



# Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity

Yifei Lang<sup>a</sup>, Wentao Li<sup>a</sup>, Zeshi Li<sup>b</sup>, Danielle Koerhuis<sup>a</sup>, Arthur C. S. van den Burg<sup>a</sup>, Erik Rozemuller<sup>c</sup>, Berend-Jan Bosch<sup>a</sup>, Frank J. M. van Kuppeveld<sup>a</sup>, Geert-Jan Boons<sup>b,d,e</sup>, Eric G. Huizinga<sup>f</sup>, Hilde M. van der Schaar<sup>c,1</sup>, and Raoul J. de Groot<sup>a,2</sup>

<sup>a</sup>Virology Division, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, The Netherlands; <sup>b</sup>Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CG Utrecht, The Netherlands; <sup>c</sup>GenDx B.V., 3584 CM Utrecht, The Netherlands; <sup>d</sup>Department of Chemistry, University of Georgia, Athens, GA 30602; <sup>e</sup>Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602; and <sup>f</sup>Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Faculty of Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands

Edited by Stanley Perlman, University of Iowa, Iowa City, IA, and accepted by Editorial Board Member Linda J. Saif August 12, 2020 (received for review April 12, 2020)

**Human coronaviruses OC43 and HKU1 are respiratory pathogens of zoonotic origin that have gained worldwide distribution. OC43 apparently emerged from a bovine coronavirus (BCoV) spillover. All three viruses attach to 9-O-acetylated sialoglycans via spike protein S with hemagglutinin-esterase (HE) acting as a receptor-destroying enzyme. In BCoV, an HE lectin domain promotes esterase activity toward clustered substrates. OC43 and HKU1, however, lost HE lectin function as an adaptation to humans. Replaying OC43 evolution, we knocked out BCoV HE lectin function and performed forced evolution-population dynamics analysis. Loss of HE receptor binding selected for second-site mutations in S, decreasing S binding affinity by orders of magnitude. Irreversible HE mutations led to cooperativity in virus swarms with low-affinity S minority variants sustaining propagation of high-affinity majority phenotypes. Salvageable HE mutations induced successive second-site substitutions in both S and HE. Apparently, S and HE are functionally interdependent and coevolve to optimize the balance between attachment and release. This mechanism of glycan-based receptor usage, entailing a concerted, fine-tuned activity of two envelope protein species, is unique among CoVs, but reminiscent of that of influenza A viruses. Apparently, general principles fundamental to virion-sialoglycan interactions prompted convergent evolution of two important groups of human and animal pathogens.**

coronavirus | hemagglutinin-esterase | spike | sialic acid | influenza virus

The subfamily *Orthocoronavirinae* comprises a group of enveloped positive-strand RNA viruses of clinical and veterinary significance. Adding to the socio-economic impact of coronaviruses (CoVs) already extant in humans and livestock, the emergence of “new” CoVs through cross-species transmission poses an ever-looming threat to public health, animal health, and food production.

Seven CoVs are known to infect humans, but not all of them have become established. The introduction of severe acute respiratory system (SARS) CoV in 2002 from horseshoe bats with masked palm civets as incidental transient hosts, was rapidly contained through quarantine measures (1). Middle East respiratory syndrome CoV, natural to dromedary camels, causes a classic zoonotic infection with limited human-to-human spread (2). December 2019, a member of the species SARS-CoV, called SARS-CoV-2 and 79.5% identical to the 2002 SARS-CoV variant, emerged in Wuhan, China (3, 4) to progress to full-scale pandemicity. Chances are SARS-CoV-2 will eventually become established in the human population.

Four other respiratory coronaviruses of zoonotic origin have succeeded in becoming true human viruses with world-wide

distribution (5–7). Among them are HKU1 and OC43 (subgenus *Embecovirus*, genus *Betacoronavirus*), related yet distinct viruses that arose from different zoonotic progenitors and entered the human population independently. OC43 is far more related to bovine coronavirus (BCoV), its presumptive ancestor, with early isolates sharing 97% genome identity (8, 9). Together with viruses of swine, canines, equines, and lagomorphs, OC43 and BCoV are considered host range variants of the virus species *Betacoronavirus-1* (collectively referred to as  $\beta$ 1CoVs throughout) (7). OC43 apparently emerged 70 to 130 y ago from a single cross-species transmission event that gave rise to a human-only virus (8–10). Like OC43, other  $\beta$ 1CoVs also exhibit host specificity (8, 11). While these observations attest to the host promiscuity and zoonotic potential of embecoviruses and  $\beta$ 1CoVs in particular, they are also indicative for the existence of host barriers, the breaching of which selects for adaptive mutations

## Significance

**Human coronavirus OC43 arose relatively recently, presumably from a bovine coronavirus spillover. Both viruses use 9-O-acetylated sialoglycans as receptors to which they attach via spike protein S. Another envelope protein, hemagglutinin-esterase (HE), serves as a receptor-destroying enzyme. We demonstrate that HE and S are functionally intertwined and that receptor destruction and receptor binding need to be carefully balanced for efficient (pre)attachment. During early emergence of OC43 this balance was reset, presumably as an adaptation to the human host. Such a two-protein mechanism for dynamic virion attachment is unique among coronaviruses, but reminiscent of that of influenza A viruses. Apparently, general principles fundamental to virion-sialoglycan interactions prompted convergent evolution of two zoonotically-relevant groups of pathogens.**

Author contributions: Y.L., W.L., and R.J.d.G. designed research; Y.L., W.L., D.K., A.C.S.v.d.B., E.R., and H.M.v.d.S. performed research; W.L., Z.L., and G.-J.B. contributed new reagents/analytic tools; Y.L., W.L., Z.L., D.K., A.C.S.v.d.B., E.R., B.-J.B., F.J.M.v.K., G.-J.B., E.G.H., H.M.v.d.S., and R.J.d.G. analyzed data; R.J.d.G. supervised and coordinated the study; and Y.L. and R.J.d.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. S.P. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

<sup>1</sup>Present address: Batavia Biosciences B.V., 2333 CL Leiden, The Netherlands.

<sup>2</sup>To whom correspondence may be addressed. Email: r.j.degroot@uu.nl.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2006299117/-DCSupplemental>.

First published September 29, 2020.

that result in host specialization and, ultimately, virus speciation. Conceivably, comparative studies of BCoV and OC43 may identify factors that promote or restrict cross-species transmission of CoVs and thus further our understanding of the requirements for colonization of the human host.

Embecoviruses, OC43 and BCoV included, differ from other CoVs in that they encode two types of surface projections. Homotrimeric “peplomers” comprised of spike protein S and extending 20 nm from the viral membrane, mediate virion attachment to entry receptors and membrane fusion (12). Interspersed are stubby 8-nm homodimeric projections comprised of the hemagglutinin-esterase (HE) (13–15), a dual-function protein typically encompassing a receptor-binding lectin domain specific for *O*-acetylated sialic acid (*O*-Ac-Sia) and a receptor destroying sialate-*O*-acetyl esterase domain (16–20). The HE lectin domain contributes to virion attachment, but at the same time enhances sialate-*O*-acetyl esterase activity toward clustered sialoglycotypes (11).

Some embecoviruses, like mouse hepatitis virus (MHV) and related CoVs in rodents, attach to 4- or 9-*O*-acetylated sialosides (4- or 9-*O*-Ac-Sias) via HE (21–25) and to a proteinaceous entry receptor via S (26, 27). Others, animal  $\beta$ 1CoVs included, bind to 9-*O*-Ac-Sias via HE (28) but, remarkably, also via S (29, 30) or, in the case of human CoVs OC43 and HKU1, even exclusively via S (11, 31, 32).

Structure function analyses of HE and S proteins have yielded a wealth of data on ligand binding, substrate selection, and protein-glycan interactions. The receptor binding sites (RBSs) of CoV HE lectin domains and those in related proteins of toro- and influenza C/D viruses (22, 33–35) differ in sequence and structure yet conform to a common architectural design with a deep hydrophobic pocket (P1) to accommodate the crucial sialate-*O*-acetyl moiety, and an adjacent pocket or depression (P2) to accept the 5-*N*-acyl group (17, 21, 22, 33). Characteristically, P1 and P2 are separated by an aromatic side chain and binding of the ligand is stabilized further through electrostatic protein-glycan interactions typically involving distinctive Sia functions, such as the Sia glycerol side chain, the 5-*N*-Acyl, and the Sia carboxylate. The RBS for 9-*O*-Ac-Sia in the S proteins of BCoV and OC43, identified by comparative structural analysis (32) and confirmed by the cryoelectron microscopy (cryo-EM) holostructure of OC43 S (36), conforms to this blueprint. Moreover, this site, located in the N-terminal subdomain S1<sup>A</sup>, is structurally and functionally conserved in HKU1 (32).

Much less is known about the functional relationship between S and HE, and the role of HE in particular remains poorly understood. In MHV, HE expression is dispensable for replication and rapidly lost during cell culture propagation (15). Conversely, in  $\beta$ 1CoVs, HE is critical for infection. In OC43, loss of HE-associated acetyl esterase activity abrogates the production of infectious virus and virus dissemination in cell culture (37). Moreover, acetyl-esterase inhibitors block BCoV replication (19), and antibodies against HE neutralize the virus in vitro and in vivo (38–40). Still, even among  $\beta$ 1CoVs there are differences in HE function, apparently correlating with host specificity. Whereas HE lectin activity is strictly maintained in BCoV (28), OC43 lost this function through progressive accumulation of mutations in the HE RBS, apparently as an adaptation to replication in the human respiratory tract (11). Nevertheless, isolates of either virus propagate in cultured cells. To better understand the consequences of loss of HE lectin function as it occurred during OC43 and also HKU1 evolution, we took a reverse genetics/forced evolution approach with BCoV as a model. The findings reveal that HE and S are functionally interdependent and that the acquisition of HE by a proto-embecovirus allowed its  $\beta$ 1CoV descendants to adopt strategies for reversible virion-sialoglycan attachment, remarkably similar to those of influenza A viruses.

## Results

**Disruption of HE Lectin Function Selects for Mutations in S.** The receptor-destroying enzyme (RDE) activity of HE facilitates virus release at the end of the infectious cycle but is also important during the very early stages of infection (32, 37). Virions attach to 9-*O*-Ac-Sia-based receptors via S with fast association and dissociation rates (36). The combination of dynamic S-mediated receptor binding and HE RDE activity results in local receptor depletion, thereby allowing virions to escape from irreversible nonproductive binding to the highly densely clustered decoy receptors in the mucus and glycocalyx. In analogy with influenza A and C viruses (41–44), the repetitive action of binding and catalytic release may translate in virion motility allowing particles to traverse the mucus layer, to pass the glycocalyx, and to reach the entry receptors at the cell surface. We recently demonstrated the importance of HE for infection during (pre)attachment in infection experiments with G-deficient vesicular stomatitis (VSV) particles that were pseudotyped with wild-type BCoV S (pVSV-S). Human rectal tumor 18 (HRT18) cells are abundantly covered with 9-*O*-Ac-Sias (28) and susceptible to BCoV and OC43. They can also be infected by pVSV-S but the efficiency of infection increases by more than fivefold when enzymatically active, soluble recombinant HE is added to the inoculum (32). Apparently, also in HRT18 cells, the sialoglycans in the glycocalyx act as decoys and their depletion by HE enhances infection.

To study the role and importance of HE during natural  $\beta$ 1CoV infection and that of the HE lectin domain especially, we developed a reverse genetics system for BCoV strain Mebus based on targeted RNA recombination (*SI Appendix, Fig. S1A*) (45, 46). Recombinant wild-type BCoVs (rBCoV) with parental type HE and S, but with accessory ORF4a replaced by the Renilla luciferase gene (rBCoV-Rluc), were readily generated upon seeding acceptor-virus-infected, donor RNA-transfected mouse LR7 cells onto monolayers of feeder HRT18 cells. rBCoV-Rluc arose and within 7 d grew to final titers routinely obtained for wild-type BCoV ( $\sim 10^8$  TCID<sub>50</sub>/mL).

Generating BCoV-Rluc derivatives defective in HE lectin function proved more cumbersome. To abolish the HE RBS, we substituted Phe<sup>211</sup>, which is crucial for ligand binding (17) (*SI Appendix, Fig. S1B*), by Ala. To prevent rapid reversion, this was done by introducing two nucleotide substitutions. Mutant viruses were recovered eventually, but in three of four successful trials, a multistep 160-h rescue did not suffice and an additional 72- to 96-h blind passage was required (Fig. 1A).

Sequence analysis of clonal virus populations, obtained by endpoint dilution, confirmed the presence of the HE Phe<sup>211</sup>Ala substitution in all cases. Surprisingly, the purified viruses all suffered single site mutations in S, clustering in domain S1<sup>A</sup> (amino acids 15 to 302) in proximity of the RBS (Fig. 1A–C and *SI Appendix, Fig. S2*). Two of the trials yielded multiple S variants and some variants (Thr<sup>83</sup>Ile and Leu<sup>89</sup>Pro) emerged independently in separate experiments (Fig. 1A). The mutations map to three distinct S RBS elements ( $\beta$ 1, 3<sub>10</sub>1, and  $\beta$ 5) that are essential for ligand binding (Fig. 1C and *SI Appendix, Fig. S2C*; nomenclature according to ref. 32). Ile<sup>26</sup>Ser and Asn<sup>27</sup>Ser locate in the S1<sup>A</sup>  $\beta$ 1 element within the N-terminal L1- $\beta$ 1-L2 segment (amino acids 15 to 33) that walls pocket P1 (32, 36). Moreover, in the OC43 S cryo-EM holostructure, the Asn<sup>27</sup> side chain hydrogen bonds with the 9-*O*-acetyl carbonyl (36). Leu<sup>89</sup>Pro in S1<sup>A</sup> element 3<sub>10</sub>1 is immediately adjacent to Trp<sup>90</sup>. The latter is arguably the most critical residue in the RBS as its indole side chain separates the P1 and P2 pockets, and its replacement precludes receptor binding and virus infectivity (32). Finally, Gly<sup>82</sup>Glu and Thr<sup>83</sup>Ile substitutions occurred in S1<sup>A</sup> element  $\beta$ 5 that interacts with the sialate carboxylate through hydrogen bonding with Lys<sup>81</sup> and Thr<sup>83</sup>.

As measured by solid-phase lectin binding assay (sp-LBA) with S1<sup>A</sup>-Fc fusion proteins and bovine submaxillary mucin

(BSM) as ligand, all mutations significantly reduced S binding to 9-*O*-Ac-Sia, albeit to widely different extents. S1<sup>A</sup>-Fc binding affinities of the mutants were 500-fold (Asn<sup>27</sup>Ser) to more than 10,000-fold (Ile<sup>26</sup>Ser; Gly<sup>82</sup>Glu) lower than that of parental BCoV S1<sup>A</sup>-Fc (Fig. 1D).

Destruction of the HE lectin RBS abolishes HE-mediated virion attachment to 9-*O*-Ac-Sia receptor determinants and causes a reduction in overall virion binding avidity. However, this is clearly not the main defect. The HE-Phe<sup>211</sup>Ala substitution selected for mutations that markedly decreased the affinity of the S RBS, thus reducing overall virion avidity even more. The data can be understood from our earlier observations that loss of HE lectin function also affects HE's receptor destroying sialate-*O*-acetyltransferase activity (11). The HE lectin domain serves as a regulator of the esterase domain, its function similar to that of the carbohydrate-binding modules that are appended to the enzyme domains of cellular glycoside hydrolases (47). It promotes receptor-destroying esterase activity toward multivalently presented glycotopes, such as present on mucins by keeping the esterase domain in prolonged close proximity of these clustered ligands. In the case of soluble HE (sHE), inactivation of HE lectin RBS decreases esterase activity toward mucin-associated sialoglycans by up to 500-fold. However, in the natural context of the intact virus particle, with spikes and HE molecules closely packed, the effects are more modest. Virion-associated esterase activity is diminished but still substantial because the lectin function of HE is partially taken over by S (11). Nevertheless, the present observations suggest that the destruction of the BCoV HE RBS creates an imbalance between S-mediated virion attachment and HE-mediated receptor destruction that can only be compensated by strongly reducing S affinity. Again, the observations fit a model in which dynamic binding of virions is crucial for infectivity and in which reversibility of attachments confers virion motility thereby enabling virions to dodge decoy receptors and to browse for entry receptors.

**HE Lectin-Deficient Recombinant BCOVs Are Genetically Stable When Grown in the Presence of Exogenous RDE.** To test whether loss of virion-associated RDE activity in rBCoV-HE-F<sup>211</sup>A/S<sup>wt</sup>/Rluc might be compensated for by adding exogenous soluble HE to the culture medium, we seeded infected/transfected LR7 cells onto HRT18 cell monolayers, supplemented the cell culture supernatant with BCoV HE-Fc (17) to final concentrations of 1 pg to 10 µg/mL, and allowed infection to proceed for 120 h. While in the absence of HE-Fc there was no sign of virus propagation as detectable by immunofluorescence assay, concentrations of exogenous sialate-*O*-acetyltransferase as low as 1 ng/mL to up to 1 µg/mL promoted virus growth (Fig. 2A).

To determine whether these conditions would allow isolation of rBCoV-HE-F<sup>211</sup>A without mutations in S1<sup>A</sup>, we performed targeted recombination and rescued recombinant viruses by 160-h multistep propagation as before, but now with culture supernatant supplemented with 100 ng/mL HE-Fc (Fig. 2B). Sanger sequence analysis of RT-PCR amplicons showed that all viruses, cloned by endpoint dilution of the 160-h stock ( $n = 4$ ), coded for mutant HE-Phe<sup>211</sup>Ala in combination with wild-type S1<sup>A</sup>. To assess the stability of clonal rBCoV-HE-F<sup>211</sup>A/S<sup>wt</sup>/Rluc, the virus stock was amplified by a low multiplicity of infection (MOI) passage for another 160 h in the presence of exogenous HE-Fc (Fig. 2B). The resulting virus population was analyzed by next-generation sequencing (NGS), which allows for the detection of low-frequency mutants. Sequence variation in HE and S1<sup>A</sup> was distributed randomly and did not exceed background levels (<0.15%). More than 99.5% of the viruses coded for HE-Phe<sup>211</sup>Ala, while preserving parental type S1<sup>A</sup> (Fig. 2B).

**Loss of HE Lectin Function Gives Rise to Mixed Virus Population with Competition and Cooperativity among S Affinity Variants.** With a clonal, virtually pure stock of rBCoV-HE-F<sup>211</sup>A/S<sup>wt</sup> available, we

performed controlled forced evolution experiments. The virus was serially passaged involving three consecutive 120-h multistep propagation rounds in HRT18 cells, but now in the absence of exogenous HE-Fc, with the initial infection performed at an MOI of 0.005 (Fig. 2C). In trial 1, viral titers in passage 1 (p1) increased only slowly to  $3 \times 10^4$  and  $2 \times 10^4$  TCID<sub>50</sub>/mL (measured with or without exogenous HE-Fc, respectively). The withdrawal of exogenous RDE during viral passage immediately selected for mutations in S1<sup>A</sup>. Virus cloning by endpoint dilution of the 120-h p1 sample yielded S RBS mutants Asn<sup>27</sup>Ser, Thr<sup>83</sup>Ile, and Ile<sup>26</sup>Ser (Fig. 2C), all three of which had been seen before (Fig. 1A). NGS analysis revealed the true complexity of the p1 population (Fig. 2C) and identified two additional S1<sup>A</sup> variants with substitutions (His<sup>173</sup>Tyr and Arg<sup>197</sup>Cys) more distal from the RBS (SI Appendix, Figs. S2D and S3 A and B). His<sup>173</sup>Tyr also reduced the relative binding affinity of S1<sup>A</sup>-Fc, albeit less dramatically than the other mutations, namely by 30-fold (Table 1). The Arg<sup>197</sup>Cys mutation seemingly falls in a separate category and presumably reduces the avidity of S homotrimers by inducing aberrant disulfide bonding and local S1<sup>A</sup> misfolding in one or more S monomers (SI Appendix, Fig. S3).

All in all, the p1 population was comprised for virtually 100% of HE-Phe<sup>211</sup>Ala mutants, 40% of which in combination with parental BCoV S, the remaining 60% with second-site mutations in S1<sup>A</sup> (Fig. 2C and Table 1). Of the latter, the ultralow-affinity variant S1<sup>A</sup>-Ile<sup>26</sup>Ser was the most abundant at 46% and the Asn<sup>27</sup>Ser variant the least at less than 1%. However, upon a subsequent round of 120-h multistep propagation, the tables were turned with S1<sup>A</sup>-Asn<sup>27</sup>Ser now comprising almost 40% of the p2 population and with the Ile<sup>26</sup>Ser variant reduced to 0.7%. In addition, four other S1<sup>A</sup> variants emerged. One of these had a mutation in S1<sup>A</sup> RBS loop L1, Val<sup>29</sup>Gly, and a relative binding affinity close to that of the Asn<sup>27</sup>Ser mutant (Fig. 2C and Table 1). We also identified at position 75 a second S1<sup>A</sup> Cys-substitution mutant, which like Arg<sup>197</sup>Cys, presumably disrupts the RBS through aberrant disulfide bonding (SI Appendix, Fig. S3). Remarkably, two other S1<sup>A</sup> variants arose with mutations (Arg<sup>88</sup>Thr and Pro<sup>174</sup>Leu) that affected the relative binding affinity only modestly to 0.25 and 0.5 of that of wild-type S1<sup>A</sup>-Fc, respectively (Table 1). Even more remarkably, upon further passage these mutants of near wild-type affinity increased to dominate the p3 population, effectively outcompeting variants with low-affinity spikes as well as those with parental spikes (Fig. 2C and D). However, when the p1, p2, and p3 stocks were cloned by endpoint dilution in the absence of exogenous HE-Fc, only virus variants with low-affinity mutations in S1<sup>A</sup> were isolated (Fig. 2C and Table 1). Strikingly, from the p3 stock, the S1<sup>A</sup>-Asn<sup>27</sup>Ser variant was isolated exclusively against all odds (10 of 10 tested;  $P < 10^{-6}$ ) when calculated from its frequency in the population (21%). Conversely, virus cloning by endpoint dilution in the presence of exogenous HE-Fc yielded high-affinity S1<sup>A</sup> mutants Pro<sup>174</sup>Leu (5 of 11) and Arg<sup>88</sup>Thr (4 of 11), parental virus rBCoV-HE-Phe<sup>211</sup>Ala/S<sup>wt</sup> (1 of 11), and intermediate S affinity variant His<sup>173</sup>Tyr (1 of 11).

Notably, the conditions selected not only for mutations in S but also in HE. Variants with an Ala<sup>211</sup>Val substitution in HE emerged in p2, rising to 17% of the p2 end population, to stabilize around this frequency in p3. As a result of this mutation, HE lectin affinity was regained albeit to levels solely detectable by high-sensitivity nanobead hemagglutination assay (HAA), while esterase activity toward clustered glycotopes in BSM increased 4-fold as compared to HE-Phe<sup>211</sup>Ala, but still remained 125-fold lower than that of wild-type HE (SI Appendix, Fig. S4). Apparently, the increase in HE function, minor as it may be, provides a selective advantage, but apparently one that benefits both low- and high-affinity S variants, because the mutation was found in cloned viruses of either type.

**Loss of HE Lectin Function Selects for Virus Swarms with Low-Affinity S Escape Mutants Promoting the Emergence and Propagation of High-Affinity S Variants.** To corroborate our observations, the controlled forced evolution experiment was repeated (*SI Appendix, Fig. S5*). As compared to the first trial, there was a much faster population built-up already in p1 at 120 h postinfection (hpi) with final titers reaching  $4 \times 10^8$  and  $3.4 \times 10^7$  TCID50/mL, when measured in the presence or absence of exogenous HE-Fc, respectively. Surprisingly, in stark contrast to trial 1, the trial 2 p1 population was comprised for about 94% of viruses expressing wild-type BCoV S. Less than 6% consisted of variants with mutations in S1<sup>A</sup>, four of low receptor binding affinity (Thr<sup>22</sup>Ile, Asn<sup>27</sup>Tyr, Val<sup>29</sup>Gly, His<sup>173</sup>Tyr), with the exception of Thr<sup>22</sup>Ile, all at positions seen before, and one of near wild-type binding affinity (Pro<sup>174</sup>Leu) (Tables 1 and 2). Consistent with our previous findings, however, virus purification through endpoint dilution in the absence of exogenous HE-Fc yielded low-affinity mutants (11 of 11 tested) exclusively (*SI Appendix, Fig. S5*). If the variants in the trial 2 p1 population were all of equal replicative fitness under the conditions applied, the odds of this result would be less than  $1.10^{-12}$ .

At first glance, the two trials would seem to differ in their outcomes. We offer, however, that the results are in fact consonant and that the main difference is in the speed with which the virus populations increased and evolved. There is an inherent stochastic element to the experimental approach and whether the developing quasi-species undergoes slow-track (Exp. 1) or fast-track evolution (Exp. 2) is likely dependent on the time of advent of the first mutant virus and its properties, for example whether it is an ultralow- (like Ile<sup>26</sup>Ser) or low-affinity variant (like Asn<sup>27</sup>Ser). The findings allow for several conclusions. 1) They confirm and firmly establish that loss of HE lectin function selects for mutations in S1<sup>A</sup> that reduce S receptor-binding affinity and virion avidity. 2) The possibilities to reduce the affinity of the S RBS through single site mutations are finite. In several independent experiments, substitutions in S1<sup>A</sup> occurred at a limited number of positions, albeit not necessarily by the same residue. For example, Asn<sup>27</sup> was replaced both by Ser and Tyr. 3) The mutations that reduce S affinity fall into different categories. Most map within or in close proximity of the RBS to affect receptor–ligand interaction directly. Others, like His<sup>173</sup>Tyr and Pro<sup>174</sup>Leu, are more distal from the RBS and apparently affect

**Table 1. Virus composition in passages p0 through p3 of controlled forced evolution Exp. 1**

S1 <sup>A</sup>	Frequency in population*				rAff <sup>†</sup>
	p0	p1	p2	p3	
Wild-type	100	40.90	13.90	3.00	1.0
Ile <sup>26</sup> Ser	0	45.83	0.70	0.03	0.0000625
Asn <sup>27</sup> Ser	0	0.94	39.80	21.09	0.004
Val <sup>29</sup> Gly	0	0	4.10	0.65	0.008
Tyr <sup>75</sup> Cys	0	0	1.60	0.02	ND
Thr <sup>83</sup> Ile	0	2.90	1.60	0.21	0.002
Arg <sup>88</sup> Thr	0	0	2.30	9.69	0.25
His <sup>173</sup> Tyr	0	5.77	5.00	8.10	0.03
Pro <sup>174</sup> Leu	0	0	13.00	46.72	0.5
Arg <sup>197</sup> Cys	0	2.24	17.25	10,0.27	0.25
Thr <sup>22</sup> Ile	NA	NA	NA	NA	0.015
Asn <sup>27</sup> Tyr	NA	NA	NA	NA	0.002

\*Percental occurrence of BCoV S1<sup>A</sup> mutations.

<sup>†</sup>Relative S1<sup>A</sup> binding affinities as measured by equilibrium endpoint solid-phase binding assay with S1<sup>A</sup>-Fc fusion proteins with that of parental BCoV S1<sup>A</sup>-Fc (“wild-type”) set at 1.0. S1<sup>A</sup> variants Thr<sup>22</sup>Ile and Asn<sup>27</sup>Tyr emerged only in Exp. 2 (*SI Appendix, Fig. S5*), but their affinities relative to that of wild-type S1<sup>A</sup> are shown for comparison. (NA, not applicable; ND, not determined).

**Table 2. Summary of S1<sup>A</sup> mutations identified in rBCoVs**

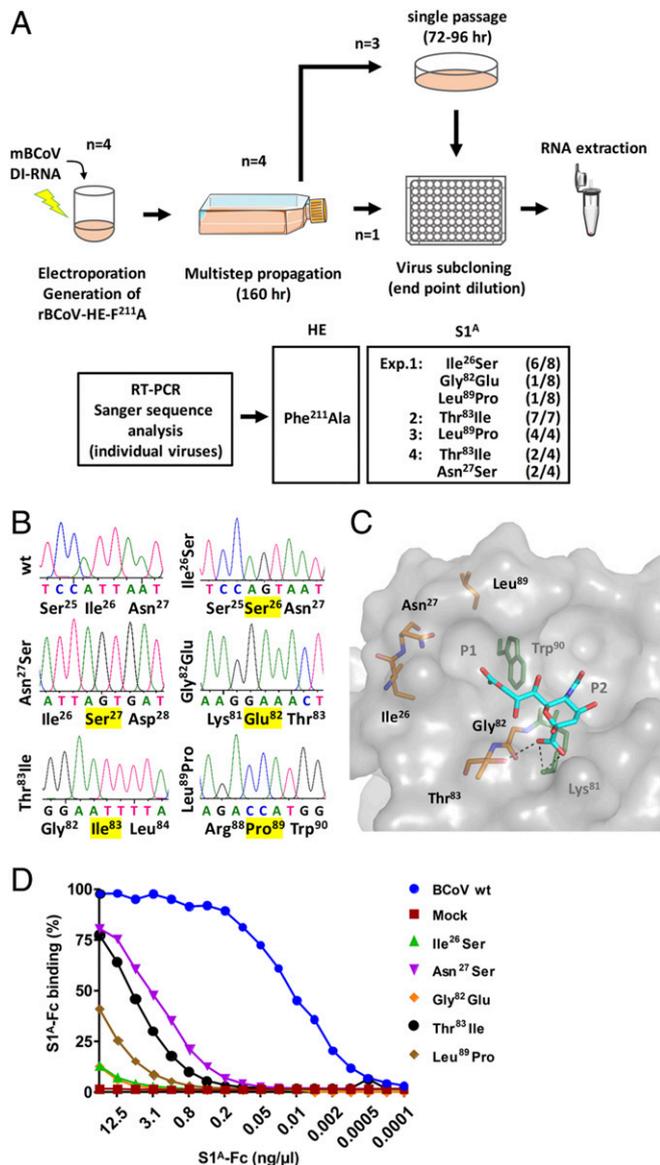
S1 <sup>A</sup> residue	E1	E2	E3	E4	E5	E6	E7
Thr-22						Ile	
Ile-26	Ser				Ser		
Asn-27				Ser	Ser	Tyr	
Val-29					Gly	Gly	
Tyr-75					Cys		
Gly-82	Glu						
Thr-83		Ile		Ile	Ile		Ile/Asn
Arg-88					Thr		
Leu-89	Pro		Pro				Pro
His-173					Tyr	Tyr	
Pro-174					Leu	Leu	
Arg-197					Cys		

Summary of S1<sup>A</sup> mutations identified in rBCoVs upon 1) introduction of a HE-Phe<sup>211</sup>Ala substitution and virus rescue by straight forward targeted recombination (E1 through E4; see also Fig. 1), 2) passage of rBCoV-HE-F<sup>211</sup>A/S<sup>wt</sup> in the absence of exogenous HE (E5 and E6), or 3) introduction of a HE-Asn<sup>114</sup>Thr substitution and subsequent viral passage (E7).

ligand binding indirectly through long range conformational effects. A third type of mutations, quasi-random Cys substitutions, apparently disrupt S1<sup>A</sup> folding by promoting nonnative disulfide-bonding. 4) Perhaps most surprisingly, quasi-species developed in which loss of HE lectin function was compensated at the level of the viral population. Minority low-affinity variants, in trial 2, constituting less than 6% of the swarm, not only sustained the replication of high-affinity variants but actually allowed the latter to flourish and amplify to become the majority phenotype. In trial 1, this took multiple passages, and spontaneous mutants with near wild-type S affinity became dominant. In trial 2, residual input virus with wild-type S presumably profited from the early emergence and rapid expansion of the initiating low-affinity variant S1<sup>A</sup>-Asn<sup>27</sup>Ser, already during the first 120 h of multistep propagation.

**S and HE Proteins Coevolve to Attain Functional Balance and Optimal Virion Avidity.** Among the first mutations fixed upon zoonotic introduction and early emergence of OC43, was a HE-Thr<sup>114</sup>Asn substitution, which created a glycosylation site at the rim of the lectin domain RBS (11) (Fig. 3A). Glycans attached to HE Asn<sup>114</sup> hamper binding to 9-*O*-Ac-Sia through steric hindrance, causing a 500-fold reduction in HE avidity (Fig. 3B) and a 125-fold in sialate-*O*-acetyltransferase-activity, respectively (Fig. 3C). We introduced the HE Thr<sup>114</sup>Asn substitution in BCoV, expecting that the glycosylation site would be rapidly lost through any of several single-nucleotide restorative mutations in HE. Indeed, NGS analysis of the virus swarm arising after targeted recombination showed the glycosylation site to be destroyed but only in 10% of the population and exclusively by Ser<sup>116</sup>Phe substitution (Fig. 3D). This mutation only partially restores HE receptor binding and receptor destruction to 0.125 and 0.17 of that of wild-type HE, respectively (Fig. 3B and C). In the vast majority of viruses, the newly introduced HE glycosylation site was retained and, instead, low-affinity S1<sup>A</sup> mutations were selected again, with S-RBS Thr<sup>83</sup> replaced either by Ile (69%)—as seen before (Figs. 1A and 2C and Table 1)—or by Asn (10%) (Fig. 3D). The latter mutation reduces S1<sup>A</sup> affinity to 0.008 of that of wild-type.

Virus cloning by endpoint dilution yielded, in three of five isolates, S1<sup>A</sup>-Thr<sup>83</sup>Asn variants with the newly introduced *N*-glycosylation site in HE still intact (HE-Thr<sup>114</sup>Asn). Furthermore, a single S1<sup>A</sup>-Thr<sup>83</sup>Ile variant was isolated, but this virus in addition had lost the *N*-glycosylation site in HE (HE-Thr<sup>114</sup>Asn/Ser<sup>116</sup>Phe) (Fig. 3D). The observations led us to entertain the possibility that the mutations in S1<sup>A</sup> and HE did not occur independently and that, even in viruses expressing low-affinity



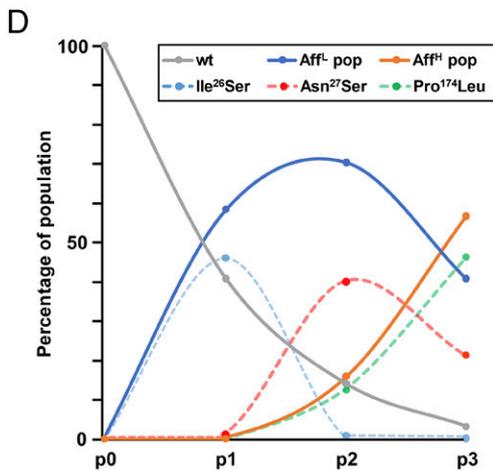
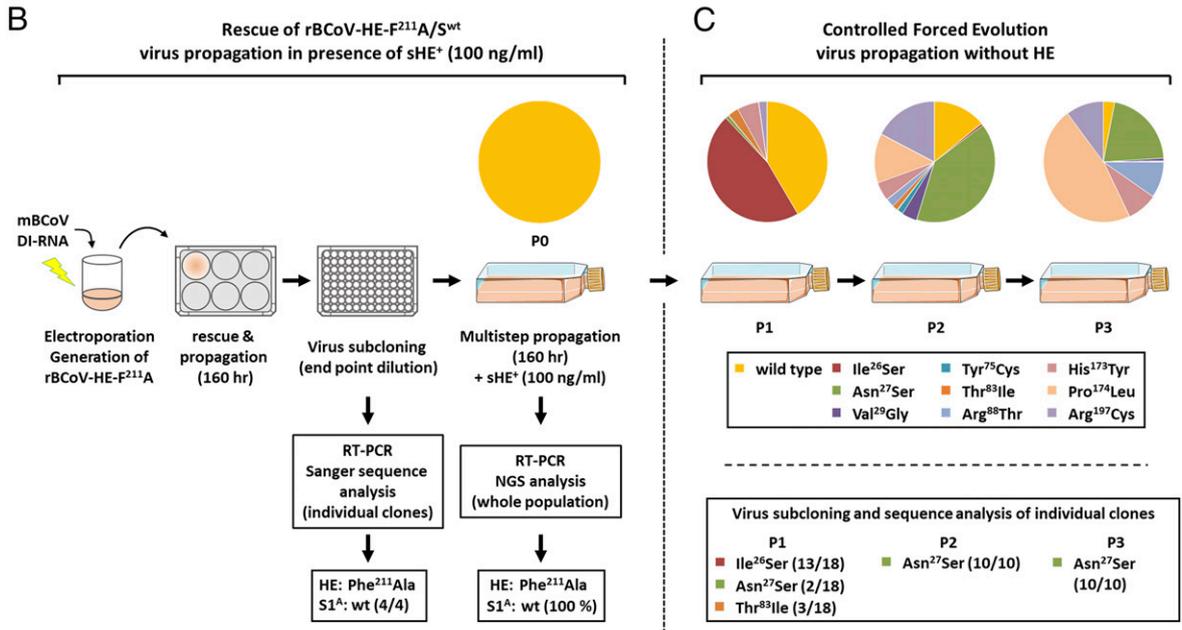
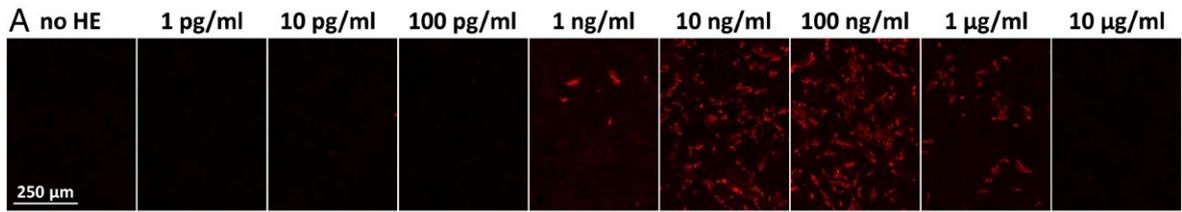
**Fig. 1.** Site-directed mutagenesis of BCoV HE by targeted recombination; loss of HE lectin activity selects for second-site mutations in domain 1<sup>A</sup> of the spike protein. (A) Schematic outline of four independent experiments, depicting each step from targeted RNA recombination and virus rescue to virus purification and genetic analysis of the resultant clonal populations by RT-PCR and Sanger sequencing. The number of virus clones, found to contain a particular S1<sup>A</sup> mutation, relative to the total number of virus clones analyzed are given (between parenthesis). (B) Relevant portions of Sanger DNA sequencing chromatograms of the S1<sup>A</sup> coding region in recombinant wild-type BCoV and in cloned rBCoV-HE-Phe<sup>211</sup>Ala derivatives. Amino acid substitutions marked in yellow. (C) Second-site mutations in S1<sup>A</sup> locate in close proximity of the RBS. Close-up of the BCoV S1<sup>A</sup> RBS (in surface representation; PDB ID code 4H14), with 9-*O*-Ac-Sia (in sticks, colored by element; oxygen, red; nitrogen, blue; carbons, cyan) modeled in the RBS (32), showing the locations of the mutations. Key elements of the RBS (hydrophobic pockets P1 and P2, and the side chains of RBS residues Lys<sup>81</sup> and Trp<sup>90</sup> in sticks, colored green) are indicated. Side chains and connecting main chains of RBS mutations also shown in sticks, but with carbon colored orange. Predicted hydrogen bonds between the Sia carboxylate moiety and the side chains of Lys<sup>81</sup> and Thr<sup>83</sup> shown as black dashed lines. (D) The S1<sup>A</sup> mutations strongly reduce binding to 9-*O*-Ac-Sia. Mutant S1<sup>A</sup>-Fc fusion proteins in twofold serial dilutions, starting at 2.5 μg per well, were tested by sp-LBA for their binding to BSM relative to that of wildtype S1<sup>A</sup>-Fc. Binding expressed in percentages with maximum binding of wildtype S1<sup>A</sup>-Fc set to 100%.

spikes, partially restorative mutations in HE would yet provide a selective advantage. To test this, the clonal S1<sup>A</sup>-Thr<sup>83</sup>Asn/HE-Thr<sup>114</sup>Asn variants were serially passaged. All three viruses independently lost HE glycosylation at Asn<sup>114</sup> over time and, saliently, again through Ser<sup>116</sup>Phe substitution exclusively. Even more remarkably, with HE-Ser<sup>116</sup>Phe mutants gaining dominance, variants emerged that had restored S affinity to (near) wild-type through substitution of S1<sup>A</sup>-Asn<sup>83</sup> either by Thr or by Ser (Fig. 3E).

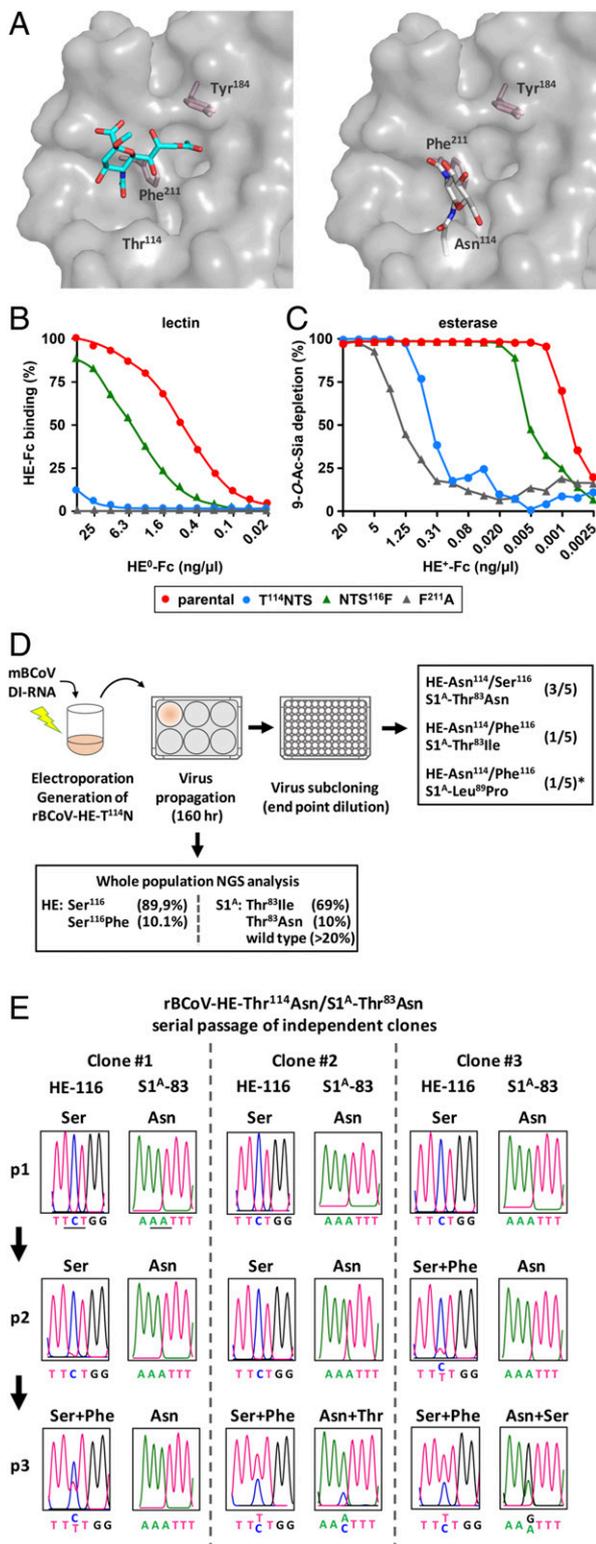
For one of the five clonal populations obtained by endpoint dilution, we unfortunately failed to determine its genotype for technical reasons. From the NGS analysis of the p1 population, we deduced that the starting mutant must have been a low-affinity S1<sup>A</sup>-Leu<sup>89</sup>Pro variant that, like the S1<sup>A</sup>-Thr<sup>83</sup>Asn/Ile variants described above, quickly lost the HE-Asn<sup>114</sup> glycan through an HE-Ser<sup>116</sup>Phe substitution. Oddly enough, the Leu<sup>89</sup>Pro substitution had not been detected by NGS in the precloning virus stock. Note, however, that this mutation had been selected before twice independently in trials with rBCoV-HE-F<sup>211</sup>A (Table 2). Possibly, it arose spontaneously during endpoint dilution procedure. Be that as it may, its *in vitro* evolution proved informative (Fig. 4A). NGS analysis of a passage p1 population, resulting from 120-h multistep propagation, showed that 100% of the viruses coded for HE-Thr<sup>114</sup>Asn/Ser<sup>116</sup>Phe in combination with S1<sup>A</sup>-Pro<sup>89</sup> (53.3%), -Thr<sup>89</sup> (40.5%), or -Ser<sup>89</sup> (2%). Note that the relationship between these variants and the course of evolution—from Leu<sup>89</sup> in the parental recombinant virus to Pro and from Pro to Thr or Ser—is evident from the codon sequences (CTA→CCA→T/ACA; mutations indicated in bold and italicized) and that the Thr<sup>89</sup> and Ser<sup>89</sup> substitutions restored S RBS affinity almost to that of wild-type RBS (Fig. 4B). All three variants (S1<sup>A</sup>-Pro<sup>89</sup>, -Thr<sup>89</sup>, and -Ser<sup>89</sup>) were readily cloned and isolated by standard endpoint dilution, and propagated independently without a requirement for exogenous RDE. p1 also contained a minor population of viruses with parental S1<sup>A</sup>, presumably regenerated from S1<sup>A</sup>-Pro<sup>89</sup>, which as for the S1<sup>A</sup>-Thr<sup>89</sup> and Ser<sup>89</sup> variants would have required only a single nucleotide substitution (CCA→CTA). Apparently, with HE lectin function partially restored, viruses that regained (near) wild-type S affinity had a selective advantage. At the end of passage p2, the low-affinity S1<sup>A</sup>-Pro<sup>89</sup> variants had dwindled to less than 1.5%, S1<sup>A</sup>-Thr<sup>89</sup> had become dominant at 75% and viruses with wild-type S1<sup>A</sup>-Leu<sup>89</sup> had rapidly risen from 2.25% in p1 to 23% (Fig. 4A).

In summary, the reduction in HE receptor-binding affinity resulting from the introduction of an *N*-linked glycosylation site triggered a series of successive mutations in S and HE. The mutations, first emerging in S, strongly decreased overall virion avidity initially, but mutants were then selected that stepwise, through loss of the HE 114-glycosylation site followed by restorative mutations in S, reverted to near wild-type avidity. Thus, the data directly demonstrate HE/S coevolution. Moreover, the findings suggest that virions are under selective pressure not only to balance receptor-binding and receptor-destroying activities in apparent relation to cell-surface receptor-densities, but also, within these constraints, to optimize virion avidity (for a schematic summary of the findings, see Fig. 4C).

**Cell Culture Adapted BCoV and OC43 Strains Differ in Their Set Point of the S/HE Balance.** Loss-of-function mutations in the BCoV HE lectin domain had a much bigger impact on BCoV propagation than expected. The prototype OC43 laboratory strain USA/1967 also lacks HE lectin function, yet in HRT18 cells it grows to titers comparable to those of BCoV reference strain Mebus. NGS of OC43 stocks revealed heterogeneity, but no indications for the existence of low S affinity minority variants that would support replication of majority high S affinity viruses. In addition, clonal



**Fig. 2.** Stable propagation and controlled directed evolution of rBCoV-HE-F<sup>211</sup>A. (A) rBCoV-HE-F<sup>211</sup>A propagation and spread is enhanced by soluble exogenous sialate-*O*-acetyltransferase. mBCoV-infected LR7 cells, donor RNA-transfected to generate rBCoV-HE-F<sup>211</sup>A, were seeded on HRT18 cell monolayers to rescue recombinant viruses with cell culture supernatants supplemented with purified BCoV HE<sup>+</sup>-Fc at concentrations indicated. Cell supernatants, harvested 120 h after seeding, were inoculated onto HRT18 cells grown on glass coverslips. After a single infectious cycle (12 hpi), infected cells were identified by immunofluorescence assay. Infected cells stained red. (B) Stable maintenance of wild-type S protein in rBCoV-HE-F<sup>211</sup>A in the presence of exogenous HE and (C) forced evolution in the absence thereof. Visual representation of experimental procedures and findings. Generation of recombinant rBCoV-HE-F<sup>211</sup>A by targeted recombination was as in Fig. 1, but with rescue, cloning, and virus amplification steps performed with tissue culture media supplemented with 100 ng/mL exogenous HE-Fc. Rescued virus was purified by endpoint dilution. Individual clonal populations were characterized for HE and S1<sup>A</sup> master sequences by extracting viral genomic RNA from the cell culture supernatant followed by RT-PCR and Sanger sequencing. One clonal population was used to grow a p0 stock of rBCoV-HE-F<sup>211</sup>A and HE and S1<sup>A</sup> diversity was assessed by NGS. The virus was used to inoculate  $5 \times 10^6$  HRT18 cells at an MOI of 0.005 TCID50 per cell and serially passaged. Cell culture supernatants were harvested at 120 hpi and virus diversity was assessed by subcloning and genetic analysis of purified viruses as in Fig. 1. In addition, viral RNA was extracted, and diversity determined by RT-PCR amplification and NGS. Frequencies of S1<sup>A</sup> variants are presented in pie charts with individual color coding as indicated. (D) Propagation of rBCoV-HE-F<sup>211</sup>A quasi-species selects for variants with (near) wild-type S affinity. Stylized graph representation depicting the emergence and decline of viral variants during serial passage of rBCoV-HE-F<sup>211</sup>A in the absence of exogenous HE. Changes in the frequencies of variants with wild-type S and groups of variants with low-affinity S (sum of Ile<sup>26</sup>Ser, Asn<sup>27</sup>Ser, Val<sup>29</sup>Gly, Thr<sup>83</sup>Ile, His<sup>173</sup>Tyr, and Arg<sup>197</sup>Cys) and high-affinity S (sum of Arg<sup>88</sup>Thr and Pro<sup>174</sup>Leu) are depicted with solid lines, colored in gray, blue, and orange, respectively. Those of individual S variants, Ile<sup>26</sup>Ser, Asn<sup>27</sup>Ser, and Pro<sup>174</sup>Leu, are shown in dashed lines and colored light blue, red, and green, respectively.



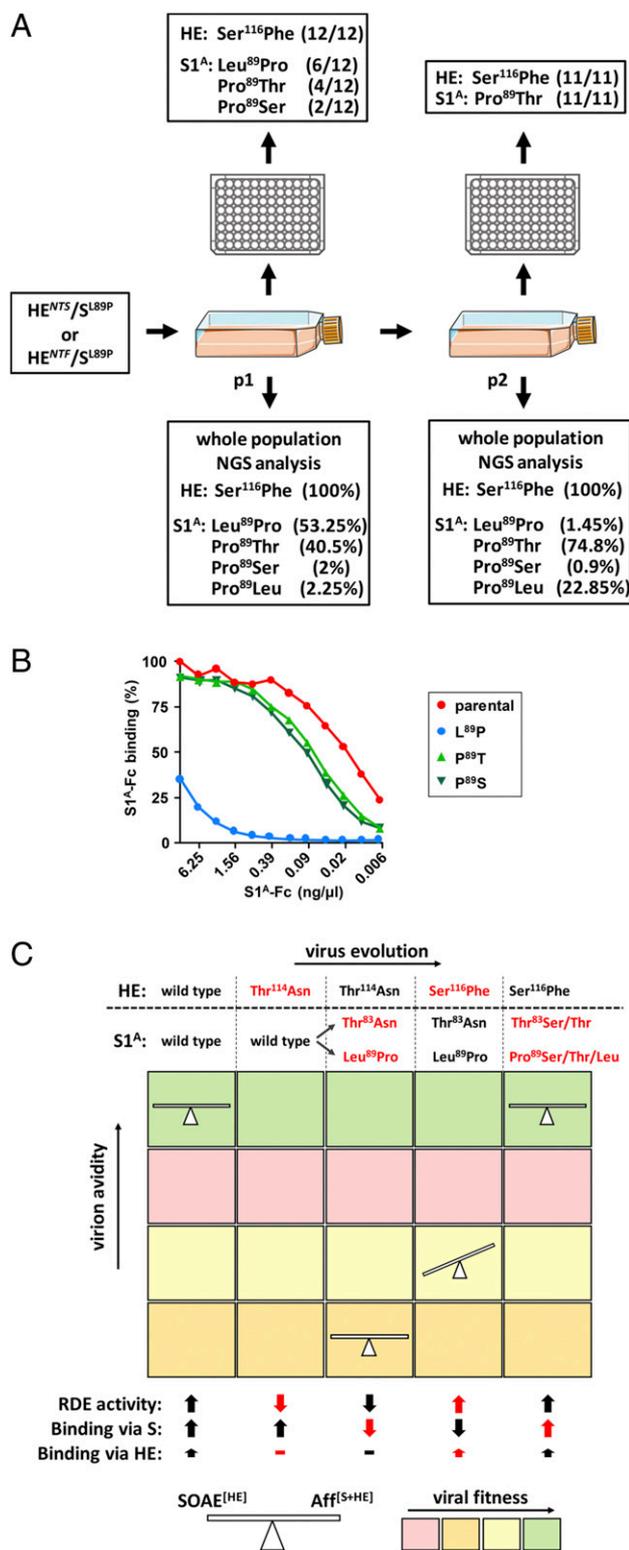
**Fig. 3.** BCoV S and HE coevolve to restore functional balance and optimal virion avidity. (A) Introduction of an *N*-glycosylation site at the rim of the HE RBS. Side-by-side close ups of the BCoV HE holoprotein (17) (PDB ID code 3CL5) in surface representation with 9-*O*-Ac-Sia (in sticks, colored by element; oxygen, red; nitrogen, blue; carbons, cyan) bound to the RBS (Left) or without the ligand and with the Thr<sup>114</sup>Asn substitution and *N*-linked glycan modeled in the RBS (in sticks, colored by element as above but with carbons in white) (Right). Modeling performed by superpositioning of the OC43 strain NL/A/2005 HE structure (PDB ID code 5N11) and confirmed in Coot. HE RBS key residues are indicated, with side chains of Tyr<sup>184</sup> and Phe<sup>211</sup> shown

virus populations obtained by endpoint dilution (10 of 10) all conformed to the S1<sup>A</sup> master sequence. We offer that instead OC43-USA/1967 may have reached a viable S/HE balance compatible with efficient *in vitro* propagation through adaptations in S that reduced receptor-binding affinity or altered receptor fine-specificity. When measured by solid-phase assay with bivalent S1<sup>A</sup>-Fc fusion proteins, binding of the S protein of OC43 USA/1967 to BSM, containing both mono- and di-*O*-acetylated  $\alpha$ 2,6-sialoglycans, is 16- to 32-fold lower than that of BCoV-Mebus (32). sp-LBA with BSM preparations, selectively depleted for either 9-*O*- or 7,9-di-*O*-Sias, showed that BCoV S, like BCoV HE (28), preferentially binds to 7,9-di-*O*-Ac-Sias (SI Appendix, Fig. S6A). OC43 USA/1967 S1<sup>A</sup> may not share this preference. Apparently due to its low affinity, detectable binding to BSM was lost upon depletion of either type of Sia (SI Appendix, Fig. S6A). Moreover, even though BCoV-Mebus S preferably binds to 7,9-di-*O*-Ac-Sia, monovalent one-on-one binding of the BCoV S1<sup>A</sup> domain to  $\alpha$ 2,6-linked 9-*O*-acetylated Sia is still threefold stronger than that of OC43-USA/1967 as measured by bilayer interferometry (SI Appendix, Fig. S6B). On a cautionary note, the isolation and complex passage history of OC43-USA/1967 (48, 49) entailed several passages in human tracheal organ culture, suckling mouse brain and many rounds of replication in cultured cells (9), which would have given the virus ample opportunity to adapt to the *in vitro* conditions. Thus, the binding characteristics of its spike may not faithfully reflect those in circulating field variants. Indeed, OC43 variants in sputum samples, contrary to the OC43 USA/1967, replicate in airway epithelial cell cultures but not in tissue culture cells (50).

## Discussion

**Coevolution and Functional Interdependence of Embecovirus S and HE Proteins.** Our findings demonstrate that in the prototypic  $\beta$ 1CoV BCoV the envelope proteins S and HE are functionally entwined and coevolve. We posit that the same holds for other members of the species *Betacoronavirus-1*, including its zoonotic descendant human CoV OC43 and related viruses of swine, rabbits, dogs, and horses, as well as for other *Embecovirus* species, most prominently among which human CoV HKU1. The data lead us to conclude that the respective activities of S and HE in receptor-binding and catalysis-driven virion elution are balanced to ensure dynamic reversible virion attachment and, thereby, efficient virus propagation. In consequence, for the viruses listed above, the roles of S and HE during natural infection cannot be understood in isolation but must be considered in unison.

in sticks. (B) Loss of HE lectin function upon introduction of an *N*-glycosylation site through HE-Thr<sup>114</sup>Asn substitution and partial restoration through a second-site Ser<sup>116</sup>Phe mutation. Sp-LBA with serial dilutions of enzyme-inactive HE<sup>0</sup>-Fc as in Fig. 1D. (C) Consequences for HE esterase activity toward clustered ligands. On-the-plate receptor-depletion assay with BSM as substrate as in ref. 11. The assay was performed with twofold serial dilutions of enzymatically active HE<sup>1</sup>-Fc and residual 9-*O*-Ac-Sia measured by sp-LBA with a fixed amount of HE<sup>0</sup>-Fc. (D) Visual representation of experimental procedures and findings as in Figs. 1 and 2. Note that of five virus clones purified by endpoint dilution, the identity of one isolate could not be established and was deduced from subsequent propagation experiments as explained in the text (marked with an asterisk; see also Fig. 4). (E) Serial passage of rBCoV-HE-Thr<sup>114</sup>Asn/S1<sup>A</sup>-Thr<sup>83</sup>Asn selects for successive mutations in HE and S to restore their function to near wild-type levels. Results are shown for three independent isolates. Viral RNA extracted from tissue culture supernatants collected at the end of each passage was characterized by RT-PCR and Sanger DNA sequencing. Relevant portions of Sanger DNA sequencing chromatograms are presented to show changes in the master sequence of the virus populations at the coding sequence for the *N*-glycosylation site (HE codon 116) and for the site of the low affinity S mutation selected initially (S1<sup>A</sup> codon 83).



**Fig. 4.** Serial passage of rBCoV-HE-T<sup>114</sup>A selects for successive mutations in HE and S to optimize HE/S functional balance and overall virion avidity. (A) Continued serial passage of rBCoV-HE-T<sup>114</sup>N+S<sup>116</sup>F/S<sup>1A</sup>-L<sup>89</sup>P. Schematic outline of the experiment and presentation of results of NGS analysis and genetic characterization of subcloned variants as in Figs. 1–3. (B) Resubstitution of S<sup>1A</sup>-Leu<sup>89</sup>Pro by Thr or Ser restores S<sup>1A</sup> binding to near wild-type levels. Sp-LBA as in Fig. 1D. (C) BCoV is under selective pressure for an optimal functional balance between virion attachment and catalysis-driven release as well as for optimal virion avidity. Schematic summary of the evidence for HE/S coevolution and our interpretation thereof in a 2D chart. The course of

Using a reverse genetics-based forced evolution approach with BCoV as a model system, we showed that loss of HE lectin function causes an offset in S/HE balance, practically incompatible with virus propagation and spread. With the HE lectin domain as modulator of esterase activity, mutations that decrease or abolish HE RBS affinity reduce virion-associated sialate-*O*-acetyltransferase activity toward clustered glycotopes on hypervalent glycoconjugates (11), such as are present in the mucus and glycocalyx in natural tissues and on the surface of cultured cells. The extent of the resultant defect is such that compensatory second-site mutations in S are selected for that dramatically reduce S RBS affinity, apparently to restore reversibility of binding and virion motility as an escape ticket from inadvertent virion attachment to decoy receptors.

The single-amino acid mutations in receptor-binding domain S<sup>1A</sup> were limited to a finite number of positions. They were either within or proximal to the RBS to directly affect protein–ligand interactions, or more distal to reduce RBS affinity through long-range effects or by disrupting local folding through aberrant disulfide-bonding. Whereas the parental recombinant viruses, defective in HE lectin function but with wild-type S RBS affinity, require an external source of receptor-destroying enzyme for propagation, their progeny escape mutants regained propagation-independence by lowering S affinity.

In expanding clonal populations of HE-defective rBCoV-HE-F<sup>211</sup>A, propagated in the absence of exogenous receptor-destroying enzyme activity, viruses with reduced S affinity gained a selective advantage initially. Eventually, however, quasi-species developed in which loss of HE lectin function was compensated at the population level. Variants that combined the HE-Phe<sup>211</sup>Ala mutation with (near) wild-type affinity S proteins increased to dominate the swarm at least numerically. Still, these high-affinity S variants for their proliferation were strictly reliant on minority low-affinity S variants. This relationship extends beyond cooperativity and group selection described for other systems (51–57) and amounts to a state of dependency. We propose that the virions of the low-affinity minority variants provide aid by serving as a source of exogenous sialate-*O*-acetyltransferase activity. They evade decoy receptors through enhanced reversibility of virion attachment, but this phenomenon increases their motility—whether by sliding diffusion or binding/rebinding—causing them to deplete cell surface 9-*O*-Ac-Sias, decoy receptors, and functional receptors alike. With increasing concentrations of low-affinity virions in the culture supernatant, high-affinity variants would profit progressively, whereas falling cell surface receptor densities would put the low-affinity viruses increasingly at a disadvantage.

The forced evolution trials performed with rBCoV-HE-F<sup>211</sup>A were restricted in course and outcome by design, because full reversion would require simultaneous mutation of two adjacent nucleotides. Moreover, the crucial role of the Phe<sup>211</sup> in ligand

evolution of rBCoV-HE-Thr<sup>114</sup>Asn (direction indicated by arrow) is shown for two types of low-affinity S escape variants (Thr<sup>114</sup>Asn and Leu<sup>89</sup>Pro) with the succession of mutations in HE and S selected for during serial passage (Top) brought in relation to 1) overall virion avidity, as mediated by S and HE, on the y axis from low to high as indicated by the arrow; 2) viral fitness, color-coded from low (pink) to high (green) as indicated in the color legend at the bottom; 3) the effect of the mutations on the function of S (attachment) and HE (RDE activity and attachment) as indicated by thick arrows (arrows pointing up, near wild-type activity; arrows pointing down, decreased function; dash, total loss of function); the size difference between arrows for S and HE reflect the difference in their contribution to virion binding; to indicate the effect of newly emerging mutations corresponding arrows for function are colored red); and 4) their effect on HE/S functional balance (as indicated by the position of the scale). SOAE<sup>[HE]</sup>, virion-associated receptor-destroying sialate-*O*-acetyltransferase activity; Aff<sup>[S+HE]</sup>, overall virion avidity as determined by the affinities of the RBSs of all S and HE proteins combined.

binding (*SI Appendix, Fig. S1*) obviates conservative substitutions (17). Although rBCoV-HE<sup>A211V</sup> variants did emerge in two separate experiments, this mutation only marginally increases HE RBS affinity and sialate-*O*-acetyltransferase activity.

In contrast to the HE-Phe<sup>211</sup>Ala mutation, the deleterious effect of *N*-glycosylation at HE-Asn<sup>114</sup> can be reversed, completely or partially, through various single-nucleotide substitutions in codons 114 and 116 and would therefore more readily allow for compensatory mutations also in HE. Indeed, serial passage of the rBCoV-HE-Thr<sup>114</sup>Asn resulted in a succession of mutations alternatingly in HE and S. The order of appearance of these mutations and their effect on protein function indicated that they were not fixed to merely restore the balance between attachment and catalysis-driven virion elution. The HE-Thr<sup>114</sup>Asn substitution initially selected for second-site mutations that reduced S affinity (Thr<sup>83</sup>Ile, Thr<sup>83</sup>Asn, and Leu<sup>89</sup>Pro), but with propagation thus recovered, derivatives rapidly emerged with increased HE lectin and esterase activity through a Ser<sup>116</sup>Phe mutation. Apparently, this created an HE/S disbalance that in turn favored the selection of viruses with revertant mutations in S that raised S RBS affinity again to wild-type (Thr<sup>83</sup> → Ile → Thr; Leu<sup>89</sup> → Pro → Leu) or near wild-type levels (Thr<sup>83</sup> → Ile → Ser; Leu<sup>89</sup> → Pro → Thr/Ser). Conjointly, our findings indicate that through an initial sharp reduction in overall avidity, compensatory to loss of HE function, virus particles regained the capacity of eluding nonproductive attachment to decoy receptors but at a fitness penalty. The decrease in S RBS affinity would predictably lower the specific infectivity of virus particles through a decrease in productive host cell attachment. The rapid selection of the HE-Ser<sup>116</sup>Phe mutation in a low-affinity S background can thus be understood to have increased virion avidity, albeit through HE and rather than through S. HE does have a dual function after all, and in influenza viruses C and D as well as in murine coronavirus-1, it is a receptor-binding protein first and foremost (22, 35, 58–61). Of note, the partial Ser<sup>116</sup>Phe reversion of HE consistently seen in multiple independent experiments suggests that a return to (near) wild-type lectin and esterase activity along with a low-affinity S would have tipped the scale too much toward catalytic virion release. We posit that in addition to an optimal balance between receptor binding and receptor destruction, the system strives toward optimal virion avidity (Fig. 4C).

Under natural circumstances, the set-point of the S/HE balance would be tailored to conditions met in the target tissues of the intact host. The spontaneous loss of HE lectin function in OC43 and HKU1 may thus be understood to have arisen through convergent evolution as an adaptation to the sialoglycan composition of the mucus in the human upper respiratory tract, and that of the glycocalyx of the respiratory epithelia (11). This change, which would predictably reduce virion-associated receptor destruction and hence decrease virion elution, might have been selected for by low density occurrence of 9-*O*-Ac-sialoglycans in the human upper airways. In accordance, limited tissue array analyses with HE-based virolectins suggested that these sugars are not particularly prevalent in the human respiratory tract and by far not as ubiquitous as in the gut (28). However, full understanding of how the S/HE balance was reset in OC43 and HKU1 upon their zoonotic introduction and why awaits further analysis of the binding properties and ligand fine-specificity of the S proteins of naturally occurring variants, as well as more quantitative and comprehensive interhost comparative analyses of airway sialoglycomes. As an added complication, virion particles encounter widely different circumstances while traversing the mucus layer, at the epithelial cell surface, during local cell-to-cell dissemination, and during transmission. It is an open question whether this selects for majority phenotypes that can cope individually and independently with each of these different conditions by striking an uneasy compromise with

regard to HE/S balance and overall virion avidity, or whether there is loco-temporal selection for swarms of variants that collectively allow the virus population as a whole to overcome each hurdle.

#### Similarities between Embeco- and Influenza A Viruses Point to Common Principles of Virion-Sialoglycan Receptor-Usage.

The embecovirus HE gene originated from a horizontal gene transfer event, presumably with an influenza C/D-like virus as donor (17, 62). Like the orthomyxovirus hemagglutinin-esterase-fusion proteins, the newly acquired coronavirus HE protein provided the acceptor virus with an opportunity to reversibly bind to 9-*O*-Ac-sialoglycans (26). This in turn would seem to have prompted a shift in the receptor-specificity of S through adaptations in S1<sup>A</sup> that created a 9-*O*-Ac-Sia binding site de novo so that virions could now attach to these receptor determinants also via S. The embecoviruses thus adopted a strategy of receptor usage entailing a concerted and carefully fine-tuned activity of two envelope proteins that is unique among coronaviruses, but uncannily similar to that of influenza A viruses. In the latter, the hemagglutinin protein HA, as a pendant of S, mediates binding to either  $\alpha$ 2,3- or  $\alpha$ 2,6-linked sialosides, while the neuraminidase (NA), like HE, is a receptor-destroying enzyme but with a substrate fine-specificity that closely matches HA ligand preference (63). For influenza A virus, the existence and biological relevance of a functional balance between receptor binding and receptor destruction is well recognized (64–67). This balance is critical for receptor-associated virus motility through the mucus and at the cell surface (41, 43, 44, 68–70). Complete or partial loss of NA activity—whether invoked spontaneously, through reverse genetics or by viral propagation in the presence of NA inhibitors—selects for mutations around the HA receptor-binding pocket that reduce HA affinity (65, 71, 72).

Furthermore, as proposed here for HE, NA contributes to virion attachment and even compensates for loss of virion avidity in mutant viruses with reduced HA affinity (73). Different from HE, NA may do so via its catalytic pocket, which doubles as a Sia-binding site (74). However, NA also possesses a second Sia binding site (75, 76), which like the HE lectin domain, regulates NA activity and which, in further analogy, is conserved or lost in apparent correlation with host tropism (77–79). Finally, among many other similarities to embecoviruses, influenza A variants with different set points in their HA-NA functional balance may cooperate to support their propagation in cultured cells (53). Our observations establish that there are common principles of virion-sialoglycan interactions that prompted convergent evolution of  $\beta$ 1CoVs and influenza A viruses. Although these two groups of viruses essentially differ in genome type and replication strategy, envelope proteins, and receptors, they seem to be subject to the same rules of engagement with respect to dynamic receptor binding, the differences between them constituting variations on a theme. This implies that observations made for the one system are informative for the other. Perhaps more importantly, insight into the overriding principles of virus-glycan interactions may open avenues to common strategies for antiviral intervention.

#### Materials and Methods

**Cells and Viruses.** HRT18 (ATCC CCL244) and mouse LR7 (45) cells were maintained in DMEM containing 10% fetal calf serum (FCS), penicillin (100 IU/mL), and streptomycin (100  $\mu$ g/mL). BCoV strain Mebus and OC43 strain USA/1967, purchased from the American Type Culture Collection (ATCC), were propagated in HRT18 cells.

**Reverse Genetics through Targeted Recombination.** A reverse genetics system based on targeted RNA recombination was developed for BCoV strain Mebus essentially as described previously (15, 45, 80). Using conventional cloning methods, RT-PCR amplicons of the 5'-terminal 601 nt and 3'-terminal 9,292 nt of the BCoV strain Mebus genome (reference GenBank sequence

U00735.2) were fused and cloned in plasmid pUC57, downstream of a T7 RNA polymerase promoter and upstream of a 25-nt poly(A) tract and a PacI site, yielding pD-BCoV1. From this construct, BCoV ORF 4a was deleted (nucleotides 27740 to 27853) and replaced by the Renilla luciferase (Rluc) gene, yielding pD-BCoV-Rluc. A second pD-BCoV1 derivative, pD-mBCoVΔHE, was created by replacing the coding sequence for the ectodomain of BCoV S (nucleotides 23641 to 27433) by the corresponding MHV-A59 sequence and by deleting the BCoV HE gene (nucleotides 22406 to 23623). The nucleotide sequences of pD-BCoV1, pD-BCoV-Rluc and pD-mBCoVΔHE, determined by bidirectional Sanger sequence analysis, were deposited in GenBank (accession codes: MT939521–MT939523).

To generate a recombinant chimeric acceptor virus, mBCoVΔHE, HRT18 cells were infected with BCoV-Mebus at MOI of 10 TCID<sub>50</sub> per cell and trypsinized and resuspended in PBS. An aliquot of this suspension, containing  $1.5 \times 10^6$  cells in 0.8 mL, was mixed with capped synthetic RNA that had been produced by in vitro transcription using the mMESSAGE mMACHINE T7 Transcription Kit (Thermo Fisher) with PacI-linearized mBCoVΔHE vector as template. The mixture was subjected to two consecutive electrical pulses of 850 V at 20 μF with a Gene Pulser II electroporator (Bio-Rad) and the cells were then seeded on a confluent monolayer of LR7 feeder cells in a 35-mm dish. Incubation was continued at 37 °C, 5% CO<sub>2</sub> for 18 h posttransfection until wide-spread cytopathic effect was apparent. The cell culture supernatant was harvested and cleared by low-speed centrifugation at 1,200 rpm, and mBCoVΔHE was purified by end-point dilution and used to generate stocks for future usage in LR7 cells.

To generate luciferase-expressing rBCoVs with the BCoV HE and S genes reconstituted (i.e., rBCoV<sup>wt</sup> or rBCoV-HE-Phe<sup>211</sup>Ala) LR7 cells, infected with mBCoVΔHE at MOI of 5, were electroporated as described above with synthetic RNA transcribed from pD-BCoV-Rluc and derivatives thereof. The infected and transfected cells were then seeded on HRT18 cell monolayers in 35-mm plates for up to 160 h. For rescue and propagation of rBCoV-HE-Phe<sup>211</sup>Ala without second-site mutations in S, the cell culture supernatants were supplemented with 100 ng/mL of BCoV HE-Fc protein (17). After 5 to 7 d of incubation at 37 °C, samples of the cell culture supernatants were tested for infectivity by transferring them to HRT18 cell monolayers grown on 12-mm glass coverslips in 15.6-mm wells. Incubation was continued for 12 h, after which the cells were fixed with paraformaldehyde and immunofluorescence staining was performed with polyclonal antiserum from a BCoV-infected cow.

**Virus Titration, Purification, and Characterization of Viral Populations.** mBCoV was titrated and cloned by endpoint dilution on LR7 cells with cytopathic effect as read-out. rBCoVs were titrated and cloned in HRT18 cells. To identify infected wells, cell supernatants were analyzed by HAA with rat erythrocytes (26) and by Renilla luciferase assay (Dual-Luciferase Reporter Assay System, Promega). Titers were calculated by the Spearman-Kärber formula. Clonal virus populations were characterized by isolating viral RNA from 150-μL aliquots of the cell culture supernatant with the NucleoSpin RNA Virus kit (Macherey-Nagel) followed by conventional RT-PCR and bidirectional Sanger sequence analysis.

**Controlled Forced Evolution Experiments.** Confluent HRT18 monolayers ( $5 \times 10^6$  cells) grown in 25-cm<sup>2</sup> flasks, were inoculated with rBCoV-HE-F<sup>211</sup>A at MOI 0.005 in PBS for 1 h at 37 °C. The cells were washed three times with PBS to remove residual exogenous HE-Fc and incubation was continued in DMEM + 10% FCS at 37 °C, 5% CO<sub>2</sub> for 120 hpi with samples collected every 24 h (p1). Subsequent 120-h passages were performed by adding 10 μL of supernatant to new cultures of HRT18 cells in 25-cm<sup>2</sup> flasks.

**Expression and Purification of HE-Fc and S1<sup>A</sup>-Fc Proteins.** BCoV HE, either enzymatically-active (HE<sup>+</sup>) or rendered inactive through a Ser<sup>40</sup>Ala substitution (HE<sup>0</sup>), and OC43 S1<sup>A</sup> were expressed as Fc fusion proteins in HEK293T cells and purified from the cell supernatant by protein A affinity chromatography, as detailed previously (17, 32). Monomeric S1<sup>A</sup> was

obtained by on-the bead thrombin cleavage (32). pCD5-BCoVHE-T-Fc vectors (17) encoding mutant BCoV HE derivatives were constructed with the Q5 Site-Directed Mutagenesis Kit per the instructions of the manufacturer.

**Pseudovirus Entry Assays.** The production of BCoV S-pseudotyped VSV-ΔG particles, their characterization by Western blot analysis, and infectivity assays in HRT18 cells were as described previously (32).

**sp-LBA.** sp-LBA was performed as described previously (32) with BSM (Sigma-Aldrich), coated to 96-well Maxisorp microtitre ELISA plates (Nunc, 0.1 μg BSM per well), serving as a ligand. Binding assays were performed with twofold serial dilutions of HE<sup>0</sup>-Fc, S1<sup>A</sup>-Fc, or mutated derivatives thereof. Receptor-destroying esterase activities of soluble HEs were measured by on-the-plate 9-O-Ac-Sia depletion assays, as described previously (11, 28).

**Hemagglutination Assay.** HAA was performed with rat erythrocytes (*Rattus norvegicus* strain Wistar; 50% suspension in PBS). Standard HAA was done with twofold serial dilutions of HE<sup>0</sup>-Fc proteins (starting at 25 ng per well) as described previously (17). High-sensitivity nanoparticle HAA (NP-HAA) was performed as in refs. 32 and 81. Briefly, self-assembling 60-meric nanoparticles, comprised of lumazine synthase (LS), N-terminally extended with the immunoglobulin Fc-binding domain of the *Staphylococcus aureus* protein A, were complexed with HE<sup>0</sup>-Fc proteins at a 1:0.6 molar ratio for 30 min on ice. The HE<sup>0</sup>-Fc-loaded nanoparticles were then twofold serially diluted and mixed 1:1 (vol/vol) with rat erythrocytes (0.5% in PBS). Incubation was for 2 h at 4 °C after which HAA titers were read.

**NGS Analysis.** Viral RNA from culture supernatants was isolated as described above. HE and S1<sup>A</sup> coding regions from viral genome of different virus populations were obtained by RT-PCR with primer sets HE<sub>F</sub> 5'-TTAGATTATGGTCTAAGCATCATG-3' and HE<sub>R</sub> 5'-TTAGATTATGGTCTAAGCATCATG-3', S1<sup>A</sup><sub>F</sub> 5'-ACCATGTTTTGATACTTTTA-3' and S1<sup>A</sup><sub>R</sub> 5'-AGATTGTGTTTACACTT-AATCTC-3', respectively. Amplicons were processed in the NGSgo workflow for Illumina according to the Instructions for Use (Edition 4), except that the fragmentation was prolonged to 40 min at 25 °C (protocol 3A). Briefly, amplicons were subjected to fragmentation and adapter ligation using NGSgo-LibrX (GenDx). Size selection and clean-up of the samples was performed with SPRI beads (Machery-Nagel). Unique barcodes were ligated to each sample using NGSgo-IndX (GenDx), after which all samples were pooled and subsequently purified with SPRI beads, resulting in a library of fragments between ~400 and 1,000 bp. The DNA fragments were denatured and paired-end sequenced on a MiSeq platform (Illumina) using a 300-cycle kit (V2). FASTQ files were analyzed in NGSengine (GenDx), which aligned the reads to the reference sequences of HE and S (reference GenBank sequence U00735.2 for BCoV strain Mebus, and NC\_006213.1 for OC43 strain USA/1967). For the characterization of each virus sample, amplicons from five independent RT-PCR reactions were analyzed in parallel and mutation frequencies were determined by averaging the results from these five replicates.

**Data Availability.** All study data are included in the article and supporting information. The nucleotide sequences of pD-BCoV1, pD-BCoV-Rluc, and pD-mBCoVΔHE, determined by bidirectional Sanger sequence analysis, were deposited in GenBank (accession codes: MT939521–MT939523).

**ACKNOWLEDGMENTS.** We thank Dr. Jolanda de Groot-Mijnes (Medical Microbiology, University Medical Center Utrecht) for critical reading of the manuscript and valuable suggestions. This study was supported by China Scholarship Council Grant 2014-03250042 (to Y.L.) and by TOP-PUNT Grant 718.015.003 of the Netherlands Organization for Scientific Research (to G.-J.B.).

1. E. de Wit, N. van Doremalen, D. Falzarano, V. J. Munster, SARS and MERS: Recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
2. C. B. E. M. Reusken, V. S. Raj, M. P. Koopmans, B. L. Haagmans, Cross host transmission in the emergence of MERS coronavirus. *Curr. Opin. Virol.* **16**, 55–62 (2016).
3. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, A novel coronavirus outbreak of global health concern. *Lancet* **395**, 470–473 (2020).
4. P. Zhou et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
5. P. C. Y. Woo, S. K. P. Lau, C. C. Y. Yip, Y. Huang, K. Y. Yuen, More and more coronaviruses: Human coronavirus HKU1. *Viruses* **1**, 57–71 (2009).

6. S. Su et al., Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* **24**, 490–502 (2016).
7. R. de Groot, S. Baker, R. Baric, *Family Coronaviridae*, (Elsevier Amsterdam, 2011), pp. 806–828.
8. L. Vijgen et al., Evolutionary history of the closely related group 2 coronaviruses: Porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43. *J. Virol.* **80**, 7270–7274 (2006).
9. L. Vijgen et al., Complete genomic sequence of human coronavirus OC43: Molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J. Virol.* **79**, 1595–1604 (2005).

10. S. K. P. Lau *et al.*, Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. *J. Virol.* **85**, 11325–11337 (2011).
11. M. J. G. Bakkers *et al.*, Betacoronavirus adaptation to humans involved progressive loss of hemagglutinin-esterase lectin activity. *Cell Host Microbe* **21**, 356–366 (2017).
12. R. J. G. Hulswit, C. A. M. de Haan, B. J. Bosch, Coronavirus spike protein and tropism changes. *Adv. Virus Res.* **96**, 29–57 (2016).
13. B. King, D. A. Brian, Bovine coronavirus structural proteins. *J. Virol.* **42**, 700–707 (1982).
14. T. E. Klenzle, S. Abraham, B. G. Hogue, D. A. Brian, Structure and orientation of expressed bovine coronavirus hemagglutinin-esterase protein. *J. Virol.* **64**, 1834–1838 (1990).
15. A. Lissenberg *et al.*, Luxury at a cost? Recombinant mouse hepatitis viruses expressing the accessory hemagglutinin esterase protein display reduced fitness in vitro. *J. Virol.* **79**, 15054–15063 (2005).
16. R. J. de Groot, Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj. J.* **23**, 59–72 (2006).
17. Q. Zeng, M. A. Langereis, A. L. W. van Vliet, E. G. Huizinga, R. J. de Groot, Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9065–9069 (2008).
18. B. Schultze, K. Wahn, H. D. Klenk, G. Herrler, Isolated HE-protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. *Virology* **180**, 221–228 (1991).
19. R. Vlasak, W. Luytjes, J. Leider, W. Spaan, P. Palese, The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. *J. Virol.* **62**, 4686–4690 (1988).
20. B. King, B. J. Potts, D. A. Brian, Bovine coronavirus hemagglutinin protein. *Virus Res.* **2**, 53–59 (1985).
21. M. J. G. Bakkers *et al.*, Coronavirus receptor switch explained from the stereochemistry of protein-carbohydrate interactions and a single mutation. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3111–E3119 (2016).
22. M. A. Langereis, Q. Zeng, B. A. Heesters, E. G. Huizinga, R. J. de Groot, The murine coronavirus hemagglutinin-esterase receptor-binding site: A major shift in ligand specificity through modest changes in architecture. *PLoS Pathog.* **8**, e1002492 (2012).
23. S. L. Smits *et al.*, Nidovirus sialate-O-acetylsterases: Evolution and substrate specificity of coronaviral and toroviral receptor-destroying enzymes. *J. Biol. Chem.* **280**, 6933–6941 (2005).
24. G. Regl *et al.*, The hemagglutinin-esterase of mouse hepatitis virus strain 5 is a sialate-4-O-acetylsterase. *J. Virol.* **73**, 4721–4727 (1999).
25. W. J. Wurzer, K. Obojes, R. Vlasak, The sialate-4-O-acetylsterases of coronaviruses related to mouse hepatitis virus: A proposal to reorganize group 2 coronaviridae. *J. Gen. Virol.* **83**, 395–402 (2002).
26. M. A. Langereis, A. L. W. van Vliet, W. Boot, R. J. de Groot, 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 (2010).
27. R. K. Williams, G. S. Jiang, K. V. Holmes, Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5533–5536 (1991).
28. M. A. Langereis *et al.*, Complexity and diversity of the mammalian sialome revealed by nidovirus virolectins. *Cell Rep.* **11**, 1966–1978 (2015).
29. B. Schultze, H. J. Gross, R. Brossmer, G. Herrler, The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. *J. Virol.* **65**, 6232–6237 (1991).
30. X. Huang *et al.*, 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 (2015).
31. F. Künkel, G. Herrler, Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* **195**, 195–202 (1993).
32. R. J. G. Hulswit *et al.*, Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2681–2690 (2019).
33. M. A. Langereis *et al.*, Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15897–15902 (2009).
34. P. B. Rosenthal *et al.*, Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* **396**, 92–96 (1998).
35. H. Song *et al.*, 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 (2016).
36. M. A. Tortorici *et al.*, Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* **26**, 481–489 (2019).
37. M. Desforges, J. Desjardins, C. Zhang, P. J. Talbot, 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 (2013).
38. D. Deregt, L. A. Babjuk, Monoclonal antibodies to bovine coronavirus: Characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. *Virology* **161**, 410–420 (1987).
39. D. Deregt *et al.*, Monoclonal antibodies to bovine coronavirus glycoproteins E2 and E3: Demonstration of in vivo virus-neutralizing activity. *J. Gen. Virol.* **70**, 993–998 (1989).
40. G. Milane, A. B. Kourtesis, S. Dea, Characterization of monoclonal antibodies to the hemagglutinin-esterase glycoprotein of a bovine coronavirus associated with winter dysentery and cross-reactivity to field isolates. *J. Clin. Microbiol.* **35**, 33–40 (1997).
41. T. Sakai, S. I. Nishimura, T. Naito, M. Saito, Influenza A virus hemagglutinin and neuraminidase act as novel motile machinery. *Sci. Rep.* **7**, 45043 (2017).
42. T. Sakai, H. Takagi, Y. Muraki, M. Saito, Unique directional motility of influenza C virus controlled by its filamentous morphology and short-range motions. *J. Virol.* **92**, e01522-17 (2018).
43. H. Guo *et al.*, Kinetic analysis of the influenza A virus HA/NA balance reveals contribution of NA to virus-receptor binding and NA-dependent rolling on receptor-containing surfaces. *PLoS Pathog.* **14**, e1007233 (2018).
44. E. de Vries, W. Du, H. Guo, C. A. M. de Haan, Influenza A virus hemagglutinin-neuraminidase-receptor balance: Preserving virus motility. *Trends Microbiol.* **28**, 57–67 (2020).
45. L. Kuo, G.-J. Godeke, M. J. B. Raamsman, P. S. Masters, P. J. M. Rottier, Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: Crossing the host cell species barrier. *J. Virol.* **74**, 1393–1406 (2000).
46. P. S. Masters, P. J. M. Rottier, Coronavirus reverse genetics by targeted RNA recombination. *Curr. Top. Microbiol. Immunol.* **287**, 133–159 (2005).
47. A. B. Boraston, D. N. Bolam, H. J. Gilbert, G. J. Davies, Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781 (2004).
48. K. McIntosh, W. B. Becker, R. M. Chanock, Growth in suckling-mouse brain of “IBV-like” viruses from patients with upper respiratory tract disease. *Proc. Natl. Acad. Sci. U.S.A.* **58**, 2268–2273 (1967).
49. K. McIntosh, J. H. Dees, W. B. Becker, A. Z. Kapikian, R. M. Chanock, Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proc. Natl. Acad. Sci. U.S.A.* **57**, 933–940 (1967).
50. R. Dijkman *et al.*, Isolation and characterization of current human coronavirus strains in primary human epithelial cell cultures reveal differences in target cell tropism. *J. Virol.* **87**, 6081–6090 (2013).
51. M. Vignuzzi, J. K. Stone, J. J. Arnold, C. E. Cameron, R. Andino, Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).
52. A. V. Borderia *et al.*, Group selection and contribution of minority variants during virus adaptation determines virus fitness and phenotype. *PLoS Pathog.* **11**, e1004838 (2015).
53. K. S. Xue, K. A. Hooper, A. R. Olldard, A. S. Dingens, J. D. Bloom, Cooperation between distinct viral variants promotes growth of H3N2 influenza in cell culture. *eLife* **5**, e13974 (2016).
54. P. Domingo-Calap, E. Segredo-Otero, M. Durán-Moreno, R. Sanjuán, Social evolution of innate immunity evasion in a virus. *Nat. Microbiol.* **4**, 1006–1013 (2019).
55. A. T. Ciota, D. J. Ehrbar, G. A. Van Slyke, G. G. Willsey, L. D. Kramer, Cooperative interactions in the West Nile virus mutant swarm. *BMC Evol. Biol.* **12**, 58 (2012).
56. L. Cao *et al.*, Coexistence of hepatitis B virus quasispecies enhances viral replication and the ability to induce host antibody and cellular immune responses. *J. Virol.* **88**, 8656–8666 (2014).
57. Y. Shirogane, S. Watanabe, Y. Yanagi, Cooperation between different RNA virus genomes produces a new phenotype. *Nat. Commun.* **3**, 1235 (2012).
58. G. Herrler *et al.*, The receptor-destroying enzyme of influenza C virus is neuraminidase-O-acetylsterase. *EMBO J.* **4**, 1503–1506 (1985).
59. G. N. Rogers, G. Herrler, J. C. Paulson, H. D. Klenk, Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *J. Biol. Chem.* **261**, 5947–5951 (1986).
60. B. M. Hause *et al.*, Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. *PLoS Pathog.* **9**, e1003176 (2013).
61. R. Vlasak, M. Krystal, M. Nacht, P. Palese, The influenza C virus glycoprotein (HE) exhibits receptor-binding (hemagglutinin) and receptor-destroying (esterase) activities. *Virology* **160**, 419–425 (1987).
62. W. Luytjes, P. J. Bredendbeek, A. F. H. Noten, M. C. Horzinek, W. J. M. Spaan, Sequence of mouse hepatitis virus A59 mRNA 2: Indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166**, 415–422 (1988).
63. J. S. Long, B. Mistry, S. M. Haslam, W. S. Barclay, Host and viral determinants of influenza A virus species specificity. *Nat. Rev. Microbiol.* **17**, 67–81 (2019).
64. R. Xu *et al.*, Functional balance of the hemagglutinin and neuraminidase activities accompanies the emergence of the 2009 H1N1 influenza pandemic. *J. Virol.* **86**, 9221–9232 (2012).
65. L. J. Mitnaul *et al.*, Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *J. Virol.* **74**, 6015–6020 (2000).
66. R. Wagner, M. Matrosovich, H. D. Klenk, Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev. Med. Virol.* **12**, 159–166 (2002).
67. L. G. Baum, J. C. Paulson, The N2 neuraminidase of human influenza virus has acquired a substrate specificity complementary to the hemagglutinin receptor specificity. *Virology* **180**, 10–15 (1991).
68. M. Cohen *et al.*, Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Viol. J.* **10**, 321 (2013).
69. M. N. Matrosovich, T. Y. Matrosovich, T. Gray, N. A. Roberts, H.-D. Klenk, Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *J. Virol.* **78**, 12665–12667 (2004).
70. Y. Yang *et al.*, Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12516–12521 (2014).

71. J. L. McKimm-Breschkin *et al.*, Generation and characterization of variants of NWS/G70C influenza virus after in vitro passage in 4-amino-Neu5Ac2en and 4-guanidino-Neu5Ac2en. *Antimicrob. Agents Chemother.* **40**, 40–46 (1996).
72. M. T. Hughes, M. Matrosovich, M. E. Rodgers, M. McGregor, Y. Kawaoka, Influenza A viruses lacking sialidase activity can undergo multiple cycles of replication in cell culture, eggs, or mice. *J. Virol.* **74**, 5206–5212 (2000).
73. F. Wen, X. F. Wan, Influenza neuraminidase: Underrated role in receptor binding. *Trends Microbiol.* **27**, 477–479 (2019).
74. K. A. Hooper, J. D. Bloom, A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. *J. Virol.* **87**, 12531–12540 (2013).
75. R. G. Webster *et al.*, Antigenic structure and variation in an influenza virus N9 neuraminidase. *J. Virol.* **61**, 2910–2916 (1987).
76. J. N. Varghese *et al.*, Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11808–11812 (1997).
77. J. C. C. Lai *et al.*, A secondary sialic acid binding site on influenza virus neuraminidase: Fact or fiction? *Angew. Chem. Int. Ed. Engl.* **51**, 2221–2224 (2012).
78. M. Dai *et al.*, Mutation of the second sialic acid-binding site, resulting in reduced neuraminidase activity, preceded the emergence of H7N9 influenza A virus. *J. Virol.* **91**, e00049-17 (2017).
79. W. Du *et al.*, The 2nd sialic acid-binding site of influenza A virus neuraminidase is an important determinant of the hemagglutinin-neuraminidase-receptor balance. *PLoS Pathog.* **15**, e1007860 (2019).
80. B. J. Haijema, H. Volders, P. J. M. Rottier, Switching species tropism: An effective way to manipulate the feline coronavirus genome. *J. Virol.* **77**, 4528–4538 (2003).
81. W. Li *et al.*, Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8508–E8517 (2017).