CHAPTER FOUR

Structural insights into coronavirus entry

M. Alejandra Tortorici\textsuperscript{a,b,c}, David Veesler\textsuperscript{a,*}
\textsuperscript{a}Department of Biochemistry, University of Washington, Seattle, WA, United States
\textsuperscript{b}Institut Pasteur, Unité de Virologie Structurale, Paris, France
\textsuperscript{c}CNRS UMR 3569, Unité de Virologie Structurale, Paris, France
*Corresponding author: e-mail address: dveesler@uw.edu

Contents
1. Introduction 94
2. Prefusion S architecture 96
3. Diversity of CoV receptors and entry mechanisms 98
4. S proteolytic cleavage 101
5. Mechanism of fusion activation 103
6. Epitope masking and glycan shielding 106
7. Concluding remarks 107
Acknowledgments 108
References 108

Abstract
Coronaviruses (CoVs) have caused outbreaks of deadly pneumonia in humans since the beginning of the 21st century. The severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in 2002 and was responsible for an epidemic that spread to five continents with a fatality rate of 10% before being contained in 2003 (with additional cases reported in 2004). The Middle-East respiratory syndrome coronavirus (MERS-CoV) emerged in the Arabian Peninsula in 2012 and has caused recurrent outbreaks in humans with a fatality rate of 35%. SARS-CoV and MERS-CoV are zoonotic viruses that crossed the species barrier using bats/palm civets and dromedary camels, respectively. No specific treatments or vaccines have been approved against any of the six human coronaviruses, highlighting the need to investigate the principles governing viral entry and cross-species transmission as well as to prepare for zoonotic outbreaks which are likely to occur due to the large reservoir of CoVs found in mammals and birds. Here, we review our understanding of the infection mechanism used by coronaviruses derived from recent structural and biochemical studies.
1. Introduction

Coronaviruses (CoVs) are enveloped viruses with a positive sense RNA genome, that belong to the subfamily Coronavirinae within the family Coronaviridae, which is part of the Nidovirales order. They are classified in four genera (α, β, γ, and δ) and four lineages are recognized within the β-CoV genus (A, B, C and D). CoVs cause a variety of respiratory and enteric diseases in mammalian and avian species. Until recently, CoVs were considered to be pathogens with a largely veterinary relevance but with limited impact on human health. However, outbreaks of severe acute respiratory syndrome (SARS) in 2002–2004 and of Middle-East respiratory syndrome (MERS) starting in 2012, with fatality rates of 10% and 35%, respectively, led CoVs to be recognized as zoonotic threats with pandemic potential. Four other CoVs are endemic in the human population and cause up to 30% of mild respiratory tract infections as well as occasional severe disease in young children, the elderly or immunocompromised individuals (Isaacs et al., 1983; Su et al., 2016). These viruses are HCoV-NL63 and HCoV-229E (α-CoVs) as well as HCoV-OC43 and HCoV-HKU1 (β-CoVs). Numerous SARS-CoV and MERS-CoV-like viruses currently circulate in bats and dromedaries making outbreaks of highly pathogenic human CoVs a global health threat (Ge et al., 2013; Haagmans et al., 2014; Hu et al., 2017; Menachery et al., 2015, 2016; Sabir et al., 2016).
The CoV virion contains at least four structural proteins: spike (S), envelope (E), membrane (M) and nucleocapsid (N). In contrast to other β-CoV lineages, lineage A CoVs also encode a hemagglutinin–esterase which serves as receptor-destroying enzyme to facilitate release of viral progeny from infected cells and escape from attachment to non-permissive host cells or decoys (Bakkers et al., 2017, 2016). S is a class 1 viral fusion protein that promotes host attachment and fusion of the viral and cellular membranes during entry (Bosch et al., 2003). As a consequence, S determines host range and cell tropism. S is also the main target of neutralizing antibodies elicited during infection and the focus of vaccine design. S is trimeric and each protomer is synthesized as a single polypeptide chain of 1100–1600 residues, depending on the CoV species. For many CoVs, S is processed by host proteases to generate two functional subunits, designated S₁ and S₂, which remain non-covalently bound in the prefusion conformation (Bosch et al., 2003). The S₁ subunit comprises the apex of the S trimer, including the receptor-binding domains, and stabilizes the prefusion state of the S₂ fusion machinery, which is anchored in the viral membrane. For all CoVs, S is further cleaved by host proteases at the so-called S₂’ site located immediately upstream of the fusion peptide. This cleavage has been proposed to activate the protein for membrane fusion via large-scale, irreversible conformational changes (Heald-Sargent and Gallagher, 2012; Millet and Whittaker, 2015).

Although several class 1 viral fusion proteins have been extensively studied, CoV S proteins have proven reluctant to structural characterization until recently. Structural studies were largely limited to X-ray crystallographic analysis of isolated receptor-binding domains in complex with viral receptor ectodomains or neutralizing antibodies (Li et al., 2005a; Lu et al., 2013; Peng et al., 2011; Prabakaran et al., 2006; Reguera et al., 2012; Wang et al., 2013; Wong et al., 2017; Wu et al., 2009; Yu et al., 2015) and of the S₂ postfusion core (Duquerroy et al., 2005; Gao et al., 2013; Supekar et al., 2004; Xu et al., 2004a, b; Zheng et al., 2006) with the exception of two low-resolution electron microscopy reports (Beniac et al., 2006, 2007). In the past few years, however, technical advances in single-particle cryo-electron microscopy (cryoEM) (Bai et al., 2013; Brilot et al., 2012; Campbell et al., 2012, 2015; Li et al., 2013; Punjani et al., 2017; Scheres, 2012) together with the implementation of strategies for the stabilization of CoV S proteins in prefusion conformation (Pallesen et al., 2017; Walls et al., 2017a) led to a surge of structural data for multiple S ectodomain trimers. We review here our current understanding of the mechanism used by CoVs to infect host
cells based on recent structural and biochemical studies of S glycoprotein ectodomains in prefusion and postfusion states as well as complexes with known receptors or neutralizing antibodies.

2. Prefusion S architecture

CryoEM studies of the S glycoproteins of mouse hepatitis virus (MHV) and HKU1 led to the first structures at high-enough resolution to obtain an atomic model of the prefusion state (Kirchdoerfer et al., 2016; Walls et al., 2016a). These structures revealed that prefusion S ectodomains are \(~160\)Å-long trimers with a triangular cross-section (Fig. 1A and B).

The S₁ subunit adopts a “V” shaped architecture for \(\beta\) and \(\gamma\) CoVs (Gui et al., 2017; Kirchdoerfer et al., 2016; Shang et al., 2018a; Tortorici et al., 2019; Walls et al., 2016a; Yuan et al., 2017) (Fig. 1C), or a square-shaped organization for \(\alpha\)- and \(\delta\)-CoVs (Shang et al., 2018b; Walls et al., 2016b; Xiong et al., 2018). The S₁ subunit folds as \(\beta\)-rich domains designated A, B, C, D. Several \(\alpha\)-CoVs harbor a likely duplication of their domain A at the N-terminus of the S glycoprotein (Hulswit et al., 2016; Walls et al., 2016b). This additional domain, designated domain 0, was visualized in the NL63 S structure and hypothesized to interact with heparan sulfate present at the host cell surface during viral entry (Milewska et al., 2014; Walls et al., 2016b). Domain A and domain 0 adopt a galectin-like \(\beta\)-sandwich fold conserved across all CoV genera (Kirchdoerfer et al., 2016; Peng et al., 2011, 2012; Walls et al., 2016a) (Fig. 1D). Domain B, which shows the highest sequence variability within CoV S₁ subunits, has a markedly different architecture between \(\alpha\)-, \(\beta\)-, \(\gamma\)- and \(\delta\)-CoVs. B domains of \(\beta\)-CoVs contain a \(\beta\)-sheet core subdomain decorated with a highly variable external subdomain mediating receptor engagement (Chen et al., 2013; Kirchdoerfer et al., 2016; Li et al., 2005b; Lu et al., 2013; Tortorici et al., 2019; Walls et al., 2016a; Wang et al., 2013) (Fig. 1E). B domains of \(\alpha\)-, \(\gamma\)- and \(\delta\)-CoV form a \(\beta\)-sandwich decorated with loops mediating receptor attachment (Reguera et al., 2012; Shang et al., 2018a, b; Walls et al., 2016b; Wong et al., 2017; Xiong et al., 2018). In the context of the S trimer, \(\beta/\gamma\) CoV B domains interact with the A and B domains of another protomer, whereas they pack against the A domain of the same protomer in \(\alpha/\delta\)-CoVs (Shang et al., 2018a; Walls et al., 2016a, b; Xiong et al., 2018).

The S₂ subunit, which is more conserved than S₁, comprises the fusion machinery and connects to the viral membrane. It is assembled from a large
number of α-helices, an antiparallel core β-sheet, a β-rich connector domain and a stem helix leading to the heptad-repeat 2 (HR2) and the transmembrane region (Fig. 1F) (Gui et al., 2017; Kirchdoerfer et al., 2016, 2018; Pallesen et al., 2017; Shang et al., 2018a, b; Tortorici et al., 2019; Walls et al., 2016a, b, 2017b; Xiong et al., 2018; Yuan et al., 2017). Key S2 features facilitating virus-cell fusion include the fusion peptide, two heptad repeat regions (named HR1 and HR2) and the transmembrane domain. In the prefusion S conformation, a central helix stretches along the threefold axis, perpendicular to the viral membrane, and is located downstream the HR1

Fig. 1 CryoEM structure of the apo-HCoV-OC43 S glycoprotein. (A) Ribbon diagrams of the apo HCoV-OC43 S ectodomain trimer (PDB: 6OHw) in two orthogonal orientations, from the side (left) and from the top, looking towards the viral membrane (right). (B) Side view of one S protomer. (C) Ribbon diagram of the HCoV-OC43 S1 subunit. (D–E) Close-up view of HCoV-OC43 domain A (D) and domain B (E). (F) Ribbon diagram of the HCoV-OC43 S2 subunit in the prefusion conformation. The N- and C-termini are labeled in panels (B–E).
motif, which-fold-as-four-consecutive-α-helices (Kirchdoerfer et al., 2016; Walls et al., 2016a). Moreover, an upstream helix runs parallel to and is zipped against the central helix via hydrophobic contacts (Fig. 1F). The CoV S2 subunit shares similarity with the pneumovirus/paramyxovirus F proteins—including a comparable 3D organization of the core β-sheet, the upstream helix and the central helix—suggesting an evolutionary relatedness between the viral fusion proteins of these different viruses and a conservation of their fusion mechanism (McLellan et al., 2013; Walls et al., 2016a, 2017b; Wong et al., 2016; Xu et al., 2015; Yin et al., 2006).

A conserved tryptophan-rich segment (Y(V/I)KWPW(Y/W)VWL) directly preceding the CoV S transmembrane region is crucial for proper trimerization. This segment is also required functionally for formation of a fusion pore (Schroth-Diez et al., 2000). Furthermore, transmembrane domain interactions within and possibly between S trimers have been proposed to be essential to complete the membrane fusion process (Schroth-Diez et al., 2000). The transmembrane domain is followed by an intraviral/cytoplasmic tail of variable length (36–46 residues) depending on the coronavirus species, which contains a palmitoylated cysteine-rich region (of about 18–24 residues with 7–10 cysteines) and a variable C-terminal end (Thorp et al., 2006). The cytoplasmic tail is involved in assembly, intracellular transport, cell-surface expression and cell-cell fusion (Bos et al., 1995; Bosch et al., 2005; Chang et al., 2000; Lontok et al., 2004; Petit et al., 2005; Ye et al., 2004; Youn et al., 2005). Currently, no structural information is available for any CoV full-length S, hindering our understanding of the influence of the transmembrane and cytoplasmic domains on the conformation of exposed antigenic sites, as previously studied for HIV-1 envelope (Chen et al., 2015; Dev et al., 2016).

3. Diversity of CoV receptors and entry mechanisms

CoV entry into susceptible cells is a complex process that requires the concerted action of receptor-binding and proteolytic processing of the S protein to promote virus-cell fusion (Heald-Sargent and Gallagher, 2012; Millet and Whittaker, 2015). Domain 0, domain A and/or domain B can act as receptor-binding domains and both attachment and entry receptors have been described, depending on the CoV species.

Lineage A β-CoVs attack their S domain A to 5′-N-acetyl-9′-O-acetyl-sialosides (9′-O-Ac-Sia) found on glycoproteins and glycolipids at the
host cell surface to promote entry into susceptible cells (Vlasak et al., 1988). These include human CoVs OC43 and HKU1, bovine CoV (BCoV) and porcine hemagglutinating encephalomyelitis virus. We recently identified and visualized by cryoEM the HCoV-OC43 S sialoside-binding site, which is located in a groove at the surface of domain A (Fig. 2A) (Hulswit et al., 2019; Tortorici et al., 2019). This site is conserved in all other CoVs known to attach to 9-O-Ac-Sia (β-CoVs, lineage A) and shares architectural similarity with the ligand-binding pockets of CoV hemagglutinin-esterases and influenza virus C/D hemagglutinin-esterase-fusion glycoproteins, highlighting common structural principles of recognition (Bakkers et al., 2017, 2016;
Hulswit et al., 2019; Rosenthal et al., 1998; Tortorici et al., 2019). The current consensus in the field is that HCoV-OC43 only utilizes 9-O-Ac-sialosides as host receptors. In line with this statement, ligand-interacting residues were shown to be essential for S-mediated viral entry (Hulswit et al., 2019; Tortorici et al., 2019) and 9-O-Ac-Sia depletion from target cells resulted in severe decrease in virus infectivity (Krempl et al., 1995; Vlasak et al., 1988). Free 9-O-Ac-Sia, however, did not trigger S conformational changes associated with membrane fusion (Tortorici et al., 2019). This observation contrasts with data for SARS-CoV S, for which addition of the human angiotensin-converting enzyme 2 (ACE2) ectodomain (the proteinaceous receptor) promoted S refolding to the postfusion state (Song et al., 2018; Walls et al., 2019). These findings suggested that either 9-O-Ac-Sia-containing receptors differ from proteinaceous receptors in their mode of action, or that an interaction with a yet unidentified proteinaceous receptor is required before or after virus internalization for HCoV-OC43 entry into target cells.

The sialoside-binding site identified in HCoV-OC43 S is not conserved among CoVs which are also known to interact with sialoglycans to initiate host cell infection but are outside of the lineage A of β-CoVs, such as MERS-CoV (β-CoV, lineage C) or infectious bronchitis virus (δ-CoV) (Li et al., 2017; Wickramasinghe et al., 2011). Some α-CoVs such as transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV) use domain 0 to attach to sialoglycans, presumably to increase virus concentration at the cell surface and enhance subsequent attachment to proteinaceous receptors (Liu et al., 2015; Schwegmann-Wessels et al., 2003). Carbohydrate binding via this domain has been proposed to be a determinant of the TGEV enteric tropism since loss of domain 0 appears to correlate with a loss of enteric tropism for porcine respiratory coronavirus (PRCoV), the latter virus being a naturally occurring TGEV variant (Krempl et al., 1997).

CoVs exploit a limited variety of proteinaceous receptors compared with the large number and diversity of viral species. All CoVs known to engage proteinaceous receptors do so using domain B with the exception of MHV, which binds CEACAM1a using domain A (Dveksler et al., 1991; Peng et al., 2011; Williams et al., 1991). Remarkably, viruses from different genera, such as HCoV-NL63 (α-CoV) and SARS-CoV (β-CoV), can recognize the same region of ACE2 (entry receptor) using structurally distinct B domains (Hofmann et al., 2005; Li et al., 2005a, 2003; Wu et al., 2009). Many α-CoVs, including HCoV-229E, TGEV and PRCV, as well as porcine δ-CoV (PDCoV) utilize aminopeptidase N (APN) as entry receptor (Delmas et al., 1992; Delmas et al., 1993; Li et al., 2018; Reguera et al., 2012;
Wong et al., 2017; Yeager et al., 1992) whereas MERS-CoV uses dipeptidyl peptidase 4 (DPP4) (Lu et al., 2013; Raj et al., 2013; Wang et al., 2013). Crystal structures of SARS-CoV, HCoV-NL63, MERS-CoV, HCoV-229E B domains in complex with their cognate receptors provided atomic details of the interacting-interface and identified key residues for cross-species transmission and infection (Li et al., 2005a; Lu et al., 2013; Wang et al., 2013; Wong et al., 2017; Wu et al., 2009). This information will be useful to guide the development of therapeutics and vaccines against human CoVs.

Recent cryoEM studies revealed that MERS-CoV S and SARS-CoV S can adopt open and closed conformations in which the receptor binding site of domain B is exposed and occluded, respectively (Gui et al., 2017; Kirchdoerfer et al., 2018; Pallesen et al., 2017; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017). In contrast, the MHV, HCoV-NL63, HCoV-HKU1, PDCoV, IBV and HCoV-OC43 S glycoproteins appear to only adopt a closed conformation (Kirchdoerfer et al., 2016; Shang et al., 2018a, b; Tortorici et al., 2019; Walls et al., 2016a, b; Xiong et al., 2018) and unknown trigger(s), besides proteolytic activation, might be necessary for these viruses to expose their receptor-binding motifs for recognition to occur. These findings suggest that CoVs have evolved a fine-tuned mechanism to balance masking of the receptor-binding motifs, putatively to avoid neutralization by the host humoral immune response, and their necessary exposure to enable receptor recognition and infection of host cells (Walls et al., 2016b, 2019; Wong et al., 2017).

Upon host recognition, CoVs are internalized via receptor-mediated clathrin-dependent, caveolin-dependent or other uptake pathways (Burkard et al., 2014; Eifart et al., 2007; Inoue et al., 2007; Nomura et al., 2004). For instance, both clathrin-dependent and clathrin/caveolae-independent entry pathways have been reported for SARS-CoV (Inoue et al., 2007; Wang et al., 2008). Feline infectious peritonitis virus was suggested to enter host cells via a clathrin/caveolin-independent internalization route (Regan et al., 2008; Van Hamme et al., 2008) whereas a caveolin-dependent endocytic uptake has been suggested for HCoV-229E and HCoV-OC43 (Nomura et al., 2004; Owczarek et al., 2018).

4. S proteolytic cleavage

Several reports have demonstrated the key role of proteolytic processing of CoV S for cell-cell fusion activity and/or virus entry into host cells using experiments of inhibition of intracellular proteases.
(Burkard et al., 2014; Frana et al., 1985; Simmons et al., 2005; Yamada and Liu, 2009) and/or substitutions of residues at the $S_1/S_2$ or $S_2'$ cleavage sites (Belouzard et al., 2009; Li et al., 2015; Millet and Whittaker, 2014; Wicht et al., 2014; Yang et al., 2015).

Prior to and/or after uptake of the virion by a host cell, the S protein is proteolytically processed by host proteases at one or two cleavage sites and both receptor-binding and proteolytic processing act in synergy to induce large-scale S conformational changes promoting CoV entry. One of the cleavage sites is located at the boundary between the $S_1$ and $S_2$ subunits ($S_1/S_2$ cleavage site), whereas the other is located immediately upstream of the fusion peptide ($S_2'$ cleavage site), reviewed in (Millet and Whittaker, 2015). Cleavage at the $S_1/S_2$ site can occur upon viral egress, such as for MHV (Frana et al., 1985), or upon encounter with a target cell, such as for SARS-CoV (Belouzard et al., 2009; Bosch et al., 2008; Shulla et al., 2011), to yield two non-covalently associated subunits. This first cleavage event, along with binding to the host receptor, promotes further cleavage at the $S_2'$ site for SARS-CoV S (Belouzard et al., 2009) and MERS-CoV S (Millet and Whittaker, 2014; Park et al., 2016). Proteolysis at the conserved $S_2'$ site is essential for fusion activation of all characterized CoV S proteins, and it can occur at the host membrane or in internal cellular compartments of the target cell (Belouzard et al., 2009; Burkard et al., 2014; Millet and Whittaker, 2015; Park et al., 2016).

Cleavage at the MERS-CoV $S_1/S_2$ site by furin during viral egress enables subsequent exposure of the $S_2'$ site upon binding to the host receptor and a second cleavage step by serine proteases anchored in the membrane of the target cells, eventually leading to fusion at the cytoplasmic membrane (early entry) (Park et al., 2016). Conversely, MERS-CoV budding with uncleaved S glycoproteins traffic to the endosomes of target cells where cathepsin L or other proteases promote membrane fusion (late entry) (Park et al., 2016). The former mechanism has been proposed to be the route of MERS-CoV entry into cell types relevant to lung infection, and therefore a significant determinant of MERS-CoV virulence (Park et al., 2016). Moreover, tetraspanin CD9 has been implicated in clustering DPP4 and transmembrane serine proteases to promote early entry of MERS-CoV (Earnest et al., 2017, 2015). PEDV, which replicates in the epithelial cells of the small intestine, undergoes S proteolytic activation by trypsin, which is highly abundant in the intestinal lumen (Wicht et al., 2014). The critical importance of cleavage at the $S_1/S_2$ site was also exemplified in studies with the MERS-CoV-related bat coronavirus HKU4. Although HKU4
S recognizes human DPP4, \textit{in vitro} infectivity assays revealed that entry into human cells required addition of exogenous trypsin, suggesting proteolytic activation of this bat virus did not occur in human cells (Wang et al., 2014). In line with these findings, various DPP4 mammalian orthologues, with variable binding affinities for the MERS-CoV S receptor-binding domain, were shown to support virus or pseudovirus entry into target cells in the presence of an activating protease (Barlan et al., 2014). These results collectively illustrate how specific S proteolytic cleavage participates in determining the intracellular site of fusion and also viral tropism and pathogenesis of CoVs. Therefore, the zoonotic potential of CoVs is not only determined by receptor engagement, but also by proteolytic processing of the S protein required for fusion activation.

5. Mechanism of fusion activation

We showed that \textit{in vitro} trypsin cleavage of MHV, SARS-CoV and MERS-CoV S, under limited proteolysis conditions, recapitulated fusion activation by inducing the pre- to postfusion transition (Walls et al., 2017b). The cryoEM structure of the MHV S₂ subunit ectodomain trimer revealed that membrane fusion involves large-scale S conformational changes that are reminiscent of the ones described for other class 1 fusion proteins, including the pneumovirus/paramyxovirus F glycoproteins (Fig. 3) (McLellan et al., 2011; Swanson et al., 2010, 2011; Walls et al., 2016b; Yin et al., 2005). These experiments also demonstrated that (i) the S₁ subunits stabilize the S₂ fusion machinery in the spring-loaded, metastable prefusion state before initiation of infection; and (ii) postfusion S is the ground state of the fusion reaction. Similarly to the organization of influenza virus hemagglutinins (Xiong et al., 2013), domain B interacts with the HR1-central helix hairpin in prefusion closed S structures likely to stabilize S₂ in the spring-loaded prefusion state. This interaction appears to coordinate receptor engagement with fusion. Upon receptor binding and proteolytic cleavage at the S₁/S₂ and S₂′ sites, the S₁ crown is likely shed (as observed for MERS-CoV S by Yuan et al., 2017) to facilitate a conformational change of S₂, which involves projection of the fusion peptide to a distance of \(\sim 100 \text{Å}\) and its insertion into the target membrane (Fig. 3) (Walls et al., 2016a, 2017b). The free energy released upon S₂ refolding from the prefusion to the postfusion state is believed to bring the viral and host membranes in close proximity and promote membrane merger (Harrison, 2008).
Recent structural work comparing recombinant S proteins from SARS-CoV and MERS-CoV in isolation and in complex with their cognate receptors or neutralizing antibodies suggested an activation mechanism for coronavirus fusion (Gui et al., 2017; Kirchdoerfer et al., 2018).

Fig. 3  CoV S conformational changes driving the fusion reaction. (A), Ribbon diagram of the MHV S2 subunit in the prefusion conformation, PDB: 3JCL. (B), Ribbon diagram of the MHV S2 subunit in the postfusion conformation, PDB: 6B3O. The prefusion to postfusion transition involves a “jack-knife” refolding of the HR1 helices and intervening regions into a single continuous helix appended to the central helix. The connector domain and HR2 in the prefusion structure and the fusion peptide in the postfusion structure of MHV were not resolved and are therefore not shown.

Recent structural work comparing recombinant S proteins from SARS-CoV and MERS-CoV in isolation and in complex with their cognate receptors or neutralizing antibodies suggested an activation mechanism for coronavirus fusion (Gui et al., 2017; Kirchdoerfer et al., 2018;
Pallesen et al., 2017; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017). Specifically, SARS-CoV and MERS-CoV S structures in complex with neutralizing antibodies isolated from survivors showed both antibodies competitively blocked receptor interaction, in agreement with previous surface plasmon resonance data (Corti et al., 2015; Rockx et al., 2008; Traggiai et al., 2004; Walls et al., 2019). The anti-SARS-CoV S230 antibody, however, functionally mimicked the receptor by promoting S fusogenic conformational rearrangements through a molecular ratcheting mechanism (Walls et al., 2019) (Fig. 4). These observations suggested that upon receptor recognition, bound B domains are locked in the open state, thereby releasing the constraints imposed on the HR1-central helix hairpin, allowing refolding of the S2 fusion machinery and membrane fusion to occur (Pallesen et al., 2017; Song et al., 2018; Walls et al., 2019; Yuan et al., 2017) (Fig. 4). Proteolytic activation is likely required to ensure that S glycoproteins will work in synergy, with proper spatial and temporal coordination, to drive fusion of the viral and host membranes.

Fig. 4 CryoEM structures of the SARS-CoV S glycoprotein in complex with the S230 neutralizing antibody. (A–B), Molecular surface representation of a complex with one open, one partially open, and one closed B domain, PDB: 6NB6 (left) and with three open B domains that do not follow threefold symmetry, PDB: 6NB7 (right). The structures are rendered with different colors for each S protomer (light blue, plum and gold) and the S230 Fab heavy (dark magenta) and light (magenta) chains (only the variable domains are shown).
6. Epitope masking and glycan shielding

A deep knowledge of the organization and chemical composition of carbohydrates obstructing the surface of CoV S glycoproteins is key for understanding accessibility to neutralizing antibodies and for guiding the rational development of subunit vaccines and therapeutics. S glycoproteins feature $\sim 20–35$ predicted N-linked oligosaccharides per protomer. A cryoEM structure of the HCoV-NL63 S ectodomain allowed to visualize for the first time the extensive N-linked glycans covering the surface of a CoV S trimer (Walls et al., 2016b) (Fig. 5). A subsequent study revealed that numerous glycosylation sites are strictly or topologically conserved between PDCoV S and HCoV-NL63 S although the two glycoproteins share only

**Fig. 5** Organization of the HCoV-NL63 S glycan shield. Ribbon representation of the S ectodomain trimer with N-linked glycans rendered as dark-blue spheres, PDB: 5SZS.
43% amino acid sequence identity and the two viruses belong to different genera infecting different hosts (Xiong et al., 2018). This observation suggested that all CoVs face similar immune pressure in their respective hosts, and that the areas that are masked by the conserved glycans might be key to the function of S. Based on the information gained from the HCoV-NL63 S structure, in which a glycan participates to masking the receptor-binding loops, it was proposed that the S glycan shield is involved in immune evasion, similarly to the well-characterized HIV-1 envelope trimer (Walls et al., 2016b).

Comparison of the N-linked oligosaccharides of full-length MERS-CoV S derived from virions produced in African green monkey VeroE6 cells, or of a purified MERS-CoV S ectodomain recombinantly produced in HEK293F cells, revealed an extensive overlap of glycan composition, including the presence of hybrid and complex glycans (Walls et al., 2019). Processed oligosaccharides were also observed decorating S trimers at the surface of authentic SARS-CoV virions (Krokhin et al., 2003; Ritchie et al., 2010). These data indicated that at least a fraction of the MERS-CoV and SARS-CoV virions produced in a cell are exposed to the glycan-processing enzymes residing in the Golgi apparatus during assembly and budding, in contrast with previous models of CoV budding (Ng et al., 2003; Stertz et al., 2007).

A common feature observed in the glycosylation patterns of S glycoproteins is the presence of less densely glycosylated regions surrounding the S1/S2 cleavage site and the conserved fusion peptide, near the S2 cleavage site, probably to allow access to activating host proteases and for membrane fusion to take place (Walls et al., 2016b; Walls et al., 2019) (Fig. 5). These “glycan holes” could be targeted for epitope-focused immunogen design or new therapeutic development against CoV, as supported by the identification of a neutralization epitope within a comparable breach of the HIV-1 envelope glycan shield (McCoy et al., 2016).

7. Concluding remarks

Recent structural and functional characterization of CoV S glycoproteins provided insights into the mechanism used by these viruses to infect host cells and suggested possible strategies for rational design of vaccines and therapeutics. Introducing stabilizing mutations, which prevent the prefusion to postfusion S transition, led to the elicitation of improved neutralization titers in mice and will be a key tool for the design of subunit
vaccines against CoVs (Kirchdoerfer et al., 2018; Pallesen et al., 2017). Furthermore, the exposure of the fusion peptide at the surface of prefusion S trimers (Walls et al., 2016a) and its conservation among CoVs indicate it might be an attractive target for broad inhibition of CoV entry. Major antigenic determinants of MHV and SARS-CoV S overlap with the fusion peptide region (Daniel et al., 1993; Zhang et al., 2004) and binding of neutralizing antibodies to this site could putatively prevent fusogenic conformational changes, as proposed for influenza virus hemagglutinin or HIV envelope (Corti et al., 2011; Kong et al., 2016; Lang et al., 2017). Finally, masking strain-specific antigenic regions via engineering of additional N-linked glycosylation sites, as implemented for the MERS-CoV domain B (Du et al., 2016), bears the promise of focusing the immune response on highly conserved epitopes and eliciting broadly neutralizing antibodies against CoVs.

Acknowledgments
We acknowledge support from the National Institute of General Medical Sciences (R01GM120553, D.V.), the National Institute of Allergy and Infectious Diseases (HHSN272201700059C, DV), a Pew Biomedical Scholars Award (D.V.), an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund (D.V.) and the Pasteur Institute (M.A.T.).

References

Bosch, B.J., Bartelink, W., Rottier, P.J., 2008. Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. J. Virol. 82, 8887.


