

The Receptor Binding Domain of the New Middle East Respiratory Syndrome Coronavirus Maps to a 231-Residue Region in the Spike Protein That Efficiently Elicits Neutralizing Antibodies

Huihui Mou,^a V. Stalin Raj,^b Frank J. M. van Kuppeveld,^a Peter J. M. Rottier,^a Bart L. Haagmans,^b Berend Jan Bosch^a

Department of Infectious Diseases and Immunology, Virology Division, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands^a; Department of Viroscience, Erasmus Medical Center, Rotterdam, The Netherlands^b

The spike (S) protein of the recently emerged human Middle East respiratory syndrome coronavirus (MERS-CoV) mediates infection by binding to the cellular receptor dipeptidyl peptidase 4 (DPP4). Here we mapped the receptor binding domain in the S protein to a 231-amino-acid fragment (residues 358 to 588) by evaluating the interaction of spike truncation variants with receptor-expressing cells and soluble DPP4. Antibodies to this domain—much less so those to the preceding N-terminal region—efficiently neutralize MERS-CoV infection.

Just 10 years following the outbreak of the severe respiratory acute syndrome coronavirus (SARS-CoV), the world is confronted with yet another deadly human coronavirus. The virus, first provisionally called human coronavirus-EMC (hCoV-EMC) (1, 2) but now named Middle East respiratory syndrome coronavirus (MERS-CoV), referring to its emergence in the Middle East and to the respiratory syndrome it causes, belongs to *Betacoronavirus* genus lineage 2c (3). As of 7 June 2013, 55 cases have been laboratory confirmed, including 31 deaths, all from—or linked to—the Arabian Peninsula (http://www.who.int/csr/don/2013_06_07/en/index.html). As with SARS-CoV, patients affected by

MERS-CoV suffer from severe and often lethal lower respiratory tract infection. The epidemiology of MERS-CoV is still enigmatic,

Received 10 May 2013 Accepted 11 June 2013

Published ahead of print 19 June 2013

Address correspondence to Berend Jan Bosch, bj.bosch@uu.nl.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.01277-13

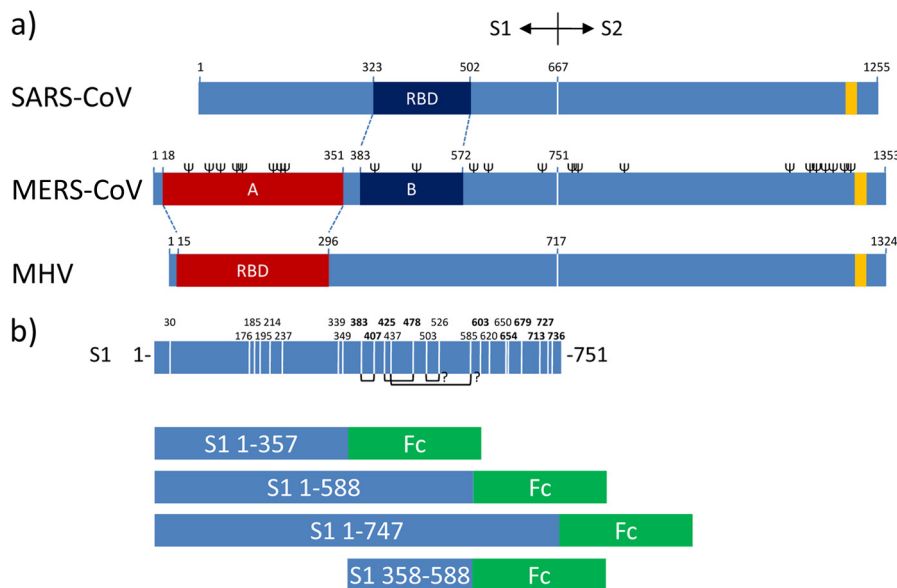
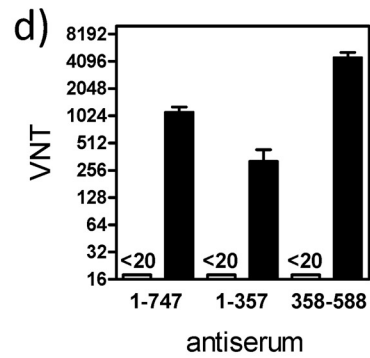
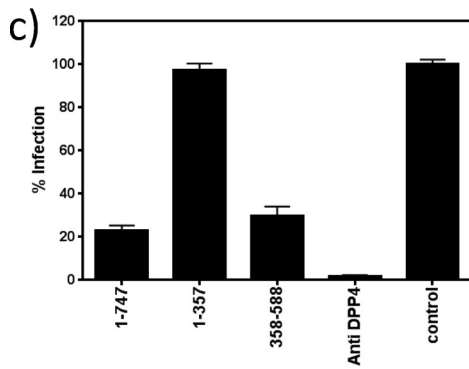
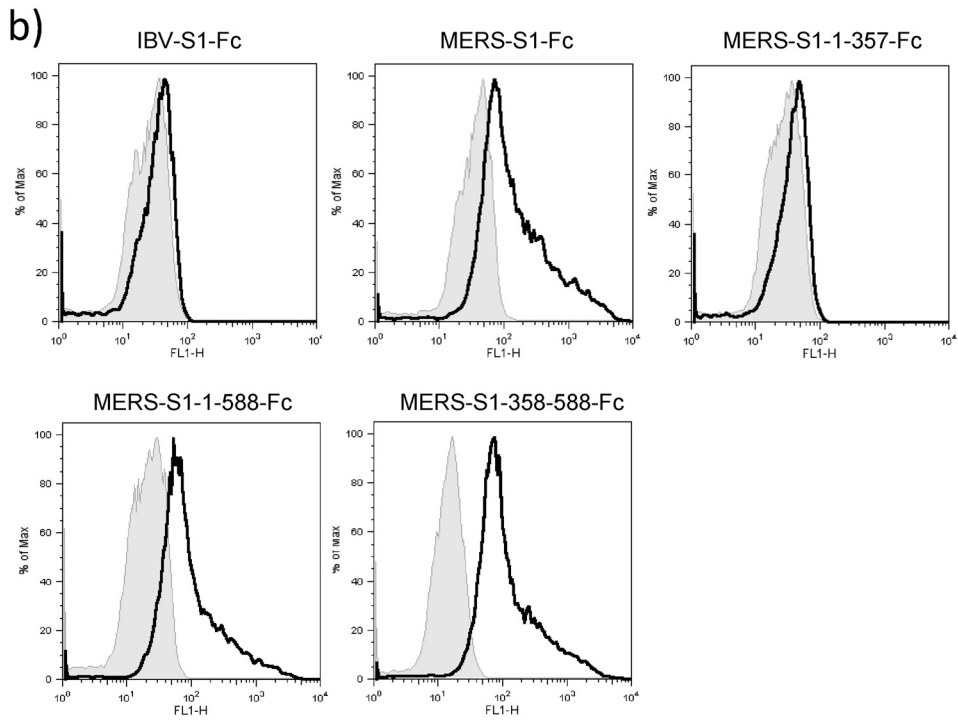
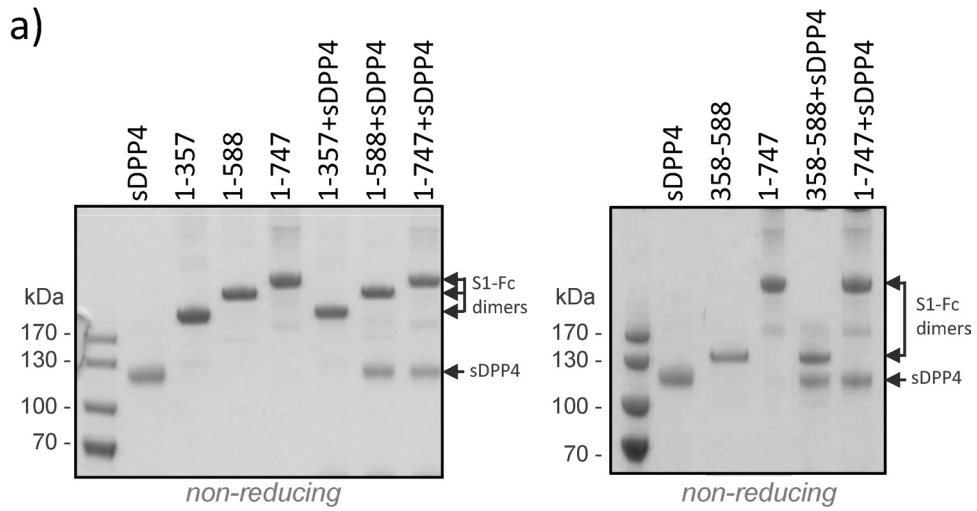


FIG 1 RBDs in betacoronavirus spike proteins and S1-Fc expression constructs. (a) Schematic representation of the betacoronavirus SARS-CoV, MERS-CoV, and MHV (strain A59) spike (S) protein sequences (drawn to scale) aligned at the S1-S2 junction. The known RBD in the S1 subunit of the MHV and SARS-CoV S proteins and their corresponding homologous regions in MERS-CoV S as defined by ClustalW alignment are indicated. The positions of the transmembrane domains (yellow bars; predicted by the TMHMM server) and of the predicted N-glycosylation sites (Ψ ; predicted by the NetNGlyc server, shown only for MERS-CoV S) are indicated. The border between the S1 and S2 subunits of the S protein is represented by a vertical white line. (b, top) Schematic representation of the MERS-CoV S1 subunit (residues 1 to 751) sequence. Cysteine positions in the S1 subunit are indicated by vertical white lines with the corresponding amino acid positions at the top. The positions of cysteines highly conserved among betacoronavirus S1 proteins are in bold. The predicted disulfide bond connections inferred from the structure of the SARS-CoV RBD are represented as connecting black lines at the bottom. (b, bottom) Domains of the MERS-CoV S1 subunit expressed as Fc chimeras.



but the geographic distribution of epidemiologically unlinked individuals points to intermittent zoonotic transmission from a so-far-unknown animal source, whereas a number of reported clusters indicate limited human-to-human spread (4).

The main determinant of coronavirus tropism is the viral spike (S) protein, as it mediates binding to a cell surface receptor. The MERS-CoV S protein, a 1,353-amino-acid type I membrane glycoprotein, assembles into trimers that constitute the spikes or peplomers on the surface of the enveloped coronavirus particle. The protein combines the two essential entry functions, namely, those of host receptor binding and membrane fusion, which are attributed to the N-terminal (S1, residues 1 to 751) and C-terminal (S2, residues 752 to 1353) halves of the S protein, respectively (Fig. 1a). Recently, we have identified dipeptidyl peptidase 4 (DPP4; also known as CD26), expressed in the human lung, as a functional receptor for MERS-CoV (5). Importantly, MERS-CoV can also use the evolutionarily conserved DPP4 protein of other species, most notably that of bats (5, 6).

Coronaviruses bind to receptors via independently folded, generally about 150- to 300-residue-long receptor binding domains (RBDs) present in their S1 subunit, the location of which within S1 can vary (7–9). Thus, for the *Betacoronavirus* mouse hepatitis virus (MHV), the binding to its carcinoembryonic antigen-related cellular adhesion molecule (CEACAM) receptor (10) has been mapped to the N-terminal ~300 amino acids of the spike protein (11, 12), whereas for the SARS-CoV, which is of the same genus, binding to the angiotensin-converting enzyme 2 (ACE2) receptor (13) maps to residues 323 to 502 of S1 (14, 15) (Fig. 1a). Identification of the RBD can hence help in the development of monoclonal antibodies and vaccines for the treatment and prevention of infection. The RBD is the most important target of neutralizing antibodies (11, 16, 17), preventing virus-receptor interaction.

We previously used the S1 domain of MERS-CoV fused to the Fc region of human IgG to demonstrate the interaction of S1 with DPP4-expressing cells and with soluble, i.e., non-membrane-anchored, DPP4 (5). To identify the RBD in the MERS-CoV S1 subunit, we generated S1-Fc protein chimeras with truncations at the C and N termini of the S1 domain. We considered a three-domain structure of the MERS-CoV S1 protein (residues 1 to 357, 358 to 588, and 589 to 747) based on the predicted location and structure of the RBD of two other *Betacoronaviruses*, MHV and SARS-CoV (11, 12, 14, 15), of which the homologous regions for MERS-CoV S map to residues 18 to 351 and 379 to 580, respectively (Fig. 1b). In addition, a soluble form of human DPP4 (residues 39 to 766)

was made that was C terminally tagged with the Fc region. These proteins were expressed in HEK-293T cells after transfection of the respective expression plasmids and subsequently affinity purified from the cell culture supernatant with protein A-Sepharose beads as described previously (5). The Fc region of purified soluble DPP4 (sDPP4)-Fc was proteolytically removed with trypsin (data not shown). First, we analyzed the S1-Fc proteins and C-terminal S1 truncations thereof for the ability to interact with sDPP4 by using a copurification assay. sDPP4 was efficiently copurified by the S1-Fc variants encompassing residues 1 to 588 and 1 to 747, whereas the residue 1-to-357 S1-Fc variant was unable to bind sDPP4 (Fig. 2a). We next generated an S1-Fc variant comprising residues 358 to 588, a region homologous to the ACE2 RBD in SARS-CoV S1 (Fig. 2a). This S1-Fc truncation variant efficiently bound soluble DPP4, indicating that the DPP4 RBD is located within the residue 358-to-588 domain of the MERS-CoV spike protein.

We subsequently tested the abilities of these S1-Fc variants to bind to HEK-293T cells transiently expressing DPP4 by using flow cytometry. The S1-Fc variants encompassing residues 1 to 588 and 358 to 588 bound to DPP4-expressing HEK-293T cells with efficiencies comparable to that of the full-length S1 protein, whereas no binding was observed with the residue 1-to-357 S1-Fc variant (Fig. 2b). These data show the amino acid 358-to-588 S1 region to be essential and sufficient for binding to DPP4-expressing cells, consistent with the results of the sDPP4 interaction study.

To confirm the observed interactions in a more biological assay, we analyzed the ability of the S1-Fc variants to prevent MERS-CoV infection. Thus, Huh-7 cells were preincubated with the different S1-Fc variants before being inoculated with MERS-CoV. We found that the variants encompassing residues 1 to 747 and 358 to 588, but not the residue 1-to-357 S1-Fc variant, inhibited infection (Fig. 2c).

Finally polyclonal antibodies were raised in rabbits against the 1-to-747, 1-to-357, and 358-to-588 S1-Fc variants (Davids Biotechnology GmbH). The sera, which displayed equal enzyme-linked immunosorbent assay titers toward their respective antigens (1:300,000; data not shown), were tested for their ability to neutralize virus infectivity. Antibodies raised against the 358-to-588 S1-Fc variant efficiently neutralized virus infectivity and were superior to those raised against the residue 1-to-747 and 1-to-357 S1-Fc variants (Fig. 2d). This indicates that neutralizing epitopes within S1 are localized primarily in the RBD region. The antibodies elicited are likely to block the interaction of the spike protein with DPP4, thereby neutralizing MERS-CoV infectivity. Of note,

FIG 2 The DPP4 binding domain is located within residues 358 to 588 of the MERS-CoV spike protein and efficiently elicits neutralizing antibodies. (a) S1-Fc chimeric proteins and the sDPP4 receptor were expressed in HEK-293T cells and purified from the culture supernatant. S1-Fc proteins were mixed with sDPP4, isolated by protein A-Sepharose affinity, analyzed on a NoVEX 4 to 12% Tris-glycine gradient gel under nonreducing conditions, and stained with GelCodeBlue reagent. The positions of the S1-Fc proteins—running as dimers under nonreducing conditions because of an Fc interchain disulfide bond—and sDPP4, as well as the sizes of the marker proteins, are indicated. Individual proteins were loaded as controls. (b) Binding of MERS-CoV S1-Fc proteins to DPP4-expressing cells. HEK-293T cells (2.5×10^5) transfected with the control pCAGGS (gray shaded area) or with the pCAGGS-DPP4 (black line) expression plasmid were incubated with the indicated S1-Fc at 15 $\mu\text{g}/\text{ml}$ and then incubated with DyLight488-labeled goat anti-human IgG antibody and analysis by flow cytometry. An Fc chimera containing S1 of infectious bronchitis virus (IBV-S1-Fc) was included as a negative control. (c) Inhibition of MERS-CoV infection by S1-Fc 1-to-747, 1-to-357, and 358-to-588 variants. Huh7 cells were preincubated with S1-Fc 1 to 747, 1 to 357, or 358 to 588 at 40 $\mu\text{g}/\text{ml}$ for 0.5 h prior to virus inoculation (1 h), all at room temperature. Mock-incubated cells (control) and cells incubated with an anti-DPP4 polyclonal antibody were included as controls. Following incubation at 37°C for 8 h, infected cells were detected by immunofluorescence and infection was quantified (relative to the control). The experiment was carried out twice, and the data from a representative experiment are shown. Error bars indicate the standard errors of the means. (d) Neutralization of MERS-CoV infection by rabbit antisera raised against the S1-Fc 1-to-747, 1-to-357, and 358-to-588 variants. Virus (200 PFU) was premixed 1:1 with serial dilutions of sera obtained (open bars) or after immunization (closed bars) prior to inoculation onto Vero cells, and virus infection was monitored by the occurrence of a cytopathic effect at 72 h postinfection. Virus neutralization titers (VNT) were determined in quadruplicate as the highest serum dilutions that completely prevented a cytopathic effect. The experiment was carried out twice, and the data from a representative experiment are shown. Error bars indicate the standard errors of the means.

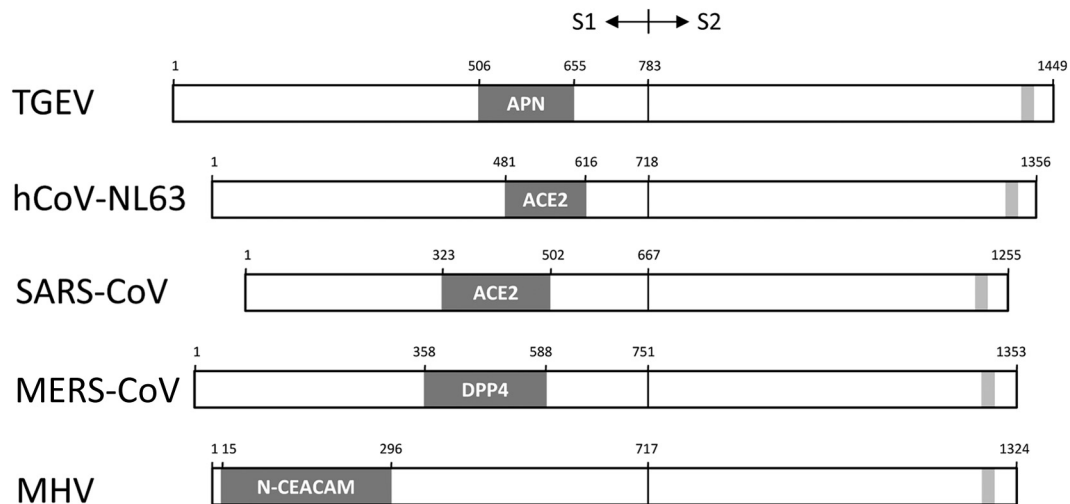


FIG 3 Localization of RBDs in coronavirus spike proteins. Shown is a schematic representation of the spike proteins of the alphacoronaviruses TGEV and hCoV-NL63 and of the betacoronaviruses SARS-CoV, MERS-CoV, and MHV (drawn to scale) aligned at the S1-S2 junction. Blue boxes represent the RBDs and indicate the receptors engaged. The RBDs of TGEV, hCoV-NL63, SARS-CoV, and MHV have been confirmed by crystallography (11, 14, 21, 25). Gray boxes indicate the transmembrane domains. Sequence identification codes: TGEV, ABG89335.1; hCoV-NL63, NC_005831.2; SARS-CoV, NP_828851.1; MERS-CoV, AFS88936.1; MHV, NC_001846.1.

antibodies raised against the MERS-CoV-S RBD did not cross-neutralize SARS-CoV infection (data not shown). The results demonstrate the potential of S1 protein and the 358-to-588 S1 polypeptide as subunit vaccines with a higher biosafety profile than vaccines based on inactivated viruses or live attenuated virus.

Except for the betacoronavirus MHV, which binds to its CEACAM receptor through a domain in the N-terminal part of its S1 protein, the RBDs of all other coronaviruses that engage protein receptors and that have been mapped occur in the C-terminal portion of the S1 subunit (Fig. 3). Examples also include the alphacoronaviruses binding to ACE2 (hCoV-NL63) and aminopeptidase N (APN; e.g., transmissible gastroenteritis virus [TGEV], hCoV-229E) (9, 18–24). In this study, we have experimentally mapped the RBD of MERS-CoV to a 231-amino-acid fragment (residues 358 to 588) within the spike protein. This domain nicely corresponds to the S1 region recently anticipated to interact with the DPP4 receptor on the basis of theoretical S1 structure predictions (25). The RBD in the MERS-CoV S1 protein localizes in the same region where the SARS-CoV S protein interacts with its ACE2 receptor (25). The SARS-CoV RBD structure displays a five-stranded β -sheet core structure (β 1 to β 4 and β 7) maintaining the overall domain conformation and a long extended loop containing two antiparallel β -sheets (β 5 and β 6) responsible for receptor binding (15). Intriguingly, compared to SARS-CoV, the RBD of MERS-CoV contains a relatively conserved core domain but a highly variable loop region, tentatively explaining the differential receptor usage (25). Crystallization and structure analysis of this MERS-CoV RBD region in complex with DPP4 will give detailed insight into the virus-receptor binding interface.

ACKNOWLEDGMENTS

We thank Ger Arkesteijn and Laura de Vries (UU, Utrecht, The Netherlands) for experimental support.

This work was supported by a fellowship from China Scholarship Council to H.M. This study was financed by European Union FP7 projects EMPERIE (contract 223498) and ANTIGONE (contract 278976).

REFERENCES

- Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367:1814–1820.
- van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, Osterhaus AD, Haagmans BL, Gorbalenya AE, Snijder EJ, Fouchier RA. 2012. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *mBio* 3(6):e00473–12. doi:10.1128/mBio.00473-12.
- de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, Fouchier RA, Galiano M, Gorbalenya AE, Memish Z, Perlman S, Poon LL, Snijder EJ, Stephens GM, Woo PC, Zaki AM, Zambon M, Ziebuhr J. 15 May 2013, posting date. Middle East respiratory syndrome coronavirus (MERS-CoV); announcement of the Coronavirus Study Group. *J. Virol.* (Epub ahead of print.) doi:10.1128/JVI.01244-13.
- Health Protection Agency (HPA) UK Novel Coronavirus Investigation Team. 2013. Evidence of person-to-person transmission within a family cluster of novel coronavirus infections, United Kingdom, February 2013. *Euro Surveill.* 18:20427. <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20427>.
- Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, Muth D, Demmers JA, Zaki A, Fouchier RA, Thiel V, Drosten C, Rottier PJ, Osterhaus AD, Bosch BJ, Haagmans BL. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* 495:251–254.
- Muller MA, Raj VS, Muth D, Meyer B, Kallies S, Smits SL, Wollny R, Bestebroer TM, Specht S, Suliman T, Zimmermann K, Binger T, Eckerle I, Tschapka M, Zaki AM, Osterhaus AD, Fouchier RA, Haagmans BL, Drosten C. 2012. Human coronavirus EMC does not require the SARS-coronavirus receptor and maintains broad replicative capability in mammalian cell lines. *mBio* 3(6):e00515–12. doi:10.1128/mBio.00515-12.
- Graham RL, Baric RS. 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* 84:3134–3146.
- Li W, Wong SK, Li F, Kuhn JH, Huang IC, Choe H, Farzan M. 2006. Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. *J. Virol.* 80:4211–4219.
- Li F. 2012. Evidence for a common evolutionary origin of coronavirus spike protein receptor-binding subunits. *J. Virol.* 86:2856–2858.
- Williams RK, Jiang GS, Holmes KV. 1991. 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.
- Kubo H, Yamada YK, Taguchi F. 1994. Localization of neutralizing

- epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. *J. Virol.* **68**:5403–5410.
12. Peng G, Sun D, Rajashankar KR, Qian Z, Holmes KV, Li F. 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proc. Natl. Acad. Sci. U. S. A.* **108**:10696–10701.
 13. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**:450–454.
 14. Wong SK, Li W, Moore MJ, Choe H, Farzan M. 2004. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J. Biol. Chem.* **279**:3197–3201.
 15. Li F, Li W, Farzan M, Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**:1864–1868.
 16. Bonavia A, Zelus BD, Wentworth DE, Talbot PJ, Holmes KV. 2003. Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E. *J. Virol.* **77**:2530–2538.
 17. He Y, Zhou Y, Liu S, Kou Z, Li W, Farzan M, Jiang S. 2004. Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine. *Biochem. Biophys. Res. Commun.* **324**:773–781.
 18. Breslin JJ, Mork I, Smith MK, Vogel LK, Hemmila EM, Bonavia A, Talbot PJ, Sjostrom H, Noren O, Holmes KV. 2003. Human coronavirus 229E: receptor binding domain and neutralization by soluble receptor at 37 degrees C. *J. Virol.* **77**:4435–4438.
 19. Hofmann H, Simmons G, Rennekamp AJ, Chaipan C, Gramberg T, Heck E, Geier M, Wegele A, Marzi A, Bates P, Pohlmann S. 2006. Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors. *J. Virol.* **80**:8639–8652.
 20. Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjostrom H, Noren O, Laude H. 1992. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* **357**:417–420.
 21. Reguera J, Ordone D, Santiago C, Enjuanes L, Casasnovas JM. 2011. Antigenic modules in the N-terminal S1 region of the transmissible gastroenteritis virus spike protein. *J. Gen. Virol.* **92**:1117–1126.
 22. Reguera J, Santiago C, Mudgal G, Ordone D, Enjuanes L, Casasnovas JM. 2012. Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog.* **8**:e1002859. doi:[10.1371/journal.ppat.1002859](https://doi.org/10.1371/journal.ppat.1002859).
 23. Yeager CL, Ashmun RA, Williams RK, Cardellicchio CB, Shapiro LH, Look AT, Holmes KV. 1992. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**:420–422.
 24. Godet M, Grosclaude J, Delmas B, Laude H. 1994. Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J. Virol.* **68**:8008–8016.
 25. Jiang S, Lu L, Du L, Debnath AK. 2013. A predicted receptor-binding and critical neutralizing domain in S protein of the novel human coronavirus HCoV-EMC. *J. Infect.* **66**:464–466.