Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Nidovirus Sialate-O-Acetyltransferases

EVOLUTION AND SUBSTRATE SPECIFICITY OF CORONAVIRAL AND TOROVIRAL RECEPTOR-DESTROYING ENZYMES*

Received for publication, August 23, 2004, and in revised form, October 26, 2004. Published, JBC Papers in Press, October 26, 2004.

This paper is available on line at http://www.jbc.org

Saskia L. Smits‡, Gerrit J. Gerwig§, Arno L. W. van Vliet‡, Arjen Lissenberg‡, Peter Briza, Johannis P. Kamerling§§, Reinhard Vlasak¶¶, and Raoul J. de Groot¶¶

From the ‡Virology Division, Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, and ¶¶Institute for Mucin Research, Academy Biomedical Center, Faculty of Veterinary Medicine, Utrecht University, 3584 CH Utrecht, The Netherlands, and ¶¶¶Department of Molecular Biology and Applied Biotechnology, Department of Cell Biology, University of Salzburg, A-5020 Salzburg, Austria

Many viruses achieve reversible attachment to sialic acid (Sia) by encoding envelope glycoproteins with receptor-binding and receptor-destroying activities. Toroviruses and group 2 coronaviruses bind to O-acetylated Sias, presumably via their spike proteins (S), whereas other glycoproteins, the hemagglutinin-esterases (HE), destroy Sia receptors by de-O-acetylation. Here, we present a comprehensive study of these enzymes. Sialate-9-O-acetylatedaminic acid, described for bovine and human coronaviruses, also occur in equine coronaviruses and in porcine toroviruses. Bovine toroviruses, however, express novel sialate-9-O-acetyltransferases, which prefer the di-O-acetylated substrate 5-N-acetyl-7(8),9-di-O-acetylneuraminic acid. Whereas most rodent coronaviruses express sialate-4-O-acetyltransferases, the HE of murine coronavirus DVIM cleaves 9-O-acetylated Sias. Under the premise that HE specificity reflects receptor usage, we propose that two types of Sias serve as initial attachment factors for coronaviruses in mice. There are striking parallels between orthomyxo- and nidovirus biology. Reminiscent of antigenic shifts in orthomyxoviruses, rodent coronaviruses exchanged S and HE sequences through recombination to extents not appreciated before. As for orthomyxovirus reassortants, the fitness of nidovirus recombinant offspring probably depends both on antigenic properties and on compatibility of receptor-binding and receptor-destroying activities.

The first and arguably most critical step of any viral infection is the attachment of the virion to the host cell. The molecules to which virions bind comprise a diverse collection of cell surface proteins, lipids, and carbohydrates (1). Quite a number of enteric and respiratory viruses use sialic acid (Sia) either as their sole receptor or as an initial attachment factor (2).

Sialic acids, a generic name for a large family of 9-carbon negatively charged monosaccharides, typically occur as terminal residues of glycoconjugates. Of the variety of Sias that occur in nature, many result from differential O-acetylation at C4, C7, C8, and C9 of the parental molecule 5-N-acetylneuraminic acid (Neu5Ac) (2, 3). Sias make ideal receptors because of their accessibility and abundance, but their wide distribution also has a downside. High avidity binding of virions to nonpermissive cells, to soluble sialylated compounds, or to sialylated mucins of the mucus barrier all might lead to loss of infectivity. In addition, loss of infectivity might result from virus self-aggregation. To ward against such hazards, some enveloped RNA viruses have developed an elegant strategy, which allows high affinity yet reversible binding to Sia receptors; they specify for virion-associated receptor-destroying enzymes (RDEs). So far, two types of viral RDEs have been identified: sialidases (neuraminidases) and sialate-O-acetyltransferases.

RDEs with sialidase activity are found in para- and orthomyxoviruses (4, 5). For example, the orthomyxoviruses, influenza virus A and B, express two envelope glycoproteins: the hemagglutinin (HA), which mediates binding to Neu5Ac and subsequent fusion of viral and host cell membranes, and the neuraminidase (NA), which catalyzes the removal of Neu5Ac from sialoglycoconjugates. A functional balance between the antagonistic HA and NA activities is critical for efficient entry (5).

The other type of RDE, sialate-O-acetyltransferase, was first discovered in influenza C virus (6, 7). This virus expresses only one envelope glycoprotein species, the hemagglutinin-esterase fusion protein (HEF), in which all three functions (receptor binding, RDE, and membrane fusion activity) are combined (7, 8). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY771998. The on-line version of this article (available at http://www.jbc.org) contains an additional figure. Supported by the Netherlands Digestive Disease Foundation Project WS98-41. Supported by the Academic Biomedical Center, Utrecht University (Expertise Center for Carbohydrate Analysis and Synthesis). Supported by the Austrian Science Fund, Project P14104-MED. ¶¶ Supported by the Expertise Center for Carbohydrate Analysis and Synthesis. © 2005 by The American Society for Biochemistry and Molecular Biology, Inc.

This is an Open Access article under the CC BY license.
The structure of the HEF spike, a homotrimer of a class I membrane glycoprotein, resembles that of HA of influenza virus A (9). Like the HA protein, the HEF protein is post-translationally cleaved. The N-terminal subunit, HEF1, forms the globular top part of the spike and contains the receptor-binding (R) and RDE esterase (E) domains; the C-terminal subunit HEF2, together with segments of HEF1, contains the fusion domain (F) and forms the large, membrane-anchored stalk (9). HEF specifically recognizes 5-N-acetyl-9-O-acetylneuraminic acid (Neu5,9Ac2) receptors, via a binding site at the tip of the globular head region. The catalytic pocket of the RDE constitutes a second Sia-binding site. The RDE, a serine hydrolase with a Ser-His-Asp catalytic triad, destroys Neu5,9Ac2 receptors by removing the 9-O-acetyl group (6, 10).

An RDE with sialate-4-O-acetyltylerase activity was recently discovered in infectious salmon anemia virus (genus Isavirus, family Orthomyxoviridae), a segmented negative-stranded RNA virus of teleosts (11). There is no obvious sequence identity between influenza virus HE and infectious salmon anemia virus hemagglutinin-esterase protein (HE), except for the regions that contain the catalytic residues (12). However, influenza C virus HEF1 does share considerable identity with the nidovirus HEs (i.e. class I envelope glycoproteins expressed by toroviruses and certain coronaviruses) (13–16).

Coronaviruses and toroviruses (separate genera in the family Coronaviridae, order Nidovirales) are evolutionarily related, enveloped (+)-strand RNA viruses of mammals and birds (17–19). All toroviruses identified so far carry an HE gene (20), but among coronaviruses, only those belonging to (sub)group 2 do. The latter are exemplified by mouse hepatitis virus (MHV) and bovine coronavirus (BCoV). In SARS-associated human coronavirus (HCoV-SARS), which is an early phylogeny split-off of group 2 (21), and in group 1 and 3 coronaviruses, represented by porcine transmissible gastroenteritis virus and avian infectious bronchitis virus, respectively, the HE gene is absent.

The role of the nidoviral HE proteins and their importance for viral replication is still subject to debate. In the case of BCoV, infection in the presence of esterase inhibitors results in reduced virus titers (22), whereas HE-specific antibodies neutralize the virus (23). Other observations, however, indicate that the HEs are nonessential for replication in tissue culture cells (24, 25). In fact, during propagation in vitro of murine coronaviruses, HE expression is readily lost (13, 26). Moreover, infection experiments with such HE-deficient MHV laboratory strains would suggest that, at least for these coronaviruses, the HE protein is also dispensable in vivo.

Entry of coronaviruses and toroviruses (i.e. attachment and subsequent fusion of viral and host cell membranes) is mediated by the spike envelope glycoprotein S (27, 28). MHV, HCoV-SARS, and group 1 coronaviruses, transmissible gastroenteritis virus, and HCoV-229E use specific cell surface glycoproteins as their main receptor (29–32), as is probably the case for the other coronaviruses and toroviruses. However, there is accumulating evidence that coronaviruses in addition use Sia receptors (33–39). Contrary to what might be expected from their similarity to influenza C virus HEF or from their name, the coronaviral HEs seem to play only a minor role in viral attachment to Sias. Rather, the spike proteins appear to be the major Sia receptor-binding proteins (35, 38, 39). Equine torovirus Berne, which lacks a functional HE gene, induces hemagglutination (40), suggesting that toroviruses utilize Sia as an attachment factor as well and again implicating the S protein as the main candidate for Sia receptor binding.

For all group 2 coronaviruses studied so far, the Sia receptor binding specificity of S matches the substrate preference of HE. For instance, BCoV and HCoV-OC43 specifically bind via their S proteins to glycosidically bound Neu5,9Ac moieties (35, 39) and, in accordance, encode HEs with sialate-9-O-acetylerase activity (22, 41). Likewise, MHV strain S binds to 5-N-acetyl-4-O-acetylneuraminic acid (Neu4,5Ac2) receptors apparently via its S protein and encodes an HE with sialate-4-O-acetylerase activity (38). Sialate-4-O-acetylerase activity is also present in the closely related rodent coronavirus strains MHV-JHM and rat sialodacryoadenitis coronavirus (RCoV-SDAV) (16, 38, 42, 43).

Toroviruses and the group 2 coronaviruses apparently acquired their HE genes from an unknown source via independent heterologous RNA recombination events (13–15). Conceivably, the acquisition of an HE protein increased viral fitness by allowing optimal usage of specific Sias as attachment factors. Still, our understanding of the evolution, function, and substrate specificities of the nidovirus HEs is incomplete. Both the toro- and the coronaviral HE genes have diverged considerably (20, 38). Coronavirus HEs have been described, which in their sequence are clearly distinct from those analyzed so far and for which the substrate specificity remains to be determined. With respect to the specificities of the toroviral HEs, no information is available at all. Here, we present a comprehensive study of nidovirus HE proteins based upon a combination of phylogenetic, biochemical, and molecular virological techniques. In addition to the sialate-9-O- and sialate-4-O-acetylerase described so far, we report the identification of a novel type of HE, which occurs in bovine toroviruses (BToVs) and displays a preference for Sias with a di-O-acetylated glycerol side chain.

We already noted that exchange of HE genes via homologous RNA recombination is common among toroviruses (20). We now show that naturally occurring rodent coronaviruses also exchange their HE as well as their S genes to an extent not appreciated before. At least on one occasion, this must have resulted not only in an antigenic shift but also in a shift in HE substrate specificity. The recombinational exchanges of surface glycoprotein genes in nidoviruses are reminiscent of the reassortment of HA and NA genome segments in influenza A and B viruses. Parallels between orthomyxo- and nidoviral biology with respect to attachment, envelope glycoprotein function, and consequences thereof are discussed.

MATERIALS AND METHODS

Phylogenetic Analyses—HE sequences were extracted from the NCBI data base for BToV strains Breda (BRV), B145 and B150 (accession numbers Y10866, AJ575379, and AJ575380), porcine torovirus (PToV) strains P-MAR, P10, and P4 (AJ575363, AJ575366, and AJ575364), MHV strains A59, JHM, and 2 (NC_001846, AAA46442, and AF201929), PuCoV (AJ005960), RCoV-SDAV (AF207551), equine coronavirus (ECoV)-NC99 (AY163600), BCoV-Mebus (U073753), HCoV-OC43 (OC43; M76375), PHEV (AY078417), and influenza C virus strains Miyagi/4/93 (C/MIY/4/93; BAB84727) and Johannesburg/66 (C/JHG/66; S07412). In MHV-A59 and MHV-2, the reading frames of the HE genes are interrupted by termination codons and, in the case of MHV-2, also interrupted by a 4-nucleotide deletion. To allow the inclusion of these strains in phylogenetic analyses, their HE genes were reconstructed in silico according to the MHV HE consensus sequence: MHV-A59, TGA (nucleotides 43–45) → GGA; MHV-2, TAA (nucleotides 292–294) → CAA; and insertion of the sequence AAC at nucleotide position 525.

Discrepancies were noted between sequences deposited for MHV-DVM in the NCBI data base, indicating that virus stocks have been exchanged. The DVM HE nucleotide sequence of Sugiyama et al. (AB008938) is 100% identical to the HE gene of MHV-S, as determined in our laboratory, but only 73% identical to the DVM HE sequence of
Table I  Oligonucleotide primers

| Oligonucleotide | Nucleotide sequence (5’–3’) | Polarity | Positions | Template |
|-----------------|----------------------------|----------|----------|----------|
| 995             | ACTATTATTTTGTGTAATG        | +        | –18–30   | MHV-S    |
| 996             | ACCGTTATCCCTAAACAGAAG      | +        | 23959–23979 | MHV-A59 |
| 1218            | CCGACACAGCGATGATGC         | +        | 1–26     | P-MAR    |
| 1240            | CTAAATAGTAAATTAAAAT        | +        | 1261–1284 | P-MAR, P10, P4 |
| 1245            | CCAGGACAAGCAGGTAATGGTCA    | +        | 27284–27299 | P10, P4 |
| 1394            | CAGCATGGTTAATGCTTGAA       | –        | 27318–27340 | B145, B150 |
| 1433            | AATTTTGGTAGGACATTTGAAC     | –        | 681–700   | BCoV     |
| 1454            | TGAGGTTGTTATCTAAGGGCTTA    | +        | 26773–26793 | B145, B150 |
| 1502            | TGTGTCAGATATCTCATCTAANAGGCAAC | +        | 221–240 | MHV-S    |
| 1503            | CGGAAATTCGCGAACGACTG       | –        | 121–140   | MHV-S    |
| 1509            | GCTGTACGTATGCTTGAAC         | +        | 127–155   | BCoV     |
| 1518            | TCGGATTACACAGGGCGAGATTAAGTGTCA | +        | 208–228  | BCoV     |
| 1555            | GTACCTATGTGGAAAATAC        | –        | 112–119   | BCoV     |
| 1556            | TTTTCCGGAATACACAAATTTGGCTGAA | +        | 193–210  | MHV-S    |
| 1560            | CCACCATGTTTGTGTTTCAATTGGTGTTTCT | +        | 1–22     | BCoV     |
| 1598            | GGCTAACAGCATCAGGCG         | +        | 1259–1278 | BCoV     |
| 2060            | GAATCCACCACTGTGGC          | –        | 1–18     | DVIM     |
| 2061            | TTATATATTATGTGAAAGTTATTTAATTTT | +        | 302–326  | DVIM     |
| 2063            | GATATCATGACCTGATCTAAT       | –        | 1276–1296 | DVIM     |
| 2078            | CTCTCCATGAATAATATAGAGGCCACCTCCAGT | –        | 283–319  | DVIM     |
| 2184            | TTGGATTCACACATGATTTTGTGCGAAAAT  | +        | 1–24     | ECoV     |
| 2185            | TTGTGCACTAAGCTACCTAGGCTATAGT | +        | 1252–1272 | ECoV     |

* a Position with respect to the EToV Berne or MHV-A59 genome.

Sia enriched for Neu5,4Ac, was purified from horse serum glycoproteins by acid hydrolysis (48). The mixture contained 3 mg of Sia/ml consisting of 69% Neu5Ac, 11% 5-N-glycolenuraminic acid, and 20% Neu4,5Ac, as determined by gas chromatography-mass spectrometry (see below). 5-N-Acetyl-4,9-di-O-acetylenuraminic acid α-methyl glycoside (αNeu4,5,9Ac2Me2) was synthesized by partial O-acetylation of Neu5Ac α-methyl glycoside (Sigma) as described previously (49).

Heterologous Expression of Nidovirus HE Proteins and Substrate Specificity Assays—OST-7–1 cells (~1 × 10⁶) were infected with recombinant vaccinia virus vTF7-3 and, at 1 h posttransfection, transfected with 2.5 μg of plasmid DNA as described (50). At 8 h postinfection, the cells were washed with phosphate-buffered saline and lysed in 200 μl of phosphate-buffered saline, 1% Triton X-100. Nuclei and cell debris were pelleted by centrifugation for 2 min at 14,000 rpm at 4 °C. Supernatants were tested for sialate-O-acetylenolase activity with p-nitrophenyl acetate (pNPA) (Sigma) as a substrate (42). Lysates from mock-transfected vTF7-3-infected cells were included as negative controls. The specific enzymatic activity per ml of lysate, corrected for endogenous acetylenolase activity present in mock-transfected cells, was calculated as described (7).

Substrate specificity assays were initially performed with equal amounts of specific enzymatic activity based upon hydrolysis of pNPA (7). However, as the activity of the acetylenolase toward pNPA proved of limited predictive value with respect to activity toward Sia derivatives, all assays were instead performed with fixed amounts of cell lysates.

Substrate specificity of nidovirus HE proteins was determined in 100 μl of reaction mixtures containing 40 μl of cell lysate and 0.5 mg/ml Neu5,9Ac, 0.3 mg/ml Neu4,5Ac, or 0.1 mg/ml Neu5,9Ac2Me in phosphate-buffered saline, 0.5% bovine serum albumin. Incubation was for 1 h at 37 °C. Reactions were stopped by snap freezing and lyophilization.

For assays with Neu5,9Ac, or Neu4,5Ac, substrates, Sias were derivatized with 1,2-diamino-4,5-methylenedioxybenzene for 2.5 h at 50 °C (51) and analyzed by fluorometric HPLC on a Cosmosil 5C18 ARIII column (Nacalai Tesque; particle size 5 μm; 4.6 × 250 mm) with isocratic elution using water/methanol/acetoniitrile (84.7:9:4.5; flow rate, 1 ml/min) and detection at excitation and emission wavelengths of 373 and 474 nm, respectively.

Assays with glycosidically bound Sias were performed essentially as described (42). Peak identities of the chromatograms were confirmed by MS analysis. Spectra were obtained on a Micromass Q-ToF Ultima Global mass spectrometer (Waters) with electrospray ionization and direct infusion. Dried samples were dissolved in 50% aqueous acetonitrile containing 1% NH₄OH and infused at a rate of 10 μl/min. Spectra were recorded in the negative ion mode for 5 min, using a scan time of 10.8-kb region comprising the 3’-end of ORF1b through the N gene, were cloned in pGEM-T Easy (Promega) and subjected to sequence analysis using the ABI Prism BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) and an ABI Prism 3100 genetic analyzer. The nucleotide sequence, determined in both orientations and for each region on at least two independent clones, was deposited in the EMBL data base (AY771998). Throughout, we use “DVIM” to refer to the Compton variant.

Amino acid alignments were generated with Multalin, employing the Dayhoff symbol comparison table, and refined manually. Gaps were excluded. Phylogenetic trees were created by the neighbor-joining method, implemented in the program MEGA version 2.1 (44); the reliability of each interior branch was tested by bootstrap analysis with 1000 resamplings.

The number of synonymous substitutions per synonymous site (Ks) was estimated by sliding window analysis for overlapping 500-nucleotide segments with a 250-nucleotide step size with the program K-Estimator (45, 46), and Ds profiles were generated as described (20).

Construction of HE Expression Vectors—Expression plasmid pBS-BRV-HE and reverse transcription-PCR cloning of the HE genes from torovirus strains B145, B150, P-MAR, P10, and P4 were described previously (47) and, for assays with Neu5,9Ac2 or Neu4,5Ac2 substrates, Sias were derivatized with 1,2-diamino-4,5-methylenedioxybenzene for 2.5 h at 50 °C (51) and analyzed by fluorometric HPLC on a Cosmosil 5C18 ARIII column (Nacalai Tesque; particle size 5 μm; 4.6 × 250 mm) with isocratic elution using water/methanol/acetoniitrile (84.7:9:4.5; flow rate, 1 ml/min) and detection at excitation and emission wavelengths of 373 and 474 nm, respectively.
2 s and 0.1-s interscan delay, combined, and centered. Only singly charged ion species were observed.

Enzymatic de-O-acetylation of αNeu4,5,9Ac2Me was analyzed by gas chromatography-mass spectrometry (52). Lyophilizates were dissolved in 40 μl of pyridine/hexamethyldisilazane/trimethylchlorosilane (5:1:1, v/v/v) and incubated for 45 min at room temperature to allow for trimethylsilylation of Sias. Samples (1 μl) were analyzed with a Fisons Instruments GC 8060/MD800 system (Interscience, Breda, The Netherlands) and an AT-1 column (30 m × 0.25 mm, Alltech, Breda, The Netherlands); the temperature program was 220 °C for 25 min, 6 °C/min to 300 °C for 10 min; the injector temperature was 220 °C; detection was done by electron impact MS with a mass range of 150–800 m/z (53).

Metabolic Labeling and Radioimmunoprecipitation Assay—Metabolic labeling of cells and subsequent radioimmunoprecipitation assay with rabbit-α-BToV BRV serum or with serum from a piglet naturally infected with PToV strain P-MAR (54) were performed as described (15, 55).

RESULTS AND DISCUSSION

Phylogenetic Relationships among Orthomyxo- and Nidovirus Sialate-O-acetyesterases—The HE proteins of toro- and coronaviruses and the influenza C virus HEF1 are evolutionary equidistant, sharing ~30% sequence identity. Coronavirus HEs form disulfide-bonded homodimers (56); the oligomeric state of the torovirus HEs is unknown. Segments corresponding to HEF1 domains E and R and to what might be vestigial remnants of domain F (9) can be distinguished in the nidovirus HEs. The nidoviral HEs are membrane-anchored via a C-terminal hydrophobic domain, which in the primary sequence of HEF would correspond to the fusion peptide at the N terminus of HEF2.

Most variation between and among orthoviral and nidoviral HEs occurs in domain R, and there is no obvious conservation of the HE receptor-binding site in the nidovirus proteins. In domain E, the catalytic Ser-His-Asp triad (Ser57, His355, Asp117) is strictly conserved. The HEF residues Gly85 and Asn117, of which the NH group and side chain, respectively, together with the NH group of Ser57 form the oxyanion hole (9), are maintained in most nidoviral HEs, but in those of the BToVs, Gly85 is substituted by Ser (see Supplemental Fig. 1).

In phylogenetic trees, the HE proteins of corona-, toro-, and influenza C viruses divide into three separate monophyletic clusters (Fig. 1; see Table II for abbreviations of virus names). This tree topology is consistent with the notion that the ancestors of torovirus, coronavirus, and influenza C virus acquired their HEs through independent heterologous recombination events (15) and that in each case the introduction of an HE gene into the viral genome occurred only once. Consequently, the coronavirus HE genes must all have evolved from one ancestral gene, which was acquired by a group II coronavirus predecessor after the split-off from HCoV-SARS. Similarly, the torovirus HEs all stem from a single ancestral O-acetyesterase.

Whereas sequence variation among influenza C virus HEFs is limited, the corona- and torovirus HEs have diverged considerably (up to 40%). Four distinct torovirus HE lineages can be distinguished, represented by the HEs of BToV-BRV, BToV-B145, PToV-P4, and PToV-P-MAR. A similar branching into four groups is seen for the coronavirus HEFs (Fig. 1) (20). Previously, only two lineages were recognized based upon phylogenetic relationships and substrate specificity, with those of lineage A (represented by the HE of MHV-S and RCoV-SDAV) exhibiting sialate-4-O-acetyesterase activity and those of lineage B (represented by BCoV and HCoV-OC43) displaying sialate-9-O-acetyesterase activity (38). Our current analysis shows that, in fact, two additional lineages, C and D, can be distinguished (represented by the HEs of MHV-DVIM and MHV-2 and that of ECoV strain NC99, respectively), for which the substrate specificity remains to be determined (see below).

To gain more insight into the evolution of the nidoviral O-acetyesterases, we studied the substrate specificities of the HEs of coronaviruses MHV-DVIM and ECoV and of representatives of each of the torovirus lineages. Since toroviruses, with Berne virus as the sole exception, cannot be propagated in tissue culture, the various HE proteins were produced in the vaccinia virus-based vTF7–3 expression system. Successful expression of each viral acetyesterase was confirmed by radioimmunoprecipitation as-

| Virus          | Strain/variant | Abbreviation |
|----------------|----------------|--------------|
| Coronavirus    |                |              |
| Bovine coronavirus | Mebus         | BCoV         |
| Equine coronavirus  | NC99          | PHEV         |
| Porcine hemagglutinating |  |              |
| encephalomyelitis virus |              |              |
| Human coronavirus   | OC43          | HCoV-OC43    |
| Rat sialodacryoadenitis coronavirus | | RCoV-SDAV |
| Puffinosis coronavirus | S            | PuCoV        |
| Mouse hepatitis virus | A59          | MHV-A59      |
|                  | JHM           | MHV-JHM      |
|                  | DVIM          | MHV-DVIM     |
|                  | 2             | MHV-2        |
| Torovirus        |                |              |
| Bovine torovirus  | Breda         | BToV-BRV     |
|                  | B145          | BToV-B145    |
|                  | B150          | BToV-B150    |
| Porcine torovirus | Markelo       | PToV-P-MAR   |
|                  | P10           | PToV-P10     |
|                  | P4            | PToV-P4      |
| Orthomyxo-ovirus | Miyagi/4/93   | C/Miy4/93    |
| influenza C virus | Johannesbur/66 | C/JHG/66    |

FIG. 1. Phylogenetic relationships among orthomyxo- and nidovirus HE proteins. An alignment of torovirus, coronavirus and influenza C virus HE sequences (see Supplemental Fig. 1) was used to generate a neighbor-joining tree. Confidence values calculated by bootstrapping (1000 replicates) are indicated at the major branching points. Branch lengths are drawn to scale; the scale bar represents 0.1 amino acid substitution/site. The four different coronavirus lineages are indicated (A–D).
Evolution and Substrate Specificities of Nidovirus HE Proteins

say (BToV strain BRV and B145 and PToV strain P-MAR) and/or by demonstrating enzymatic activity in cell lysates with pNPA as a substrate (Fig. 2 and Table III). Note that antisera against BToV-BRV also detected the HE of BToV-B145 but not that of BToV-B150 (Fig. 2A), illustrating that the two BToV HE lineages (Fig. 1) are antigenically distinct.

Substrate preferences of the nidoviral HEs were studied with mixtures of purified free mono- and di-O-acetylated Sias, enriched for Neu4,5Ac2 or Neu5,9Ac2. In agreement with published observations (22, 41, 42), vTF7-3-expressed HEs of MHV-S and BCoV-Mebus, taken along as model enzymes, preferentially removed 4-O- and 9-O-acetyl groups, respectively. When tested with αNeu4,5,9Ac2Me, the HEs of both PToV and BToV exclusively hydrolyzed the 9-O-acetyl group (Fig. 4 and Table III).

These experiments tentatively identified the toroviral HEs as sialate-9-O-acetylerases. We noted, however, a consistent difference between BToV and PToV enzymes; although the HEs of BToV strains BRV and B145 and those of PToV strains P-MAR and P10 hydrolyzed the synthetic substrate pNPA to similar extents, the BToV enzymes were reproducibly less active toward free Sia substrates (Table III). This raised the question of whether a difference in substrate specificity between the PToV and BToV O-acetylerases might yet exist. We therefore compared the reactivity of these enzymes toward bovine submaxillary mucin (BSM), a model substrate, which contains a mixture of glycosidically bound Sias (Table IV). The HE of PToV P-MAR hydrolyzed 9-O-acetyl-5-N-glycolyneuraminic acid, Neu5,9Ac2, Neu5,8Ac2, and possibly also Neu5,7Ac2 but showed little reactivity toward 5-N-acetyl-7(8,9-di-O-acetylneuraminic acid (Neu5,7,8,9Ac4) (Fig. 5A and data not shown). In contrast, the HE of BToV-BRV exhibited only modest activity toward mono-O-acetylated Sias but converted Neu5,7,8,9Ac4 very efficiently (Fig. 5A), apparently to Neu5,7Ac2 and Neu5Ac (Fig. 5B). Similar observations were made for the HEs of BToV strains B145 (not shown) and B150 (Fig. 5A). None of the enzymes showed any reactivity when tested with horse serum glycoproteins, a model substrate containing glycosidically bound Neu4,5Ac2 (not shown). Our findings suggest that the PToV and BToV HEs indeed differ in substrate specificity, displaying distinct preferences for Sia substrates with either a mono-O- or di-O-acetylated glycerol side chain, respectively.

Coronavirus HEs: The HE Protein of ECoV-NC99 Is a Sialate-9-O-acetylerase—The HE proteins of MHV strains DVIM-2 (lineage C) and ECoV-NC99 (lineage D) are distinct from each other and from those of the coronavirus HE lineages A and B (Fig. 1). The HE of ECoV is most closely related to the sialate-9-O-acetylerases of the BCoV/HCoV-OC43 cluster, but since 4-O-acetylated Sias are prominently expressed in the horse (2), we entertained the possibility of ECoV HE being a sialate-9-O-acetylerase. However, in cleavage assays with the free Sia substrates Neu4,5Ac2, Neu5,9Ac2, and αNeu4,5,9Ac2Me, ECoV HE was indistinguishable from BCoV HE and preferentially cleaved at C9 (Figs. 3 and 4; Table III). Also, when tested with glycosidically bound Sias of BSM, ECoV HE preferentially de-O-acetylated Neu5,9Ac2 (not shown). These observations identify ECoV-NC99 HE as a classical sialate-9-O-acetylerase.

Coronavirus HEs: The HE Protein MHV-DVIM Is Not a Sialate-4-O-acetylerase—We anticipated MHV-DVIM, like all other rodent coronaviruses studied so far, to encode a
sialate-4-O-acetylesterase. However, in contrast to MHV-S HE, DVIM HE did not cleave the 4-O-acetyl groups from the free Sia substrates Neu4,5Ac2 and Neu4,5,9Ac32Me (Figs. 3 and 4) nor from glycosidically bound Neu4,5Ac2 present in horse serum glycoproteins (not shown). Instead, DVIM HE hydrolyzed the 9-O-acetyl moieties from free Neu5,9Ac2 and from αNeu4,5,9Ac2,2Me. Puzzlingly, the enzyme showed little reactivity toward any of the glycosidically bound Sias of BSM (Fig. 5). Substrate specificity of DVIM HE might be restricted to a subset of 9-O-acetylated Sias not present in BSM (e.g. efficient cleavage might require additional Sia modifications or might be linkage-dependent). Recently, we have found that treatment of chicken erythrocytes with DVIM HE destroyed BCoV receptors (i.e. Neu5,9Ac2), thus preventing BCoV-mediated hemagglutination; conversely, treatment of rat erythrocytes with BCoV HE abrogated agglutination by MHV-DVIM.4 The combined observations strongly suggest that MHV-DVIM binds to (a subset of) 9-O-acetylated Sia receptors and that its HE is a sialate-9-O-acetylesterase. We conclude that murine coronaviruses occur in the field in two types, which differ with respect to Sia receptor usage and HE substrate specificity.

Evidence for Horizontal Transfer of HE Genes among Coronavirus—It is important to stress that the phylogenetic relationships in Fig. 1 reflect the evolutionary history of the HE modules and not necessarily that of the viruses from which these genes were derived, since nidoviruses frequently engage in homologous RNA recombination (57). Indeed, there is evidence for recurrent horizontal transfer of HE genes among toroviruses; ancestors of bovine and porcine field strains apparently replaced their original HE sequences via intertypic RNA recombination with hitherto unidentified toroviruses,4 A. L. W. van Vliet, M. Langereis, and R. J. de Groot, unpublished observations.
thus giving rise to the BToV B150 and (presumably) the PToV P-MAR/P10 lineages (20). The fact that in the coronavirus branch, the HEs of the MHV strains DVIM and MHV-2 cluster separately from those of other MHV strains (JHM, S, and A59) and of RCoV-SDAV (Fig. 1) prompted us to consider the possibility that, also during coronavirus evolution, an exchange of HE genes had occurred.

As the sequence data available for MHV-DVIM are incomplete and conflicting (also see “Materials and Methods”), we determined the nucleotide sequence for the 3' one-third of the genome for the DVIM stock obtained from the Compton laboratory. Phylogenetic trees generated per gene revealed incongruencies, indicative for recombination (Fig. 6A). In the HE gene, MHV strains DVIM and MHV-2 grouped separately from the other rodent coronavirus strains (A59, JHM, and RCoV-SDAV). In the S gene, MHV-DVIM and MHV-2 were also distinct from strains A59 and JHM but now grouped together with RCoV-SDAV. Finally, in the surrounding genes (ORF1b, ns2, E, M, and N), all rodent coronaviruses were closely related and in essence formed one single cluster.

Further support for recombination was obtained by analysis of the distribution of synonymous substitutions (\(D_{\alpha_s}\); Fig. 6B). \(D_{\alpha_s}\) profiles based upon pairwise comparison of MHV-DVIM and -2 with JHM and A59 showed that within the HE and S genes, \(K_s\) values (i.e., the estimated number of synonymous substitutions per synonymous site) (46) are \(\geq\)6-fold higher than in the flanking genes. Comparison of RCoV-SDAV versus MHV-A59 yielded roughly constant \(K_s\) values across the ns2, HE, E, M, and N genes, but in the S gene, the number of synonymous substitutions per synonymous site was 5-fold higher. Most interestingly, pairwise comparison of MHV-DVIM and MHV-2 with RCoV-SDAV revealed that these three viruses are very closely related. RCoV-SDAV only differs from these MHV strains in the HE gene as indicated by a 5-fold increase in the \(K_s\) values. Surprisingly, however, MHV-DVIM and MHV-2 also seem to be recombinant with respect to each other as indicated by an abrupt 3-fold rise in \(K_s\) values in the HE gene.

The combined observations indicate that the divergence of the rodent coronaviruses was indeed accompanied by multiple recombination events during which both HE and S sequences were replaced. Apparently, related strains as well as hitherto unidentified coronaviruses served as donors. Our current data do not allow conclusions as to which of the MHV strains represent parental types and which represent recombinant progeny. The exchange among rodent coronaviruses of S and HE genes is highly reminiscent of the reassortment of HA and NA genome segments, which gives rise to antigenic shifts in influenza A viruses. We propose that, in analogy, the observed murine coronavirus diversity, brought about by RNA recombination, is a direct consequence of immune pressure for antigenic variation. In the case of the MHV HE proteins, the genetic exchanges would, however, not only have resulted in an antigenic shift, but, at least on one occasion, also in a shift in substrate preference from 4-O- to 9-O-acetylated Sias or vice versa.

The Evolution of Nidoviral Acetyesterases: A Complex Tale of Divergence and Recombination—In comparison with other nidovirus structural proteins, the HEs have received little attention, probably because they are dispensable for replication in vitro and, in the case of the coronaviruses, do not belong to the standard protein repertoire. Yet, the mere fact that in the course of evolution, two groups of nidoviruses independently acquired an HE gene indicates that these virion-associated enzymes must provide a selective advantage during the natural infection (15). Our current observations strongly support this view: under field conditions, HE genes may be exchanged via homologous RNA recombination both among coronaviruses and among toroviruses, but HE expression per se is strictly maintained. This is in stark contrast with the loss of HE expression, which occurs during virus propagation in tissue culture cells (13, 15, 26).

The evolution of the coronavirus and torovirus HE proteins has yielded a variety of sialate-O-acetyesterases. Current diversity seems to have arisen not only from antigenic drift but also as a result from adaptations to new hosts and/or new target tissues. Under the assumption that HE substrate preference reflects Sia receptor usage, nidoviruses apparently employ Neu5,9Ac2, Neu4,5Ac2, and Neu5,7(8)9Ac2 as initial attachment factors. As first reported here, HEs with a preference for the latter Sia subtype occur in bovine toroviruses, both in the BRV/B145 and B150 lineages. The latter observation is of particular interest, since the B150-type BToVs are recombinant viruses, which apparently originated from an exchange of HE sequences between...
dendrogram within the rodent coronavirus clade. The number of synonymous substitutions per synonymous site (Ks) values against nucleotide position. As a reference, the genes for 1b, ns2, HE, S, E, M, and N drawn to scale, are depicted as boxes (top). C, simplified schematic representation of the recombinant genome composition of rodent coronaviruses. Genes are represented as boxes. Related S and HE genes are color-coded. The genome structure indicated by a question mark has not yet been identified in the field.

**FIG. 6.** Evidence for multiple exchanges of S and HE sequences during the radiation of rodent coronaviruses. A, rooted neighbor-joining trees depicting the phylogenetic relationships among rodent coronavirus strains. Trees were constructed for the 3’-most 1476 residues of ORF1b and for the ns2, HE, S, E, M, and N genes with BCoV serving as an outgroup. Confidence values are indicated at the major branching points. Branch lengths are drawn to scale; the scale bar represents 0.05 nucleotide substitution per site. B, distribution of synonymous substitutions (Dsyn) in the 1b, ns2, HE, S, E, M, and N genes of rodent coronavirus variants. Sequences were compared pairwise by sliding window analysis; the number of synonymous substitutions per synonymous site (Ks) was estimated for overlapping 500-nucleotide gene segments with a 250-nucleotide step size. Dsyn profiles were gene-boxed color-coded. The
Evolution and Substrate Specificities of Nidovirus HE Proteins

...previous studies showed that the HEs may serve as novel tools to...ative analysis of nidovirus sialate-O-acetylemesterases, which should aim at the elucidation of the molecular basis for substrate specificity and explore the potential of HE-directed antiviral drugs. Moreover, the HEs may serve as novel tools to study sialic acid biology.

Acknowledgments—We thank Ulrike Vilas for excellent technical assistance and Jolanda de Groot-Mijnes and Peter Rottier for critical reading of the manuscript.

REFERENCES

1. Smith, A. R., and Helenius, A. (2004) Science 304, 237–242
2. Angata, T., and Varaki, A. (2002) Chem. Rev. 102, 439–469
3. Schauer, R. (2004) Zoology 107, 49–64
4. Cavanagh, D. (1997) J. Gen. Virol. 78, 839–846
5. Wagner, R., Matrosovich, M., and Klenk, H. D. (2002) Rev. Med. Virol. 12, 159–166
6. Herrler, G., Rott, R., Klenk, H. D., Muller, H. P., Shukla, A. K., and Schauer, R. (1985) EMBO J. 4, 1503–1506
7. Vlasak, R., Krystal, M., Nacht, M., and Palese, P. (1999) Virology 269, 419–425
8. Herrler, G., Durkop, I., Becht, H., and Klenk, H. D. (1988) J. Gen. Virol. 69, 839–846
9. Rosenthal, P. B., Biehl, S. H., Meier-Ewert, H., Skehel, J. J., and Wiley, D. C. (1998) Nature 396, 92–96
10. Vlasak, R., Muster, T., Lauro, A. M., Powers, J. C., and Palese, P. (1989) J. Virol. 63, 2056–2062
11. Hellebre, A., Vilas, U., Falk, K., and Vlasak, R. (2004) J. Virol. 78, 3055–3062
12. Falk, K.,ronov, S., Dubois-Dalcq, M., Rottier, P., Rossen, J., and Holmes, K. V. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5533–5536
13. Herrler, G., Rott, R., Klenk, H. D., Muller, H. P., Shukla, A. K., and Schauer, R. (1985) EMBO J. 4, 1503–1506
14. Snijder, E. J., den Boon, J. A., Horzinek, M. C., and Spaan, W. J. (1989) J. Virol. 63, 5277–5286
15. Klausegger, A., Strobl, B., Regl, G., Kaser, A., Iwersen, M., Schmid, H., Kohla, G., Strobl, B., Vilas, U., Schauer, R., and Vlasak, R. (1999) J. Virol. 73, 4721–4727
16. Schauer, R., Bredenbeek, P. J., Dobbe, J. C., Thiell, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J., and Gorbunova, A. E. (2003) Virology 311, 991–1004