex. In their report, Peng et al. (17) stated that all efforts to determine the S1^A holo-structure had been unsuccessful. We considered that the complications the authors encountered might be related to crystal packing, and therefore opted to perform crystallization trials not with BCoV or OC43 S1^A, but with that of PHEV. Because its sequence varies from that of BCoV S1^A by 22%, we hoped that crystals of different packing topology would be produced, as was indeed the case. Crystals formed under a variety of conditions, but all with space group $P3_121$ with two S1^A molecules per asymmetric unit. These crystals consistently disintegrated within seconds during soaking attempts with receptor analog methyl-5-*N*-acetyl-4,9-di-*O*-acetyl-α-neuraminoside. Crystals flash-frozen immediately upon ligand addition showed poor diffraction that did not extend beyond 10-Å resolution. The observations are suggestive of ligand-binding interfering with crystal stability (*SI Appendix*, Fig. S2).

As an alternative, we performed extensive cocrystallization screenings with PHEV S1^A, both with native and EndoH$_f$-treated protein samples, together with methyl-5,9-di-*N*-acetyl-α-neuraminoside (Neu5,9Ac$_2$α2Me), a receptor analog chemically more stable than the 9-*O*–acetylated compound (22, 34). These attempts also remained unsuccessful as no crystals were formed. In a final effort to obtain crystals of altered packing that perchance would be compatible with ligand soaking, each of the four *N*-glycosylation sites in PHEV S1^A were systematically removed, separately and in combination, by Asn-to-Gln substitutions. However, the resulting proteins were prone to aggregation and hence unsuitable for crystallization.

The diffraction data obtained for noncomplexed PHEV S1^A eventually allowed us to solve the apo-structure to 3.0-Å resolution (PDB ID 6QFY; *SI Appendix*, Fig. S1). While the results do not provide direct clues to the location of the β1CoV S RBS, they do permit a side-by-side comparison of the S1^A domains of two divergent β1CoVs. The difficulties met by us and others to identify the RBS by crystallography led us to switch strategy and to follow an alternative approach based on comparative structural analysis and visual inspection of S1^A domains using the general design of HE 9-*O*-Ac-Sia binding sites as a query.

An Alternative RBS Candidate Identified by Comparative Structural Analysis. As explicated by Neu et al. (35), structural comparison of viral attachment proteins in complex with their sialoglycan-based ligands may aid to define common parameters of recognition, and in turn these "rules of engagement" may be used to predict the location of Sia-binding sites in viral proteins for which such structural information is still lacking. Studies by us and others on corona-, toro-, and orthomyxoviral HE(F) proteins

provide a wealth of structural data on how proteins recognize *O*-Ac-Sias (18–24). The *O*-acetyl Sia binding sites as they occur in HE lectin domains and esterase catalytic sites—despite major differences in structure and composition—conform to a common design in which ligand/substrate recognition is essentially based on shape complementarity and hydrophobic interactions, supported by protein–sugar hydrogen bonding involving characteristic Sia functions, such as the glycerol side chain, the 5-*N*-acyl moiety and, very often, the carboxylate. Typically, the critical *O*-acetyl moiety docks into a deep hydrophobic pocket (P1) and the 5-*N*-acyl in an adjacent hydrophobic pocket/depression (P2). P1 and P2, ∼7 Å apart in 9-*O*-Ac-Sia RBSs, are separated by an aromatic side chain, placed such that in the bound state the side chain intercalates between the *O*- and *N*-acyl groups, often with the ring structure positioned to allow for CH/π interactions with the *O*-acetyl-methyl moiety (18–20, 22–24). Site A previously proposed as the BCoV $S1^A$ RBS (17) clearly does not conform to this signature. Visual inspection of the PHEV and BCoV $S1^A$ structures, however, identified a region distal from site A that in many aspects does resemble a typical *O*-Ac-Sia binding site, with two hydrophobic pockets separated by the Trp^{90} indole (Fig. 3*A*). We will refer to this location as site B. Interestingly, in the BCoV $S1^A$ apo-structure, at the rim of site B, a sulfate ion is bound through hydrogen bonding with Lys^{81} and Thr^{83}. Its presence is of considerable significance as oxoanions in apo-structures often are indicative of and informative on interactions between the RBS and ligand-associated carboxylate moieties (36–38). Thus, in apo-structures of other Sia-binding proteins, sulfate and phosphate anions were found to mimic the Sia carboxylate in topology and sugar–protein hydrogen bonding (37, 38). Automated docking of 9-*O*, 5-*N*-Ac-Sia with the Sia carboxylate anchored at the position of the sulfate ion showed that the ligand would fit into site B, with the 9-*O*-acetyl group, most critical for ligand recognition (15, 39), docking into the more narrow and deeper pocket of the two (Fig. 3*B*). The 5-*N*-acyl moiety would be accommodated by a hydrophobic patch within the adjacent pocket, which is wide enough to also accept a 5-*N*-glycolyl group.

To explain the architecture of site B, *SI Appendix*, Fig. S1 presents the overall structure of BCoV $S1^A$ and the designation of secondary structure elements. Central to $S1^A$ is a β-sandwich scaffold, with one face packing against a separate domain containing the C-terminal region and the other covered by loop excursions that form a topologically distinct layer. The putative binding site is located within this layer at an open end of the β-sandwich with at its heart the $\beta5\text{-}3_{10}1$ region (residues 80–95). Leu^{80}, Leu^{86}, Trp^{90}, Phe^{91}, and Phe^{95} form a well-packed hydrophobic core that on one side interacts with the underlying β-sandwich. On the other side, along the protein surface, it is wrapped by residues from the N-terminal loop 1–β1-loop 2 segment (L1β1L2) to form site B pocket P1. L1β1L2 is locked in position through disulfide bonding of Cys^{21} and β12-residue Cys^{165}, and through extensive intersegment hydrogen bonding with $\beta5\text{-}3_{10}1$ residues, involving Ser^{25} and Asn^{27} via main-chain interactions with Leu^{86} and Arg^{88}, Val^{29} via main-chain–side-chain interaction with Ser^{87}, and Thr^{31} via side-chain–side-chain interaction with Trp^{90}.

Site B is conserved in all three β1CoVs and, from S cryo-EM structures obtained for other lineage A betacoronaviruses (26, 27), should be readily accessible also in the context of the fully folded intact spike (*SI Appendix*, Fig. S3). In the PHEV and BCoV $S1^A$ crystals (17), however, site B—but notably not site A—is occluded by packing contacts, in the case of BCoV $S1^A$ even via coincidental intermonomeric site B–site B interactions (*SI Appendix*, Figs. S2 and S4). Thus, the rapid disintegration of PHEV $S1^A$ crystals during soaking may well be explained by disruption of crystal contacts in result of ligand-binding.

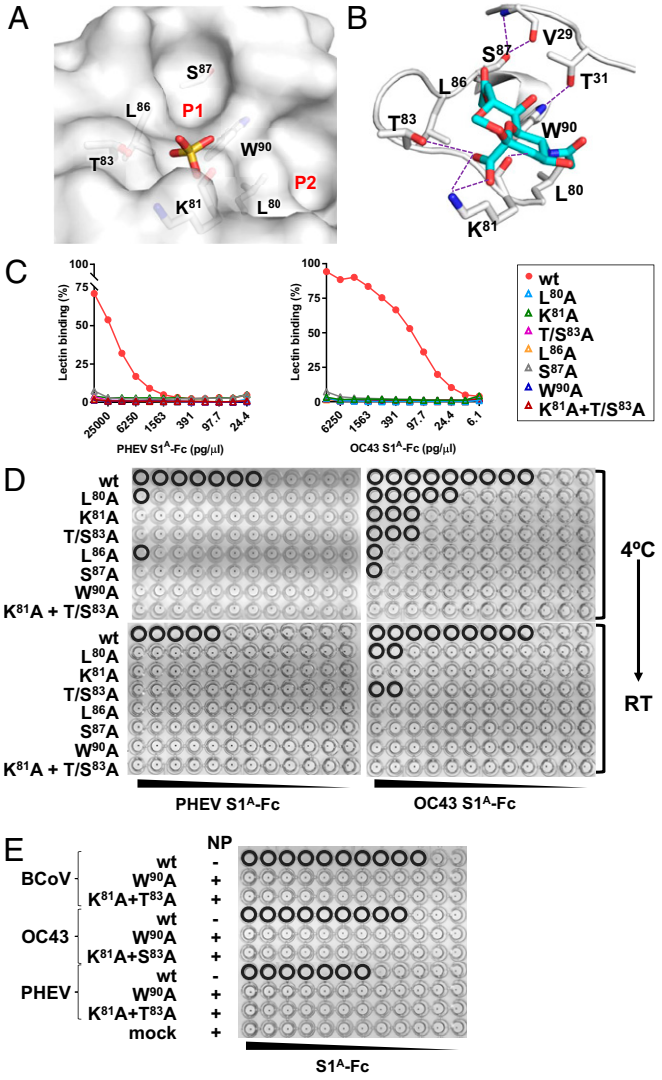

Fig. 3. Evidence from comparative structural analysis and structure-guided mutagenesis to suggest that the $S1^A$ RBS locates at site B. (*A*) Close-up of BCoV $S1^A$ site B in surface representation with hydrophobic pockets P1 and P2 indicated, and with side chains of residues, proposedly involved in receptor-binding, shown in sticks and colored by element. A sulfate ion bound to site B in the BCoV $S1^A$ apo structure (PDB ID code 4H14) is also shown (oxygen, red; nitrogen, blue; carbons, gray; sulfur, yellow). Hydrogen bonds between the sulfate ion and site B residues Lys^{81} and Thr^{83} are shown as purple dashed lines. (*B*) Cartoon representation of BCoV $S1^A$ in complex with 9-*O*-Ac-Sia as modeled by automated molecular docking with Autodock4 (22). Residues predicted to be involved in ligand-binding are shown in stick representation and colored as in *A*. Sia-9-*O*-Ac is also shown in sticks, but with carbon atoms colored cyan. Predicted hydrogen bonds between Lys^{81} and Thr^{83} side chains and the Sia carboxylate moiety, and between the Lys^{81} main chain and the Sia-5-*N*-Ac moiety are shown as purple dashed lines. (*C*) Substitutions of site B residues in PHEV and OC43 $S1^A$–Fc result in complete loss of binding as detected by sp-LBA, performed as in Fig. 2*B*. (*D*, *Upper*) Substitutions of site B residues in PHEV and OC43 $S1^A$–Fc result in substantial to complete loss of binding as measured by conventional HAA at 4 °C. (*Lower*) Mutations, resulting in residual binding, render the $S1^A$ RBS thermolabile. HAA shown (*Upper*) after a temperature shift-up to 21 °C and continued 16-h incubation. (*E*) Substitution of proposed ligand contacting residues Trp^{90}, Lys^{81}, and Thr^{83}/Ser^{83} results in total loss of detectable binding of BCoV, OC43 and PHEV $S1^A$–Fc even when assayed by high-sensitivity nanoparticle HAA. Assays performed with $S1^A$–Fc protein multivalently presented on 60-meric protein A-lumazine synthase icosahedral shells (pA-LS) as in ref. 40.

Mutational Analysis of β1CoV S1^A Site B. To test whether site B represents the true RBS, we again performed structure-guided mutagenesis of PHEV and OC43 S1^A (Fig. 3). Of note, the mutant proteins tested were expressed and secreted to wild-type levels, arguing against gross folding defects (for a quantitative comparison of the expression levels of key mutant proteins and their melting curves, see *SI Appendix*, Fig. S5). From the docking model, Trp^{90}, Lys^{81}, and Thr^{83} (Ser^{83} in OC43 S1^A) are predicted to be critically involved in ligand binding. In accordance, their replacement by Ala led to complete loss of detectable binding by sp-LBA (Fig. 3*C*). Upon substitution of Trp^{90} by Ala, binding was no longer observed for either PHEV or for OC43 S1^A, not even by HAA, be it conventional or high-sensitivity nanoparticle (NP-)HAA (40) (Fig. 3 *D* and *E*). Substitution of either Lys^{81} or Thr^{83} in PHEV S1^A led to loss of detectable binding also as measured by conventional HAA at 4 °C, but for OC43 S1^A-Ser^{83}Ala and OC43 S1^A-Lys^{81}Ala residual binding was detected. Saliently, HA by OC43 S1^A-Lys^{81}Ala fully resolved after a shift-up to room temperature, indicating that the mutation renders the receptor–ligand interaction thermolabile and that binding affinity is more severely affected under physiological conditions (Fig. 3*D*). Importantly, combined substitution of Lys^{81} and Ser^{83} in OC43 S1^A resulted in complete loss of detectable binding even by NP-HAA (Fig. 3*E*). Also, Ser^{87}Ala substitution, which disrupts the hydrogen bond with Val^{29} and thereby weakens the association between L1β1L2 and β5-3$_{10}$1, caused loss of binding as detected by sp-LBA, and a considerable decrease in binding as measured by HAA. Similar results were obtained upon substitution of Leu^{86} and Leu^{80}, which line the proposed P1 and P2 pockets, respectively (Fig. 3). Finally, replacement of L1β1L2 residue Thr^{31}, thus abolishing the hydrogen bond with the β5-3$_{10}$1 residue Trp^{90}, reduced binding affinity in sp-LBA by more than a 1,000-fold, as demonstrated for OC43 S1^A (*SI Appendix*, Fig. S6).

To further test our model, we asked whether we might also identify gain-of-function mutations (i.e., substitutions within or proximal to site B that would increase S1^A binding affinity). Among the three β1CoVs, the proposed RBS itself is highly conserved. OC43 and BCoV S1^A, for example, differ at site B only at position 83, which is either a Thr (in BCoV) or Ser (in OC43). Despite this near identity, BCoV S1^A is the stronger binder of the two as indicated by a consistent 32-fold difference in binding efficiency as measured by sp-LBA (Fig. 1*C*). The difference at position 83 is particularly intriguing, as this residue is predicted to be involved in ligand binding through hydrogen bond formation with the Sia carboxylate (Fig. 3*B*). Indeed, Ser^{83}Thr substitution in OC43 S1^A resulted in a profound increase in binding affinity almost to that of BCoV S1^A (Fig. 4*A*).

Comparative sequence analysis of BCoV and PHEV S1^A revealed several amino acid differences proximal to site B—among which are Thr^{88}Arg, Thr^{22}Asn, and Val^{24}Ser—that would alter hydrogen bonding within L1 and between L1 and β5-3$_{10}$1 (Fig. 4 and *SI Appendix*, Fig. S1). Moreover, the latter two together create an additional *N*-glycosylation site in PHEV S1^A. To test whether and how these differences affect receptor binding, PHEV residues were replaced by their BCoV orthologs (Fig. 4*B* and *C*). The mutations again led to gain-of-function. Thr^{88}Arg substitution, disrupting a Thr^{88}-Asp^{28} hydrogen bond absent in BCoV S1^A, resulted in a consistent twofold increase in binding affinity as measured by sp-LBA. Asn^{22}Thr substitution, eliminating the glycosylation site in PHEV S1^A, raised binding affinity even 50-fold. However, Ser^{24}Val substitution, also destroying the glycosylation site, did not have an effect in isolation. Yet, Ser^{24}Val and Asn^{22}Thr in combination raised PHEV S1^A-binding affinity by a further 30-fold almost to that of BCoV (Fig. 4*B*). Apparently, the difference in S1^A-binding affinity between PHEV and BCoV can be ascribed to the architecture of site B as determined by the L1–β5-3$_{10}$1 hydrogen bonding network rather than to the absence or presence of a glycan at position 22 (Fig. 4*B*). Changes in S1^A L1–β5-3$_{10}$1 association may alter pocket P1 and thereby affect ligand binding (*SI Appendix*, Fig. S7; note the difference in the topology of Leu^{86} and Trp^{90} side chains in PHEV and BCoV S1^A).

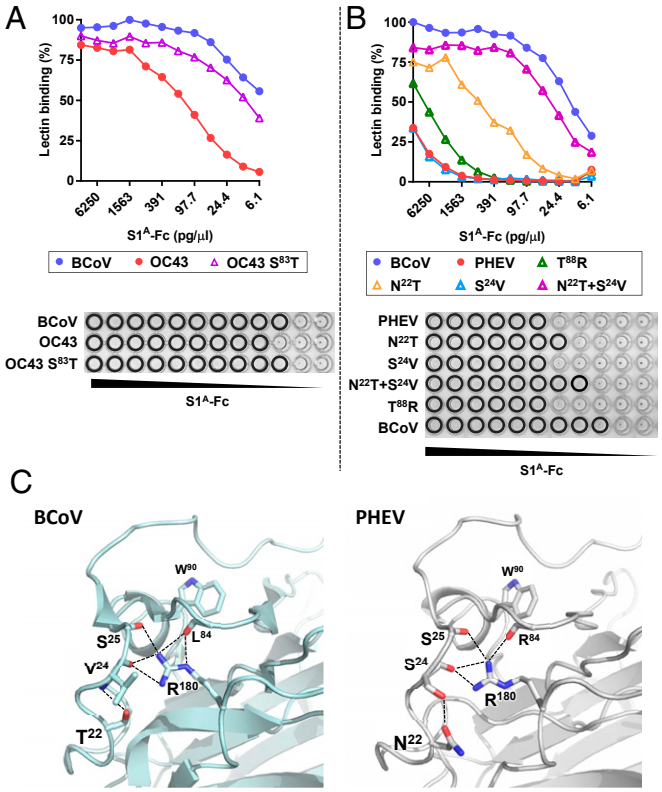

Fig. 4. Increased RBS binding affinity of OC43 and PHEV S1^A–Fc upon replacement of site B residues by BCoV orthologs. (*A*) Comparison of binding affinity of OC43 S1^A-Ser^{83}Thr to that of parental OC43 S1^A and BCoV S1^A by sp-LBA and conventional HAA. (*B*) Comparison of binding affinity of PHEV S1^A mutants (T^{88}R, N^{22}T, S^{24}V individually, and N^{22}T and S^{24}V combined) to that of parental PHEV S1^A and BCoV S1^A by sp-LBA and conventional HAA. Assays performed and results shown as in Fig. 1. (*C*) Differences in PHEV and BCoV S1^A binding affinity correlate with amino acid variations in loop L1 and the resulting changes in an intricate site B-organizing hydrogen bonding network. Side view of S1^A site B in BCoV (*Left*) and PHEV (*Right*) in combined cartoon and sticks representation to highlight differences in intra- and intersegment hydrogen bonds that fix the central site B β5-3$_{10}$1 segment through interactions with loop L1 and with the 3$_{10}$2-β12′ loop (*SI Appendix*, Fig. S1). Note the pivotal role of the 3$_{10}$2-β12′ loop Arg^{180} side chain bridging the L1 and β5β6 loops through multiple interactions with main chain carbonyls (depicted in sticks), and the difference between BCoV and PHEV S1^A in intraloop L1 hydrogen bonding between residues 24 and 26.

One remaining question is how Tyr^{162}, Glu^{182}, Trp^{184}, and His^{185} in site A affect the RBS and why their substitution reduces ligand-binding affinity, albeit modestly compared with mutations within site B. Within the structure the sites are spaced relatively closely together and indirect effects of site A mutations can be envisaged. One possible explanation is the location of strand β12′, comprising Glu^{182} and Trp^{184} right next to the first residues of L1. As we show, mutation of L1 residues Asn^{22}, Val^{24}, Thr^{31}, or of residues that interact with L1 like Ser^{87} have a drastic effect on ligand binding. Through a similar mechanism, mutations that destabilize strand β12′ may in turn affect L1 and ligand binding at site B.

In summary, our analyses revealed both loss-of-function and gain-of-function mutations within or in close proximity of site B, each of which in accord with our model for S1^A binding of 9-*O*-Ac-Sia in β1CoVs. Moreover, they provide a plausible explanation for

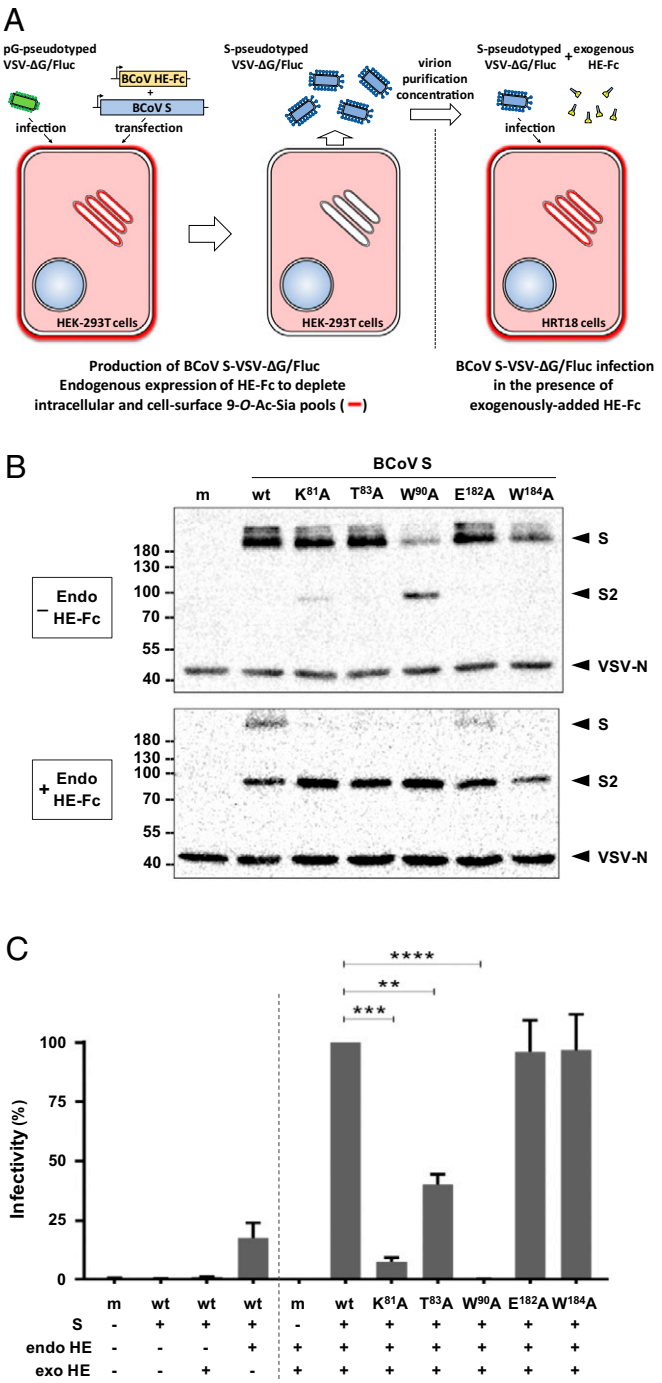

Fig. 5. Mutations in site B but not site A reduce infectivity of BCoV S-pseudotyped VSV particles. (*A*) Schematic outline of the production of BCoV S-pseudotyped VSV-ΔG/Fluc and of the infection experiments. (*B*) Protein composition of (mock) pseudotyped VSV particles. Western blot analysis of purified and concentrated virus particles, secreted by mock-transfected, VSVΔG/Fluc-transduced cells (m) or by transduced cells, transfected to transiently express Flag-tagged BCoV strain Mebus S (wild-type) or mutants thereof. BCoV S was either expressed exclusively (– Endo HE–Fc; *Upper*) or coexpressed with HE-Fc (+ Endo HE–Fc; *Lower*). Arrowheads indicate noncleaved S (S), the C-terminal subunit S2 resulting from intracellular proteolytic cleavage of S (S2), and the N protein of VSV. Note that coexpression of HE–Fc promotes cleavage of BCoV S, presumably by preventing inadvertent receptor association during S biosynthesis. Also note that in the absence of endogenous HE–Fc, site B mutants are cleaved to extents inversely related to RBS affinity. (*C*) Infectivity assays with BCoV S-pseudotyped viruses. Virus particles, pseudotyped with BCoV S or mutants thereof and produced in the presence or absence of intracellular HE–Fc (endo HE) were used to inoculate HRT18 cells, either with (+) or without (–) exogenous HE (exo HE) added to the inoculum. "Infectivity" is expressed in RLUs detected in lysates from inoculated cells at 18 h after inoculation. RLU values were normalized with those measured for VSV-S^{wt} set at 100%. The data shown are averages from three independent experiments, each of which performed with technical triplicates. SDs and significant differences, calculated by Welch's unequal variances *t* test, are indicated (**$P \leq 0.01$; ***$P \leq 0.001$, ****$P \leq 0.0001$).

the modest yet reproducible loss of binding affinity upon substitution of residues previously identified by Peng et al. (17). The results therefore lead us to conclude that site B identified here represents the true $S1^A$ RBS.

Site B Is Essential for Infectivity. To directly test the relevance of sites A and B for infectivity, we pseudotyped G protein-deficient, luciferase-expressing recombinant vesicular stomatitis virus (VSV) with Flag-tagged BCoV S or with mutant derivatives (see Fig. 5*A* for a schematic outline of the experiment). Considering that for β1CoVs HE is an essential protein (10, 16), we expressed wild-type BCoV S either exclusively or together with a secretory HE–Fc fusion protein. Because HE–Fc is not membrane-associated, it will not be included in the VSV envelope, but would deplete endogenous 9-*O*-Ac-Sia pools in the exocytotic compartment and from the cell surface. Thus, virus release would be facilitated and inadvertent receptor association during S biosynthesis and receptor-mediated virion aggregation would be prevented. Wild-type S in the absence of HE–Fc was incorporated in VSV particles in its noncleaved 180K form (Fig. 5*B*, *Upper*). However, the resulting particles—purified and concentrated from the tissue culture supernatant by sucrose cushion ultracentrifugation—were noninfectious as measured by luciferase assay (Fig. 5*C*). VSV particles, produced in cells that coexpressed S and HE–Fc, carried spikes that resembled those described for BCoV and OC43 virions in that a large proportion had been proteolytically cleaved into subunits S1 and S2 (Fig. 5*B*, *Lower*) (41, 42). Moreover, these particles were infectious. Values of relative light units (RLUs) detected in inoculated cells were more than 40-fold above background. Strikingly, their infectivity was increased by a further 5.7-fold upon addition of "exogenous" HE–Fc to the inoculum (Fig. 5 *A* and *C*). Hence, these conditions were chosen to determine the effects of mutations in BCoV S. Site B mutations caused complete loss of infectivity ($Trp^{90}Ala$), or significantly reduced infectivity to extents correlating with the effect of each mutation on S binding affinity ($Trp^{90}Ala > Lys^{81}Ala > Thr^{83}Ala$) (Figs. 3 and 5*C*). Particles pseudotyped with site A mutants were as infectious as S^{wt}-VSV. The combined data reinforce our conclusion that site B is the RBS and provide direct evidence that this site is essential for BCoV infectivity.

HCoV-HKU1 also Binds to 9-*O*-Ac-Sia via $S1^A$ Site B. HKU1 is a lineage A betacoronavirus but separated from the β1CoVs by a considerable evolutionary distance. To illustrate, $S1^A$ of HKU1 is only 55–60% identical to those of BCoV and OC43. However, HKU1, like the β1CoVs, uses 9-*O*-Ac-sialoglycans for attachment and the binding to these receptors is essential for infection (14). According to recent publications, the RBS would not be in $S1^A$, but in a downstream domain (31, 32), the main argument being that HKU1 $S1^A$ does not detectably bind to glycoconjugates (30). Indeed, we also did not detect binding of HKU1 $S1^A$–Fc, at least not by standard sp-LBA and conventional HAA (Fig. 6 *A* and *E*). However, when HKU1 $S1^A$–Fc was tested by high-sensitivity NP-HAA with multivalent presentation of the HKU1 $S1^A$-Fc proteins, agglutination of rat erythrocytes was observed (Fig. 6*A*). Moreover, depletion of cell surface 9-*O*-Ac-Sia receptors by pretreatment of the erythrocytes with BCoV HE completely prevented HA (Fig. 6*A*). These findings conclusively demonstrate that, in fact, HKU1

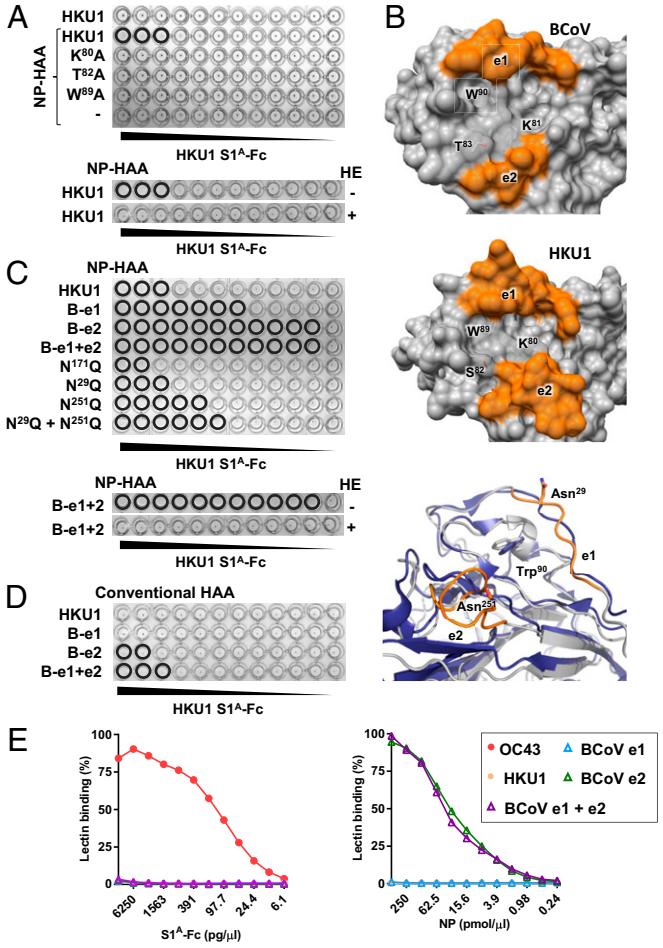

Fig. 6. HKU1 S1^A site B mediates binding to 9-*O*-Ac-Sias. (*A*) HKU1 S1^A–Fc binds to 9-*O*-Ac-Sia through low-affinity, high-specificity interactions. (*Upper*) Conventional HAA (top row) and high-sensitivity pA-LS NP-HAA (bottom rows) with native HKU1 S1^A-Fc (HKU1) or mutants thereof (Lys^{80}Ala, Thr^{82}Ala, Trp^{89}Ala). Noncomplexed nanoparticles were included as negative control (–). (*Lower*) NP-HAA with erythrocytes (mock)depleted for 9-*O*-Ac-Sias by prior sialate-*O*-acetylesterase treatment. (*B*) Surface representation of site B in BCoV (*Top*) and HKU1 (*Middle*) [PDB ID code 5I08 (27)], respectively. Protein structures were aligned with site B residues Lys^{81}, Thr^{83}, and Trp^{90} as query. Side chains of site B residues shown in sticks. BCoV residues 29–34 (element e1) and 246–253 (e2), and corresponding elements in HKU1 indicated in orange. (*Bottom*) Overlay of cartoon representations of BCoV (purple) and HKU1 (white) S1^A centered on site B. In the HKU1 protein, e1 and e2 are colored orange and marked. Side chains of Asn^{29} and Asn^{251}, acceptors of *N*-linked glycosylation, indicated in sticks. BCoV Trp^{90} and its HKU1 ortholog, also in sticks, shown as a reference point to site B. (*C*, *Upper*) Replacement of HKU1 S1^A e1 and e2 by corresponding BCoV elements increases binding affinity to 9-*O*-Ac-sialoglycans present on rat erythrocytes. NP-HAA with native HKU1 S1^A–Fc (HKU1) or with HKU1 S1^A–Fc derivatives, with e1 and e2 replaced by the corresponding BCoV elements separately (B-e1, B-e2) or in combination (B-e1+e2). For comparison, mutant HKU1 S1^A–Fc proteins were included with *N*-glycosylation sites eliminated in site A (N^{171}Q) or in elements e1 (N^{29}Q) and e2 (N^{251}Q), individually and in combination (N^{29}Q + N^{251}Q). (*C*, *Lower*) NP-HAA with erythrocytes (mock)depleted for 9-*O*-Ac-Sias. (*D*) Replacement of HKU1 S1^A e1 and e2 by corresponding BCoV elements increases binding affinity to 9-*O*-sialoglycans as to allow detection by conventional HAA. (*E*) Replacement of HKU1 S1^A e1 and e2 increases binding affinity as to allow detection of 9-*O*-Ac-sialoglycans by nanoparticle sp-LBA. (*Left*) Conventional sp-LBA with soluble S1^A-Fc. (*Right*) Nanoparticle sp-LBA with S1^A–Fc complexed to pA-LS icosahedral shells in twofold serial dilutions, starting at 50 nmol pA-LS per well. Data points are mean averages of independent duplicate experiments, each performed in triplicate.

S1^A does bind to 9-*O*-Ac-Sia in a 9-*O*-Ac–dependent fashion apparently through low-affinity high-specificity interactions. The detection of such interactions critically depends on multivalency of receptors and receptor-binding proteins (10, 40), and thus they are easily missed. This requirement for multivalency may also explain why preincubation of cells with soluble S1^A–Fc fusion proteins did not block HKU1 infection (32).

Interestingly, the RBS identified for β1CoV S1^A is conserved in HKU1 in sequence and structure (Fig. 6*B* and *SI Appendix*, Figs. S1*B* and S8), raising the question as to whether binding of HKU1 S to its sialoglycan receptor occurs through the same site. Indeed, Ala substitution of HKU1 S1^A Lys^{80}, Thr^{82} or Trp^{89}—orthologs of Lys^{81}, Thr^{83}, and Trp^{90} in BCoV S1^A—resulted in a complete loss of detectable binding (Fig. 6*A*; see *SI Appendix*, Fig. S9 for an analysis in the context of S1–Fc). To further study whether this site is involved in 9-*O*-Ac-Sia binding and to understand the structural basis for the considerable difference in affinity between HKU1 and β1CoV S1^A domains, we performed a side-by-side structure comparison. While in β1CoV S1^A domains the RBS is readily accessible, the corresponding site in HKU1 S1^A is located at the bottom of a canyon because it is flanked by protruding parallel ridges comprised of HKU1 S1^A residues 28 through 34 [element (e)1 corresponding to BCoV L2β2 residues 29–35], and 243 through 252 (e2 corresponding to BCoV residues 253–256 in the β18-β19 loop), each decorated with an *N*-linked glycan (Fig. 6*B*; see also *SI Appendix*, Fig. S1*B*) (27). Conceivably, this assembly would hamper binding particularly to short sialoglycans, such as the *O*-linked STn sugars (Sia-α2,6GalNAc-α1-*O*-Ser/Thr) predominantly present on BSM (43). Individual exchange of e1 and e2 in HKU1 S1^A by the corresponding segments from BCoV indeed increased receptor-binding as measured by NP-HAA by 16- or 256-fold for e1 and e2, respectively (Fig. 6*C*). Exchange of e2 in fact enhanced receptor affinity to levels that now also allowed detection of binding by conventional HAA, while simultaneous replacement of both elements increased binding affinity even further (Fig. 6*D*). Importantly, pretreatment of the erythrocytes with BCoV HE to deplete receptors by Sia-de-*O*-acetylation prevented hemagglutination (Fig. 6*C*), indicating that the enhanced binding was still 9-*O*-Ac-Sia–dependent. Replacement of e2 increased receptor affinity to such extent that also binding to BSM by sp-LBA assay became detectable, albeit only with nanoparticle-bound and not with free S1^A (Fig. 6*E*).

To study whether the *N*-glycans attached to e1 and e2 hinder receptor-binding and thus contribute to the apparent low affinity of HKU1 S1^A, we expressed the protein in *N*-acetylglucosaminyltransferase I-deficient HEK293S GnTI$^-$ cells (44). Replacement of complex *N*-glycans by high-mannose sugars resulted in an eightfold increase in binding affinity as measured by NP-HAA (*SI Appendix*, Fig. S9). In agreement, combined disruption of the *N*-linked glycosylation sites in e1 and e2 through Gln substituting for Asn^{29} and Asn^{251} increased binding affinity to a similar extent, with the latter mutation exerting the largest effect (Fig. 6*C*). The deletion of glycosylation sites, however, did not enhance affinity to that of the HKU1-BCoV S1^A/e1+e2 chimera, indicating that receptor binding is hampered not only by the *N*-linked glycans, but also by the local protein architecture. Of note, removal of the glycan attached to Asn^{171}, proximal to the binding site originally predicted by Peng et al. (17), did not enhance but actually lowered the binding affinity of HKU1 S1^A (Fig. 6*C*), possibly by affecting protein folding.

In summary, our results provide direct evidence that HKU1 binds to 9-*O*-Ac-Sias via S domain A. The data again argue against a role for site A (17), and on the basis of loss- and gain-of-function mutations strongly indicate that HKU1 binds its sialoglycan receptor via site B here identified for β1CoVs. On a critical note, we are aware that in the published HKU1 S apostructure, the orientation of the Lys^{81} side chain differs from that in BCoV S1^A such that hydrogen bonding with the Sia carboxylate

would be precluded. Whether this is an inaccuracy in the structure, whether the orientation of the side chain changes during ligand binding, or whether this is indeed a factual difference between HKU1 and β1CoV RBSs remains to be seen. Be it as it may, our findings do imply that in two human coronaviruses, related yet separated by a considerable evolutionary distance, the RBS and general mode of receptor-binding has been conserved. It is tempting to speculate that the apparently low affinity of HKU1 $S1^A$ is a virus-specific adaptation to replication in the human tract. Conceivably, occlusion of the HKU1 S RBS by e1 and e2 might allow selective high-affinity binding to particular sialoglycans with the 9-*O*-Ac-Sia terminally linked to extended glycan chains as to prevent nonproductive binding to short stubby sugars—such as present on BSM and other mucins—that otherwise would act as decoys and inhibitors. In analogy, influenza A H3N2 variants and 2009 pandemic H1N1 (Cal/04), initially assumed to carry low-affinity HAs, were recently reported to have evolved a preference for a subset of human-type α2,6-linked sialoglycan-based receptors comprising branched sugars with extended poly-*N*-acetyl-lactosamine (poly-LacNAc) chains (45). Whether such an adaptation would translate in distinctive differences between OC43 and HKU1 with respect to dynamic virion–receptor interactions and cell tropism merits further study.

Materials and Methods

Protein Design, Expression, and Purification. The 5′-terminal sequences of the S genes of BCoV strain Mebus (GenBank: P15777.1), OC43 strain ATCC VR-759 (GenBank: AAT84354.1), PHEV strain UU (GenBank: ASB17086.1), HKU1 strain Caen1 (GenBank: ADN03339.1), and mouse hepatitis virus (MHV) strain A59 (GenBank: P11224.2), encoding the signal peptide and adjacent $S1^A$ domain, were cloned in expression vector pCAGGS-Tx-Fc (46). Domain A, defined on the basis of the cryo-EM structure of the MHV-A59 S ectodomain (26), corresponds to S amino acid residues 15–294, 15–298, and 15–294 of BCoV, OC43, and PHEV, respectively. The amino acid sequences of these $S1^A$–Fc fusion proteins were deposited in GenBank (MG999832-35). Recombinant $S1^A$ proteins, genetically fused to the Fc domain of human IgG, were purified by protein A affinity chromatography from the supernatants of transiently transfected HEK293T cells as in Zeng et al. (18).

HAA. Standard HAA was performed as in Zeng et al. (18). Twofold serial dilutions of CoV $S1^A$–Fc proteins (starting at 2.5 μg per well; 50 μL per well) were prepared in V-shaped 96-well microtiter plates (Greiner Bio-One) unto which was added 50 μL of a rat erythrocyte suspension (*Rattus norvegicus* strain Wistar; 0.5% in PBS). High-sensitivity NP-HAA was performed as in Li et al. (40). In brief, $S1^A$–Fc proteins were complexed with pA-LS (a self-assembling 60-meric lumazine synthase nanoparticle, N-terminally extended with the Ig Fc-binding domain of the *Staphylococcus aureus* protein A) at a 0.6:1 molar ratio for 30 min on ice, after which complexed proteins were twofold serially diluted and mixed 1:1 with rat erythrocytes (0.5% in PBS). HA was assessed after 2-h incubation on ice unless stated otherwise. For specific depletion of cell-surface *O*-Ac-Sias, erythrocytes (50% in PBS pH 8.0) were (mock-)treated with soluble HE^+ (0.25 μg/μL BCoV-Mebus and 0.25 μg/μL PToV-P4) for 4 h at 37 °C comparable to as described in Langereis et al. (47).

Sp-LBA. The 96-well Maxisorp microtitre ELISA plates (Nunc) were coated with 0.1-μg BSM (Sigma-Aldrich) per well. Conventional sp-LBAs were performed using twofold serial dilutions of CoV $S1^A$–Fc proteins, as described previously (20, 47). In nanoparticle sp-LBA experiments, BSM-bound nanoparticles were detected with StrepMAb-HRP via the C-terminally appended Strep-tag of the pA-LS proteins, as described previously (40). Receptor-depletion assays were performed by (mock-)treatment of coated BSM with twofold serial dilutions of soluble HE^+ in PBS (100 μL per well) starting at 0.6 ng/μL (BCoV-Mebus) for 2 h at 37 °C, as in Langereis et al. 47. Receptor destruction was assessed by sp-LBA with β1CoV $S1^A$-Fc at fixed concentrations (BCoV 0.3 ng/μL; OC43 1.2 ng/μL).

Glycoside Synthesis and NMR Analysis. The starting material Neu5Ac2αMe (48) was treated with p-toluenesulfonyl chloride (1.8 eq.) in pyridine overnight and then with sodium azide (5 eq.) in DMF at 70 °C. The resulting intermediate was subjected to Staudinger reduction by 1 M trimethylphosphine in toluene (3 eq.) in the presence of potassium hydroxide. Upon completion, acetyl chloride (4 eq.) was added directly. After 5 min, 1 M potassium hydroxide solution was added to quench the reaction. The mixture was neutralized with acidic resin. The residue was purified with silica gel and then p-2 biogel to give pure Neu5,9$NAc_2$2αMe at a yield of 31% over four steps. The product was dissolved in 50 mM ammonium bicarbonate and freeze-dried, which was repeated three times to give the ammonium salt form. The final product was analyzed by NMR spectroscopy (*SI Appendix*, Fig. S10).

Crystallization. Crystallization conditions were screened by the sitting-drop vapor-diffusion method using a Gryphon (Art Robins). Drops were set up with 0.15 μL of $S1^A$ dissolved to 11 mg/mL in 10 mM BisTris-propane, 50 mM NaCl, pH 6.5, and 0.15-μL reservoir solution at room temperature. Diffracting crystals were obtained from the $JCSG^+$ screen [200 mM $MgCl_2$, 100 mM Bis·Tris pH 5.5, 25% (wt/vol) PEG3350] (JCSG Technologies) at 18 °C. Crystals were flash-frozen in liquid nitrogen using reservoir solution with 30% (wt/vol) glycerol as the cryoprotectant. Crystals were soaked with methyl-5,9-di-*N*-acetyl-α-neuraminoside and methyl-5-*N*-acetyl-4,9-di-*O*-acetyl-α-neuraminoside, as described, for the latter ligand, with batches previously used to solve HE holo-structures (18–20, 22).

Data Collection and Structure Solution. Diffraction data of PHEV $S1^A$ crystals were collected at European Synchrotron Radiation Facility station ID30A-3. Diffraction data were processed using XDS (49) and scaled using Aimless from the CCP4 suite (50). Molecular replacement was performed using PHASER (51) with BCoV $S1^A$ residues 35–410 as template (PDB ID code 14H4). NCS-restrained structure refinement was performed in REFMAC (52), alternated with model building in Coot (53); water molecules were built in difference density peaks of at least 5.0 σ located at buried sites conserved and occupied by water in the 1.5-Å resolution structure of BCoV $S1^A$. Molecular graphics were generated with PYMOL (pymol.sourceforge.net). See Table 1 for X-ray data statistics for PHEV $S1^A$.

Molecular Docking. Molecular docking of 9-*O*-Sia in the crystal structure of apo-BCoV $S1^A$ (PDB ID code 4H14) was performed with AutoDock4 (22)

Table 1. X-ray data statistics PHEV $S1^A$

Data collection and refinement	Statistic
Data collection	
Wavelength, Å	0.9677
Space group	$P3_121$
Cell dimensions	
a, b, c, Å	112.57, 112.57, 141.44
α, β, γ, °	90, 90, 120
Resolution range, Å*	97.5–3.0 (3.07–2.97)
No. unique reflections	20,445 (1,640)
Redundancy	7.3 (4.1)
Completeness, %	93.5 (79.9)
R_{merge}	0.154 (1.56)
I/σI	14.0 (1.0)
$CC_{1/2}$	0.994 (0.462)†
Refinement	
R_{work}/R_{free}	0.218/0.253
No. molecules in the asymmetric unit	2
No. atoms	
Protein	4,536
Carbohydrate	193
Water	30
Average B/Wilson B, $Å^2$	98.1/94.5
RMS deviations	
Bond lengths, Å	0.0078
Bond angles, °	1.33
NCS-restrained atoms	0.054
Ramachandran plot; favored, allowed, outliers, %	92.6, 6.9, 0.5

*Numbers between brackets refer to the outer resolution shell.
†Diffraction is highly anisotropic: $CC_{1/2}$ in the outer resolution shell is 0.845 for reflections within 20°from c*, whereas it is 0.00 within 20° from the a*b* plane.

similar to the procedure described by Bakkers et al. (22). The ligand molecule was extracted from BCoV HE (PDB ID code 3CL5). During docking, the protein was considered to be rigid. An inverted Gaussian function (50-Å half-width; 15-kJ energy at infinity) was used to restrain the Sia-carboxylate near the position occupied by the sulfate ion in the BCoV $S1^A$ crystal structure. The initial ligand conformation was randomly assigned, and 10 docking runs were performed.

Preparation of BCoV S-Pseudotyped VSV Particles and Infection Experiments. Recombinant G protein-deficient VSV particles were pseudotyped as described previously (54), with the S protein of BCoV strain Mebus. To allow transport of S to the cell surface and its incorporation into VSV particles, a truncated version of S was used from which the C-terminal 17 residues, comprising an endoplasmic reticulum-retention signal, had been removed. To facilitate detection, the S protein was provided with a Flag-tag by cloning its gene in expression vector pCAGGS-Flag. HEK 293T cells at 70% confluency were transfected with PEI-complexed plasmid DNA, as described previously (18). For coexpression of BCoV S and HE–Fc, S expression vectors and pCD5-BCoV HE–Fc (18) were mixed at molar ratios of 8:1. At 48 h after transfection, cells were transduced with VSV-G–pseudotyped VSVΔG/Fluc (54) at a multiplicity of infection of 1. Cell-free supernatants were harvested at 24 h after transduction, filtered through 0.45-μm membranes, and virus particles were purified and concentrated by sucrose cushion ultracentrifugation at approximately 100,000 × *g* for 3 h (10). Pelleted virions were resuspended in PBS and stored at −80 °C until further use. Relative virion yields were determined on the basis of VSV-N content by Western blot analysis by using anti–VSV-N monoclonal antibody 10G4 (Kerafast). Uptake of BCoV S into virus particles was detected by Western blot analysis with monoclonal antibody ANTI-FLAG M2 (Sigma). Inoculation of HRT18 monolayers in 96-well cluster format was performed with equal amounts of S-pseudotyped VSVs, as calculated from VSV-N content (roughly corresponding to the yield from 2×10^5 cells from each transfected and transduced culture), diluted in 10% FBS-supplemented DMEM. For virus infections with "exogenous" HE, 12.5 mU/mL of BCoV HE–Fc protein was added to the inoculum. At 18 h postinfection, cells were lysed using passive lysis buffer (Promega). Firefly luciferase expression was measured using a homemade firefly luciferase assay system, as described previously (55). Infection experiments were performed independently in triplicate, each time with three technical replicates.

ACKNOWLEDGMENTS. We thank Louis-Marie Bloyet for help in preparing the structure heat map conservation figure; Dr. Yoshiharu Matsuura (Osaka University, Japan) for providing VSV-G pseudotyped VSVΔG/Fluc; and the European Synchrotron Radiation Facility for providing beamline facilities and the beamline scientists at ID30A-3 for help with data collection. This study was supported by TOP Project Grant 40-00812-98-13066 funded by ZonMW (to F.J.M.v.K. and B.-J.B.); TOP-Punt Grant 718.015.003 of the Netherlands Organization for Scientific Research (to G.-J.B); China Scholarship Council Grant 2014-03250042 (to Y.L.); and by ECHO Grant 711.011.006 of the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (to R.J.d.G.).

1. Corman VM, et al. (2016) Link of a ubiquitous human coronavirus to dromedary camels. *Proc Natl Acad Sci USA* 113:9864–9869.
2. Vijgen L, et al. (2006) Evolutionary history of the closely related group 2 coronaviruses: Porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43. *J Virol* 80:7270–7274.
3. Vijgen L, et al. (2005) Complete genomic sequence of human coronavirus OC43: Molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J Virol* 79:1595–1604.
4. Huynh J, et al. (2012) Evidence supporting a zoonotic origin of human coronavirus strain NL63. *J Virol* 86:12816–12825.
5. Tao Y, et al. (2017) Surveillance of bat coronaviruses in Kenya identifies relatives of human coronaviruses NL63 and 229E and their recombination history. *J Virol* 91:1–16.
6. Yuen KY, Lau SK, Woo PC (2012) Wild animal surveillance for coronavirus HKU1 and potential variants of other coronaviruses. *Hong Kong Med J* 18:25–26.
7. Woo PC, et al. (2005) Clinical and molecular epidemiological features of coronavirus HKU1-associated community-acquired pneumonia. *J Infect Dis* 192:1898–1907.
8. McIntosh K, et al. (1970) Seroepidemiologic studies of coronavirus infection in adults and children. *Am J Epidemiol* 91:585–592.
9. Morfopoulou S, et al. (2016) Human coronavirus OC43 associated with fatal encephalitis. *N Engl J Med* 375:497–498.
10. Bakkers MJG, et al. (2017) Betacoronavirus adaptation to humans involved progressive loss of hemagglutinin-esterase lectin activity. *Cell Host Microbe* 21:356–366.
11. Hulswit RJG, de Haan CAM, Bosch B-J (2016) Coronavirus spike protein and tropism changes. *Adv Virus Res* 96:29–57.
12. Heald-Sargent T, Gallagher T (2012) Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. *Viruses* 4:557–580.
13. Matrosovich M, Herrler G, Klenk HD (2015) Sialic acid receptors of viruses. *Top Curr Chem* 367:1–28.
14. Huang X, et al. (2015) Human coronavirus HKU1 spike protein uses *O*-acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. *J Virol* 89:7202–7213.
15. Vlasak R, Luytjes W, Spaan W, Palese P (1988) Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc Natl Acad Sci USA* 85:4526–4529.
16. Desforges M, Desjardins J, Zhang C, Talbot PJ (2013) The acetyl-esterase activity of the hemagglutinin-esterase protein of human coronavirus OC43 strongly enhances the production of infectious virus. *J Virol* 87:3097–3107.
17. Peng G, et al. (2012) Crystal structure of bovine coronavirus spike protein lectin domain. *J Biol Chem* 287:41931–41938.
18. Zeng Q, Langereis MA, van Vliet ALW, Huizinga EG, de Groot RJ (2008) Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc Natl Acad Sci USA* 105:9065–9069.
19. Langereis MA, et al. (2009) Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases. *Proc Natl Acad Sci USA* 106:15897–15902.
20. Langereis MA, Zeng Q, Heesters BA, Huizinga EG, de Groot RJ (2012) The murine coronavirus hemagglutinin-esterase receptor-binding site: A major shift in ligand specificity through modest changes in architecture. *PLoS Pathog* 8:e1002492, and correction (2012) 8, 10.1371/annotation/a1e2a2e4-df03-40db-b10b-fd0cfcf78d3c.
21. Cook JD, Sultana A, Lee JE (2017) Structure of the infectious salmon anemia virus receptor complex illustrates a unique binding strategy for attachment. *Proc Natl Acad Sci USA* 114:E2929–E2936.
22. Bakkers MJG, et al. (2016) Coronavirus receptor switch explained from the stereochemistry of protein-carbohydrate interactions and a single mutation. *Proc Natl Acad Sci USA* 113:E3111–E3119.
23. Song H, et al. (2016) An open receptor-binding cavity of hemagglutinin-esterase-fusion glycoprotein from newly-identified influenza D virus: Basis for its broad cell tropism. *PLoS Pathog* 12:e1005411, and correction (2016) 12:e1005505.
24. Rosenthal PB, et al. (1998) Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* 396:92–96.
25. Walls AC, et al. (2016) Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol* 23:899–905.
26. Walls AC, et al. (2016) Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* 531:114–117.
27. Kirchdoerfer RN, et al. (2016) Pre-fusion structure of a human coronavirus spike protein. *Nature* 531:118–121.
28. Yuan Y, et al. (2017) Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* 8:15092.
29. Gui M, et al. (2017) Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* 27:119–129.
30. Peng G, et al. (2011) Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proc Natl Acad Sci USA* 108:10696–10701.
31. Ou X, et al. (2017) Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat Commun* 8:15216.
32. Qian Z, et al. (2015) Identification of the receptor-binding domain of the spike glycoprotein of human betacoronavirus HKU1. *J Virol* 89:8816–8827.
33. Li F (2016) Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol* 3:237–261.
34. Brossmer R, Isecke R, Herrler G (1993) A sialic acid analogue acting as a receptor determinant for binding but not for infection by influenza C virus. *FEBS Lett* 323:96–98.
35. Neu U, Bauer J, Stehle T (2011) Viruses and sialic acids: Rules of engagement. *Curr Opin Struct Biol* 21:610–618.
36. Kaiser JT, et al. (2000) Crystal structure of a NifS-like protein from Thermotoga maritima: Implications for iron sulphur cluster assembly. *J Mol Biol* 297:451–464.
37. Pronker MF, et al. (2016) Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nat Commun* 7:13584.
38. van der Vries E, et al. (2012) H1N1 2009 pandemic influenza virus: Resistance of the I223R neuraminidase mutant explained by kinetic and structural analysis. *PLoS Pathog* 8:e1002914.
39. de Groot RJ (2006) Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj J* 23:59–72.
40. Li W, et al. (2017) Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc Natl Acad Sci USA* 114: E8508–E8517.
41. King B, Brian DA (1982) Bovine coronavirus structural proteins. *J Virol* 42:700–707.
42. Hogue BG, King B, Brian DA (1984) Antigenic relationships among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. *J Virol* 51:384–388.
43. Kozak RP, Royle L, Gardner RA, Fernandes DL, Wuhrer M (2012) Suppression of peeling during the release of *O*-glycans by hydrazinolysis. *Anal Biochem* 423:119–128.
44. Reeves PJ, Callewaert N, Contreras R, Khorana HG (2002) Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc Natl Acad Sci USA* 99:13419–13424.
45. Peng W, et al. (2017) Recent H3N2 viruses have evolved specificity for extended, branched human-type receptors, conferring potential for increased avidity. *Cell Host Microbe* 21:23–34.
46. Li W, et al. (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426:450–454.

47. Langereis MA, et al. (2015) Complexity and diversity of the mammalian sialome revealed by nidovirus virolectins. *Cell Rep* 11:1966–1978.
48. Lubineau A, Le Gallic J (1991) Stereoselective syntheses of ALKYL- and ALKYL-2-THIO-α-sialosides. *J Carbohydr Chem* 10:263–268.
49. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
50. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
51. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40:658–674.
52. Vagin AA, et al. (2004) REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr Sect D Biol Crystallogr* 60:2184–2195.
53. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
54. Kaname Y, et al. (2016) Acquisition of complement resistance through incorporation of CD55/decay-accelerating factor into viral particles bearing baculovirus GP64. *J Virol* 84:3210–3219.
55. Burkard C, et al. (2014) Coronavirus cell entry occurs through the endo-/lysosomal pathway in a proteolysis-dependent manner. *PLoS Pathog* 10:e1004502.
56. Langereis MA, van Vliet ALW, Boot W, de Groot RJ (2010) Attachment of mouse hepatitis virus to *O*-acetylated sialic acid is mediated by hemagglutinin-esterase and not by the spike protein. *J Virol* 84:8970–8974.