Viral destruction of cell surface receptors

Andrew D. Mesecar* and Kiira Ratia

Center for Pharmaceutical Biotechnology and Department of Medicinal Chemistry and Pharmacognosy, University of Illinois, Chicago, IL 60607

Viral infection is initiated by the attachment of the virus to the appropriate host cells. This process involves a series of dedicated virion proteins that have evolved to specifically recognize one, or a small number, of cell-surface molecules. Although a number of virus–host attachment mechanisms involve direct protein–protein interactions, carbohydrate molecules such as sialic acids (SAs) may also serve as receptor-binding determinants. The binding of viral envelope glycoproteins to carbohydrates on cell membranes plays a significant role in infection by many viruses. In general, the glycoproteins of several lipid-enveloped viruses, including orthmyxoviruses (influenza A, B, and C), toroviruses, and coronaviruses, have three important functions: to recognize the receptor on the cell surface, to mediate viral fusion with the cell membrane, and to destroy the receptor. In the highly infectious influenza A and B viruses, the receptor-binding and membrane-fusion activities of cell entry are carried out by the glycoprotein hemagglutinin (HA) (Fig. 1a). The receptor-destroying enzyme (RDE) activity important for virus release is conducted by the glycoprotein/enzyme neuraminidase (NA). In influenza C virus, a single glycoprotein, the hemagglutinin-esterase-fusion (HEF) protein, possesses all three functions. For a number of toroviruses and group 2a coronaviruses, the glycoprotein hemagglutinin esterase (HE) has both receptor-destroying and receptor-binding activities. However, the receptor-binding activity of HE is considered accessory to that of the spike protein (S), a receptor-binding and fusion protein (Fig. 1*b*). Our understanding of the structure, mechanism, and evolution of HA, HEF, NA, and S at the molecular level has increased substantially over the past two decades because of the availability of numerous x-ray structural models of these molecules in the unliganded or receptor-bound complexes. In contrast, a lack of detailed structural studies on HE has hindered our understanding of its function and evolution. In this issue of PNAS, Zeng *et al.* (1) report the x-ray structures of HE from bovine coronavirus (BCoV) in both its unliganded and liganded forms. These structures and the associated biochemical data reported provide us with new insight and clues into the evolutionary relationships among corona-, toro-, and influenza viruses.

Corona- and toroviruses (family *Coronaviridae*, order *Nidovirales*) are enveloped, positive-stranded RNA viruses that have been implicated in many different respiratory and enteric diseases. Toroviruses are known to cause mild enteric infections in animals such as swine and cattle, and possibly humans. In contrast, coronaviruses, including the human coronavirus that causes severe-acute respiratory syndrome (SARS-CoV), are now recognized as important human and animal pathogens. Despite their potential clinical and veterinary relevance, little is known about molecular details of the viral entry process of these pathogens.

The involvement of the spike (S) glycoprotein in host-cell recognition and virus–host membrane fusion is well documented (reviewed in ref. 2). Spike is a trimeric protein (Fig. 1*b*) that specifically recognizes different cell-surface receptor glycoproteins depending on the specific coronavirus involved. For example, some coronaviruses of group 1 are capable of recognizing aminopeptidase N (APN), whereas others, along with members of the group 2b coronaviruses, bind to angiotensin-converting enzyme (ACE2). The group 2a coronaviruses, including BCoV, human coronavirus strain OC43 (HCoV-OC43), and mouse hepatitis virus DVIM (MHV-DVIM), bind to the carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1). A large number of the coronaviruses in group 2a also express a second surface glycoprotein, HE, that functions as both a hemagglutination protein and acetylesterase (Fig. 1*a*). Depending on the coronavirus, HE binds to specific SAs, causing hemagglutination, and its acetyltransferase activity cleaves off specific *O*-acetyl groups from these SAs on either glycoproteins or glycolipids. For example, BCoV, MHV-DVIM, and HCoV-OC43 all have sialic acid-*O*-acetylesterase activity, whereas mouse hepatitis virus S (MHV-S) and infectious salmon anemia virus have sialic acid-4-*O*-acetylesterase activity (Fig. 1*c*) (reviewed in ref. 3).

Zeng *et al.* (1) first demonstrated that their recombinant version of BCoV has sialic acid-9-*O*-acetyltransferase activity on synthetic substrates and that it is capable of destroying BCoV receptors on rat erythrocytes. Armed with this information, they crystallized the wild-type HE protein and determined the x-ray structure. Next, they generated a catalytically inactive form of HE by replacing the active site serine with an alanine residue, and then they soaked crystals of this HE mutant with the receptor-ligand analog Neu4,5,9Ac32Me. These x-ray structures yield a number of surprising observations that were not predicted on the basis of the structures of the related HA and HEF proteins.

From the structures it is evident that HE is not a trimeric protein as has been predicted on the basis of the trimeric structures for HEF and HA. Rather, HE is a dimeric protein with unique subunit contacts (Fig. 1*d*). Dimerization of HE is achieved by interaction of the receptor domains (R) and the membrane-proximal domains (MP). The MPs of HE are formed from a remnant portion of the F1 fusion domain of HEF and HA, and the F2 domain. Contacts between the enzyme domains (E) were not observed (Fig. 1*d*). For HEF and HA, trimerization results from significant interactions between the entire F domain inclusive of F1, F2, F3, and MP (Fig. 1*e*) (4–6). The trimeric coiled-coil fusion domain for HEF and HA is necessary for fusion of influenza A, B, and C viruses to the host cell. However, because the fusion of corona- and toroviruses is mediated through spike, there has been no evolutionary pressure on HE to retain all of the F domain. Instead, HE retained only a small portion of F1, all of F2, and none of F3 (Fig. 1*a*). In addition, the fusion peptide domain (FP) evolved into a transmembrane domain (TM) to anchor HE to the virion. During the course of this evolution, the MP domain was used to form stable dimers.

A second and significant observation is that Neu4,5,9Ac$_3$2Me binds to the receptor binding site (R), as would be expected for this class of glycoproteins. However, within the site, the ligand is

Author contributions: A.D.M. and K.R. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 9065.

*To whom correspondence should be addressed at: Center for Pharmaceutical Biotechnology, University of Illinois, 900 South Ashland Avenue M/C 870, Chicago, IL 60607. E-mail: mesecar@uic.edu.

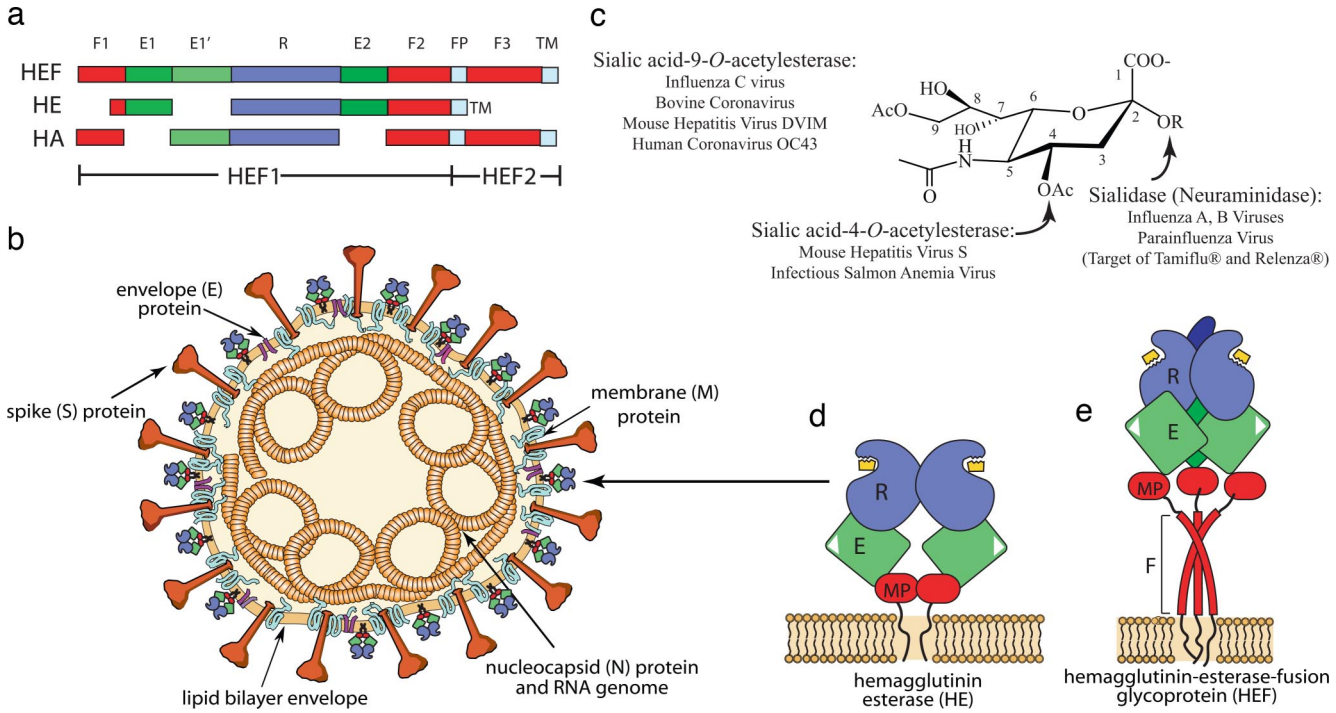

Fig. 1. Comparison of the structural and functional activities of HE, HEF, and HA. (*a*) Linear order of the sequence segments in HEF, HE, and HA color-coded by domains. Red segments, F1, F2, and F3; green, E1 and E2; lighter green, E′; blue, R; light blue, fusion peptide (FP) or transmembrane domain (TM). The HEF subunit 1 (HEF1) and subunit 2 (HEF2) are also indicated. HEF2 is absent from HE. This figure is adapted from refs. 1, 3, and 5. (*b*) Illustration of a group 2a coronavirus with the indicated virion proteins. Individual hemagglutinin molecules (HE) are color coded as described for *a*. (*c*) Substrate specificity of viral HE, HEN, and NA toward SAs. Figure is adapted from ref. 3. (*d*) Schematic illustration of the HE dimer. Colors are as described for *a*. The membrane-proximal domain (MP) is also shown. The catalytic triad is represented in the E domain as a white triangle. Binding of SA to the R domain is shown as a yellow box. The TM domain is shown embedded in the virion envelope. (*e*) Schematic illustration of the HEF trimer. The F domain comprising F1, F2, F3, and FP that forms a coiled-coil is indicated. Other features are as described for *d*.

bound in a unique conformation that is rotated ≈150° from similar ligands bound to HEF (5). The structure of HE surrounding the Neu4,5,9Ac$_3$2Me binding site is significantly different from that of HEF as a result of substantial differences in loop sizes and conformations. The plasticity of the receptor binding site and different orientation of the ligand are surprising because both HEF and HE evolved to bind 9-*O*-acetylated SAs. It would be expected on the basis of binding identical or similar ligands that these binding sites would have evolved to have similar structures. In contrast, because HEF and HA recognize different receptor-ligands (Neu5,9Ac$_2$ versus α2,3- or α2,6-linked Neu5Ac), it would be predicted that these proteins would have significantly different structures. However, Zeng *et al.* (1) have shown that HEF and HA have more closely related receptor binding sites than do HEF and HE even though they have more evolutionary distance (1).

In contrast to the receptor binding site, the active site architecture of enzyme domain E is very similar in HE and HEF. Superposition of the catalytic sites [see figure 4*a* of Zeng *et al.* (1)] indicates that the orientations of the catalytic triad residues are identical, within experimental error, as are other surrounding residues. The more canonical active site structure of E compared with R would be expected for enzymatic function because the geometric requirements for catalysis are more stringent than those for binding (7).

In summary, the structures of HA, HEF, and now HE should serve as a basis for furthering our understanding of how viruses recognize their receptors on the cell surface, mediate viral fusion with the cell membrane, and then destroy their receptors. A number of different respiratory and enteric viruses have evolved virion-associated receptor-destroying enzymes (HEs or HEFs) as part of an elaborate system to help them avoid irreversible binding to "decoy receptors" or to the already infected host cell, processes that would ultimately result in an overall loss of viral infectivity. With numerous emerging infectious diseases on the horizon, the work by Zeng *et al.* (1) should help to advance the development of new antiviral compounds that will help to treat these diseases.

1. Zeng Q, Langereis MA, van Vliet ALW, Huizinga EG, de Groot RG (2008) Structure of coronavirus hemagglutinin-esterase offers insight into corona- and influenza virus evolution. *Proc Natl Acad Sci USA* 105:9065–9069.
2. Wentworth DE, Holmes KV (2007) Coronavirus binding and entry, *Coronaviruses: Molecular and Cellular Biology*, ed Thiel V (Caister Academic, Norfolk, UK), pp 3–31.
3. de Groot RJ (2006) Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj J* 23:59–72.
4. Ha Y, Stevens DJ, Skehel JJ, Wiley DC (2002) H5 avian and H9 swine influenza virus haemagglutinin structures: Possible origin of influenza subtypes. *EMBO J* 21:865–875.
5. Rosenthal PB, *et al.* (1998) Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* 396:92–96.
6. Wilson IA, Skehel JJ, Wiley DC (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289:366–373.
7. Mesecar AD, Stoddard BL, Koshland DE, Jr (1997) Orbital steering in the catalytic power of enzymes: Small structural changes with large catalytic consequences. *Science* 277:202–206.