JVI Accepted Manuscript Posted Online 25 October 2017
J. Virol. doi:10.1128/JVI.01556-17
Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Accepted Manuscript Posted Online
Journal of Virology
JVI
Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Cryo-EM structure of porcine delta coronavirus spike protein in the pre-fusion state

Jian Shang [1, *], Yuan Zheng [1, *], Yang Yang [1, *], Chang Liu [1], Qibin Geng [1], Wanbo Tai [2, 3], Lanying Du [2], Yusen Zhou [3], Wei Zhang [4, 5, #], Fang Li [1, #]

[1] Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, MN 55108, USA

[2] Lindsley F. Kimball Research Institute, New York Blood Center, New York, NY 10065, USA

[3] State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China

[4] Department of Diagnostic and Biological Sciences, School of Dentistry, University of Minnesota, Minneapolis, MN 55455, USA

[5] Characterization Facility, College of Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA

* These authors contributed equally to this work.

Correspondence:
Fang Li (lifang@umn.edu); Wei Zhang (zhangwei@umn.edu)

Key words: viral spike glycoprotein, pre-fusion, receptor binding, membrane fusion, immune evasion, cryo-electron microscopy, single-particle reconstruction

Running title: Structure, function, and evolution of PdCoV spike

Accepted Manuscript Posted Online

Abstract

Coronavirus spike proteins from different genera are divergent, although they all mediate coronavirus entry into cells by binding to host receptors and fusing viral and cell membranes. Here we determined the cryo-EM structure of porcine delta coronavirus (PdCoV) spike protein at 3.3-angstrom resolution. The trimeric protein contains three receptor-binding S1 subunits that tightly pack into a crown-like structure and three membrane-fusion S2 subunits that form a stalk. Each S1 subunit contains two domains, N-terminal domain (S1-NTD) and C-terminal domain (S1-CTD). PdCoV S1-NTD has the same structural fold as alpha- and beta-coronavirus S1-NTDs as well as host galectins, and it recognizes sugar as its potential receptor. PdCoV S1-CTD has the same structural fold as alpha-coronavirus S1-CTDs, but its structure differs from that of beta-coronavirus S1-CTDs. PdCoV S1-CTD binds to an unidentified receptor on host cell surfaces. PdCoV S2 is locked in the pre-fusion conformation by structural restraint of S1 from a different monomeric subunit. PdCoV spike possesses several structural features that may facilitate immune evasion by the virus, such as its compact structure, concealed receptor-binding sites, and shielded critical epitopes. Overall, this study reveals that delta-coronavirus spikes are structurally and evolutionally more closely related to alpha-coronavirus spikes than to beta-coronavirus spikes; it also has implications for the receptor recognition, membrane fusion, and immune evasion by delta-coronaviruses as well as coronaviruses in general.

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Journal of Virology JVI Journal of Virology

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

52

53 **Significance**

54 In this study we determined the cryo-EM structure of porcine delta coronavirus
55 (PdCoV) spike protein at 3.3 angstrom. This is the first atomic structure of a spike protein
56 from the delta coronavirus genus, which is divergent in amino acid sequences from the
57 well-studied alpha- and beta-coronavirus spike proteins. In the current study, we
58 described the overall structure of the PdCoV spike and the detailed structure of each of its
59 structural elements. Moreover, we analyzed the functions of each of the structural
60 elements. Based on the structures and functions of these structural elements, we discussed
61 the evolution of PdCoV spike protein in relation to the spike proteins from other
62 coronavirus genera. This study combines the structure, function, and evolution of
63 coronavirus spike proteins, and provides many insights into the receptor recognition,
64 membrane fusion, immune evasion, and evolution of PdCoV spike protein.

65

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Introduction

Coronaviruses are large enveloped RNA viruses that can be classified into four genera: α, β, γ, and δ (1). Both α- and β-coronaviruses infect mammals, γ-coronaviruses infect birds, and δ-coronaviruses infect mammals and birds (1). Representative coronaviruses include: human NL63 coronavirus (HCoV-NL63) and porcine transmissible gastroenteritis coronavirus (TGEV) from α genus; mouse hepatitis coronavirus (MHV), bovine coronavirus (BCoV), SARS coronavirus (SARS-CoV) and MERS coronavirus (MERS-CoV) from β genus; avian infectious bronchitis virus (IBV) from γ genus; porcine delta coronavirus (PdCoV) from δ genus (2). Coronaviruses from different genera demonstrate distinct serotypes, mainly due to the divergence of their envelope-anchored spike proteins (3). The spike proteins mediate viral entry into host cells by first binding to host receptors through their S1 subunit and then fusing host and viral membranes through their S2 subunit (4). Hence they are critical determinants of viral host range and tissue tropism, and also induce most of the host immune responses (5). Knowing the structure and function of the spike proteins from different genera is critical for understanding cell entry, pathogenesis, evolution, and immunogenicity of coronaviruses (6).

The receptor recognition pattern by coronaviruses is complicated (7). The S1 subunits from α- and β-coronavirus spikes contain two domains, the N-terminal domain (S1-NTD) and C-terminal domain (S1-CTD). Depending on the virus, either one or both of the S1 domains can function as the receptor-binding domain (RBD) by binding to host receptors. On the one hand, S1-CTDs from α- and β-coronaviruses have different tertiary

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

structures, but they share a common structural topology, indicating a common evolutionary origin and subsequent divergent evolution of S1-CTDs (7). α-coronavirus S1-CTDs recognize either angiotensin-converting enzyme 2 (ACE2) or aminopeptidase-N (APN) as their protein receptor, whereas β-coronavirus S1-CTDs recognize either ACE2 or dipeptidyl peptidase 4 (DPP4) (8-16). Hence S1-CTDs likely have undergone further divergent evolution to recognize different receptors. On the other hand, S1-NTDs from α- and β-coronaviruses both have the same structural fold as human galectins, and they recognize either sugar receptors or a protein receptor CEACAM1 (17-23). Hence it has been suggested that coronavirus S1-NTDs originated from host galectins and have undergone divergent evolution to recognize different receptors (7). These studies on receptor recognition by coronaviruses have revealed complex evolutionary relationships among the spikes from different genera.

The membrane fusion mechanism for coronavirus spikes is believed to be similar to those used by "class 1" viral membrane-fusion proteins (24, 25). The best studied such protein is hemagglutinin (HA) from influenza virus (26, 27). Influenza HA exists in two structurally distinct conformations. Its "pre-fusion" conformation on mature virions is a trimer, already cleaved by host proteases into receptor-binding subunit HA1 and membrane fusion subunit HA2 that remain associated. During the membrane fusion process, HA1 dissociates and HA2 undergoes a dramatic conformational change to reach its "post-fusion" conformation: two heptad repeat (HR) regions from each HA2 subunit, HR-N and HR-C, refold into a six-helix bundle, and a previously buried hydrophobic fusion peptide (FP) becomes exposed and inserts into host membrane. The cryo-EM structures of α- and β-coronavirus spikes in the pre-fusion conformation have recently

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

111 been determined (28-31). The overall architecture of α- and β-coronavirus spikes is
112 similar to, albeit more complex than, that of influenza HA. Biochemical studies have
113 identified parts of S2 that form six-helix bundle structures and hence likely correspond to
114 HR-N and HR-C respectively (32-34), and another part of S2 that associates with
115 membranes and hence likely corresponds to FP (35, 36). It was demonstrated that α-
116 coronavirus spikes are heavily glycosylated, with S2 more heavily glycosylated than S1,
117 as a viral strategy for immune evasion (29). These studies on membrane fusion by α- and
118 β-coronavirus spikes have suggested a common molecular mechanism for membrane
119 fusion shared by coronavirus spikes and other class 1 viral membrane fusion proteins (37,
120 38).

121 PdCoV from the δ genus is a highly lethal viral pathogen in piglets (39-41).
122 Compared to the extensive studies on α- and β-coronavirus spikes, much less is known
123 about the structure and function of δ-coronavirus spikes. It is not clear which of their S1
124 domains functions as the RBD, where the structural elements of S2 are located, how δ-
125 coronavirus spikes are structurally and evolutionarily related to the spikes from other
126 genera, or what strategies δ-coronavirus spikes use to evade host immune surveillance.
127 This study fills in these critical gaps by determining the cryo-EM structure of PdCoV
128 spike and revealing its functions in receptor binding, viral entry and immune evasion.

129 **Results and Discussion**

130 **Overall structure of PdCoV spike**

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

131 To capture PdCoV spike in the pre-fusion conformation, we constructed and
132 prepared PdCoV spike ectodomain (S-e) without the transmembrane anchor or
133 intracellular tail (Fig. 1A). We also excluded a short pre-transmembrane region (PTR)
134 because this region is hydrophobic and can adversely affect protein solubility (42).
135 Instead, we replaced these regions with a GCN4 trimerization tag followed by His_6 tag.
136 We expressed PdCoV S-e in insect cells, and purified it to homogeneity. We collected
137 cryo-EM data on PdCoV S-e, and determined its structure at 3.3Å resolution (Table 1;
138 Fig. 1B, Fig. 2).

139 The atomic structure of pre-fusion PdCoV S-e contains residues from 52 to 1017,
140 covering all of the key structural elements except HR-C (Fig. 1A). The overall trimeric
141 structure of PdCoV spike is similar to, but more compact than, those of α- and β-
142 coronavirus spikes: PdCoV spike has a length of 130Å from S1 to S2 and a width of 50Å
143 at S2 (Fig. 1C). S2 itself spans 100Å in length (Fig. 1D). Three S1 subunits form a
144 crown-like structure and sit on top of the trimeric S2 stalk (Fig. 1C, 1D). Three S1-CTDs
145 are located at the top and center of the spike trimer, whereas three S1-NTDs are located
146 on the lower and outer side of S1-CTDs (Fig. 3A, 3B, 3C, 3D). The S1-CTD mainly
147 stacks with the S1-NTD from the same monomeric subunit, although there also exist
148 inter-subunit interactions between S1-CTDs from different subunits and between S1-
149 CTD and S1-NTD from different subunits. In contrast, the S1 trimer of β-genus MHV
150 spike has an intertwined quaternary structure, with S1-CTD from one subunit mainly
151 stacking with S1-NTD from another subunit (Fig. 4A) (30). Like PdCoV spike, the S1-
152 CTD in α-genus HCoV-NL63 spike also mainly stacks with the S1-NTD from the same
153 subunit (Fig. 4B) (29). Moreover, whereas each subunit of PdCoV S1 contains only one

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

S1-NTD, each subunit of HCoV-NL63 S1 contains two, possibly resulting from gene duplication (Fig. 4B) (29). Connecting S1 and S2 are two subdomains, SD1 and SD2, and a long loop (Fig. 3A, 3B). The structure of PdCoV S2 is in the pre-fusion conformation and can be aligned well with those of α- and β-coronavirus S2 fragments (Fig. 4A, 4B). HR-C is missing in both the current PdCoV S2 structure and previously published α- and β-coronavirus S2 structures, suggesting that this region is poorly ordered. Our structural model also includes glycans N-linked to 39 residues on the trimer (13 on each monomeric subunit). In this article, we will illustrate the structures and functions of each of the structural elements in PdCoV spike.

Structure, function, and evolution of PdCoV S1-NTD

PdCoV S1-NTD adopts a β-sandwich fold identical to human galectins (Fig. 5A). Its core structure consists of two anti-parallel β-sheet layers: one is seven-stranded and the other is six-stranded. On top of the core structure is a short α-helix. Underneath the core structure is another three-stranded β-sheet and another α-helix. The S1-NTDs from α- and β-coronaviruses have the same galectin fold (Fig. 5B, 5C). Like PdCoV S1-NTD, α-coronavirus S1-NTDs contain a short α-helix on top of the core structure, but β-coronavirus S1-NTDs contain a ceiling-like structure in the same location. The galectin fold of PdCoV S1-NTD suggests that like some of the α- and β-coronavirus S1-NTDs, PdCoV S1-NTD may recognize sugar as host receptors to facilitate initial viral attachment to cells, and hence it may function as a viral lectin.

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

We investigated the sugar-binding capability of PdCoV S1-NTD. To this end, we expressed and purified recombinant PdCoV S1-NTD containing a C-terminal His_6 tag, and carried out an ELISA assay to examine whether it binds sugar (Fig. 5D). More specifically, PdCoV S1-NTD was incubated with mucin, which contains a variety of sugar chains on its surface; subsequently, the mucin-bound PdCoV S1-NTD was detected using antibodies recognizing its His_6 tag. The result showed that PdCoV S1-NTD bound to mucin. Thus, PdCoV S1-NTD bound to the sugar moiety of mucin and can potentially recognize sugar as its receptor. The sugar-binding site in PdCoV S1-NTD is currently unknown. Because the sugar-binding site in β-genus BCoV S1-NTD and the galactose-binding site in human galectins are both located on top of the core structure (18, 43), the sugar-binding site in PdCoV S1-NTD may also be located in the same region (Fig. 5A, 5C).

The above structural and functional analyses of PdCoV S1-NTD provide insight into the evolution of coronavirus S1-NTDs from different genera. Previously, based on the structures and functions of β-coronavirus S1-NTDs, we hypothesized that ancestral coronaviruses acquired a galectin gene from the host and incorporated it into their spike gene, which began to encode S1-NTD; we further predicted that the S1-NTDs from other genera also contain the galectin fold. Both the structure of PdCoV S1-NTD presented here and the structures of α-coronavirus S1-NTDs determined by recent studies confirmed our earlier prediction and lent further support to our previous hypothesis. Hence, coronavirus S1-NTDs from different genera likely all have the same evolutionary origin, which might be the host galectin, and have conserved the galectin fold through evolution.

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Structure, function, and evolution of PdCoV S1-CTD

PdCoV S1-CTD adopts a β-sandwich fold also containing two β-sheet layers: one is a three-stranded anti-parallel β-sheet and the other is a three-stranded mixed β-sheet (Fig. 6A). Its structure is similar to the β-sandwich core structure of α-coronavirus S1-CTDs, but different from the core structure of β-coronavirus S1-CTDs that contains a single β-sheet layer (Fig. 6B, 6C). We previously showed that despite their different structural folds, α- and β-coronavirus S1-CTDs share the same structural topology (i.e., connectivity of secondary structural elements) (7). Similarly, PdCoV S1-CTD also shares the same structural topology with β-coronavirus S1-CTDs. Because α- and β-coronaviruses widely use their S1-CTD as the main RBD by recognizing protein receptors, PdCoV S1-CTD may also recognize a protein receptor and function as the main RBD.

We examined the possibility of PdCoV S1-CTD recognizing a receptor on the surface of mammalian cells. To this end, we expressed and purified recombinant PdCoV S1-CTD containing a C-terminal Fc tag, and performed a flow cytometry assay to detect the binding of PdCoV S1-CTD-Fc to mammalian cells (Fig. 6D). Here the cell-bound PdCoV S1-CTD was detected using antibodies recognizing its Fc tag. The result showed that PdCoV S1-CTD-Fc bound to both human and pig cells with significantly higher affinity than Fc alone, suggesting that PdCoV S1-CTD binds to a receptor on the surface of both human and pig cells. Although PdCoV S1-CTD demonstrates higher affinity for human cells than for pig cells, it is unknown whether PdCoV infects human cells since receptor recognition is only one of several factors that can impact coronavirus infections.

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

We further investigated whether PdCoV S1-CTD recognizes ACE2 or APN, two known protein receptors for α-coronavirus S1-CTDs. To this end, we prepared and purified recombinant PdCoV S1-CTD containing a C-terminal His_6 tag, and carried out a dot-blot assay to examine whether it binds ACE2 or APN (Fig. 6E). The result showed that PdCoV S1-CTD does not bind ACE2 or APN. As positive controls, TGEV S1-CTD binds APN, whereas SARS-CoV S1-CTD binds ACE2. Taken together, these results demonstrate that PdCoV S1-CTD likely functions as the main RBD and binds a yet-to-be-identified receptor on the surface of human and pig cells.

The receptor-binding site in PdCoV S1-CTD is currently unknown. In α-coronavirus S1-CTDs, the three loops on the top of the β-sandwich core function as receptor-binding motifs (RBMs) by binding to their respective protein receptor, ACE2 for HCoV-NL63 and APN for TGEV. In PdCoV S1-CTD, the same three loops are structurally similar to their counterparts in α-coronavirus S1-CTDs. Hence, these three loops in PdCoV S1-CTD may bind to a protein receptor and function as RBMs. In the current structure, the S1-CTD is in a closed conformation, with its putative RBMs pointing towards the S1-NTD and unavailable for receptor binding. To bind its receptor, the S1-CTD would need to switch to an open conformation by "standing up" on the spike trimer and rendering the putative RBMs available for receptor binding.

Based on the above structural and functional analyses, we discuss the evolution of coronavirus S1-CTDs. Because S1-CTD is located on the tip of the pre-fusion spike trimer, it is the most exposed region on the surface of virions and thereby is under heavy immune pressure to evolve. Possibly as a consequence of immune pressure, S1-CTD is

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

structurally divergent among different coronavirus genera: α- and δ-coronavirus S1-CTDs have a β-sandwich core, whereas β-coronavirus S1-CTDs have a β-sheet core. The RBMs are located on the very tip of S1-CTDs, and are even more structurally divergent than the core structure of S1-CTDs. The RBMs in α- and δ-coronavirus S1-CTDs are three short discontinuous loops; depending on the virus, their RBM loops can bind APN (as in TGEV), ACE2 (as in HCoV-NL63), or a yet-to-be-identified receptor (as in PdCoV). The RBM in β-coronavirus S1-CTDs is a long continuous subdomain; depending on the virus, their RBM can bind ACE2 (as in SARS-CoV) or DPP4 (as in MERS-CoV). Despite their structural divergence, the S1-CTDs from different genera share the same structural topology in their cores (7). These results suggest that these S1-CTDs have a common evolutionary origin and have undergone divergent evolution. Moreover, our study demonstrates that PdCoV S1-CTD is structurally and evolutionarily more closely related to α-coronavirus S1-CTDs than to β-coronavirus S1-CTDs.

Structures, functions, and evolution of S1 subdomains

The structures of SD1 and SD2 are similar to their counterparts in α- and β-coronavirus spikes (Fig. 3B). SD1 adopts a small β-sandwich fold containing two antiparallel β-sheets: one is two-stranded and the other is five-stranded. SD2 also adopts a small β-sandwich fold containing two three-stranded β-sheets: one is antiparallel and the other is mixed. Interestingly, both SD1 and SD2 consist of discontinuous regions: majority of their sequences are to the C-terminus of S1-CTD, but they also each contain a region to the N-terminus of S1-CTD. Based on these structural data, SD1 and SD2 might have evolved later than S1-NTD and S1-CTD. The main function of the two S1

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

subdomains is to connect S1 and S2, but SD1 also plays a role in membrane fusion as discussed below.

Structure, function, and evolution of S2

The overall structure of the pre-fusion trimeric PdCoV S2 is similar to those of α- and β-coronaviruses. Two central helices, CH-N and CH-C, from each subunit form a six-helix inter-subunit interface. Based on previous biochemical and structural studies using isolated regions in S2, HR-N corresponds to a region consisting of four helices and connecting loops, and HR-C corresponds to a disordered region (Fig. 7A, 7B) (30). The exact location of FP is uncertain, but it may correspond to a region consisting of two helices and a connecting loop (30). Examination of the pre-fusion and post-fusion structures of influenza HA2 suggests that during the conformational changes of PdCoV S2, HR-N from each subunit in the pre-fusion conformation would need to fold into one long central helix as part of the six-helix bundle of the post-fusion structure (Fig. 7C). Hence, like influenza HA2, part of the CH-C in PdCoV S2 should also be part of the HR-N, such that the other parts of HR-N can anchor upon CH-C and extend towards the membrane-distal direction (Fig. 7A). Like the FP in influenza HA2, the FP in PdCoV S2 would also need to change its conformation, spring out towards the membrane-distal direction, and insert into the target membrane. The reason why HR-N and FP are locked in their pre-fusion conformation is likely because S1-CTD and SD1 from another subunit sit on top of them respectively, and prevent them respectively from extending towards the membrane-distal direction. The stacking between S1 and S2 from two different subunits contributes to the compact structure of PdCoV spike trimer. Two protease cleavages, one

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

285 at the S1/S2 boundary and the other on the N-terminus of FP, can potentially remove the
286 structural restraint of S1 on S2, allowing the conformational changes of S2 to occur (30,
287 37, 44). Both the structural and mechanistic similarities between coronavirus S2 and
288 influenza HA2 suggest that the two viral membrane-fusion proteins are evolutionarily
289 related (4). The above analysis will need to be confirmed by the atomic structure of post-
290 fusion PdCoV S2.

291 **Immune evasion strategies by PdCoV spike**

292 The structure of PdCoV spike suggests immune evasion strategies by PdCoV
293 spike. First, the PdCoV spike has a compact structure. The six domains and six
294 subdomains of trimeric S1 are tightly packed (Fig. 3B, 3C), which reduces the surface
295 area of the spike protein. Despite its compact structure, S1 maintains the two-RBD
296 system, giving the virus more options in receptor selections than a single-RBD system
297 would do. Second, in the current structure, PdCoV S1-CTD is in a closed conformation
298 with its putative RBM loops facing S1-NTD and inaccessible to the host receptor (Fig.
299 3D). Upon infecting host cells, S1-CTD would need to switch to an open conformation to
300 render the putative RBM loops accessible to the host receptor. The closed-to-open
301 conformational change of S1-CTD has been observed for β-genus MERS-CoV and
302 SARS-CoV spikes (28). This mechanism can minimize the exposure of the putative RBM
303 loops to the immune system. Third, our structural model of PdCoV spike contains
304 glycans N-linked to 39 residues (13 on each subunit); there are also another 24 predicted,
305 but not observed, N-linked glycosylation sites (8 on each subunit) (Fig. 8A, 8B). Most of
306 these sites are located on the surface of S1, which is in contrast to α-genus HCoV-NL63

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

spike where S2 is more heavily glycosylated than S1. Thus, while it was previously suggested that HCoV-NL63 spike evades host immune surveillance mainly by glycan shielding its S2 epitopes (29), PdCoV spike appears to evade host immune surveillance mainly by glycan shielding its S1 epitopes. For example, the putative sugar-binding site in PdCoV S1-NTD is surrounded by glycans, which reduces the accessibility of this site to the immune system (Fig. 8C). As a comparison, the sugar-binding site in β-genus BCoV S1-NTD is also shielded, not by glycans, but by the ceiling-like structure on top of the core structure (18). Taken together, PdCoV spike has several structural features that may facilitate viral immune evasion, such as reducing surface areas, concealing receptor-binding sites, and shielding critical S1 epitopes.

Conclusions

In this study we determined the cryo-EM structure of PdCoV spike at 3.3 Å. To our knowledge, this is the first atomic structure of a spike protein from the δ coronavirus genus, which is divergent in amino acid sequences from the well-studied α- and β-coronavirus spikes. Our study reveals a compact PdCoV spike trimer locked in the pre-fusion conformation. The trimeric S1 contains six domains (three copies of S1-NTD and S1-CTD each) and six subdomains (three copies of SD1 and SD2 each) that tightly pack into a crown-like structure. PdCoV S1-NTD has the same galectin fold as α- and β-coronavirus S1-NTDs; it binds sugar and can potentially recognize sugar as its receptors. These results expand our knowledge on the structures and functions of S1-NTDs from different coronavirus genera, and provide further evidence on the common host origin of coronavirus S1-NTDs. PdCoV S1-CTD has the same β-sandwich fold as α-coronavirus

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

S1-CTDs, and this structural fold differs from the β-sheet fold of β-coronavirus S1-CTDs. However, S1-CTDs from all coronavirus genera share the same structural topology, suggesting a common evolutionary origin of coronavirus S1-CTDs. PdCoV S1-CTD binds to an unidentified receptor on mammalian cell surfaces, and may function as the main RBD. Moreover, PdCoV S1-CTD is in a closed conformation with its putative receptor-binding sites buried; it would need to switch to an open conformation for receptor binding. The structures of both S1-NTD and S1-CTD of PdCoV are more similar to those of α-coronaviruses than to those of β-coronaviruses, and hence PdCoV spike is evolutionarily more closely related to α-coronavirus spikes than to β-coronavirus spikes. The trimeric PdCoV S2 forms the stalk of the spike protein. Each of the S2 subunits is locked in the pre-fusion conformation by structural constraint of S1 from a different monomeric subunit. More specifically, HR-N and FP are prevented from re-folding into their post-fusion conformation by the steric restrictions from S1-CTD and SD1, respectively, of another subunit. PdCoV spike possesses several structural features that appear to facilitate its evasion from host immune surveillance, such as its compact structure, the closed conformation of its S1-CTD, and heavy glycosylation near critical epitopes in S1. Overall, our study combines the structure and function of PdCoV spike, and provides many insights into the receptor recognition, membrane fusion, immune evasion, and evolution of PdCoV spike as well as coronavirus spikes in general.

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Acknowledgements

This work was supported by NIH grant R01AI089728 (to F. Li) and fund from University Minnesota (to W. Zhang). Initial cryo-EM images were collected using FEI Tecnai TEMs maintained at the Characterization Facility of the University of Minnesota. Final cryo-EM data were collected at the EM facility of Purdue University. We thank Valorie Bowman, Thomas Klose, Yue Liu and Steve Wilson for help in data collection and image processing, and Hinh Ly for comments on the manuscript. Initial image analysis and computation work was performed using the workstations at the Basic Sciences Computing Laboratory of the University of Minnesota Supercomputing Institute. The cryoEM map has been deposited in the Electron Microscopy Data Bank (EMD) under accession codes 7063. The atomic model has been deposited in the Protein Data Bank (PDB) under accession codes 6B7N.

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

362 **References**

363

364 1. **Gonzaalez JM, Gomez-Puertas P, Cavanagh D, Gorbalenya AE, Enjuanes L.**
365 2003. A comparative sequence analysis to revise the current taxonomy of the
366 family Coronaviridae. Archives of Virology **148:**2207-2235.

367 2. **Perlman S, Netland J.** 2009. Coronaviruses post-SARS: update on replication
368 and pathogenesis. Nature Reviews Microbiology **7:**439-450.

369 3. **Du LY, He YX, Zhou YS, Liu SW, Zheng BJ, Jiang SB.** 2009. The spike
370 protein of SARS-CoV - a target for vaccine and therapeutic development. Nature
371 Reviews Microbiology **7:**226-236.

372 4. **Li F.** 2016. Structure, Function, and Evolution of Coronavirus Spike Proteins.
373 Annu Rev Virol **3:**237-261.

374 5. **Graham RL, Baric RS.** 2010. Recombination, reservoirs, and the modular spike:
375 mechanisms of coronavirus cross-species transmission. J Virol **84:**3134-3146.

376 6. **Graham RL, Donaldson EF, Baric RS.** 2013. A decade after SARS: strategies
377 for controlling emerging coronaviruses. Nat Rev Microbiol **11:**836-848.

378 7. **Li F.** 2015. Receptor recognition mechanisms of coronaviruses: a decade of
379 structural studies. J Virol **89:**1954-1964.

380 8. **Li WH, Moore MJ, Vasilieva N, Sui JH, Wong SK, Berne MA,**
381 **Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H,**
382 **Farzan M.** 2003. Angiotensin-converting enzyme 2 is a functional receptor for
383 the SARS coronavirus. Nature **426:**450-454.

384 9. **Raj VS, Mou HH, Smits SL, Dekkers DHW, Muller MA, Dijkman R, Muth**
385 **D, Demmers JAA, Zaki A, Fouchier RAM, Thiel V, Drosten C, Rottier PJM,**
386 **Osterhaus A, Bosch BJ, Haagmans BL.** 2013. Dipeptidyl peptidase 4 is a
387 functional receptor for the emerging human coronavirus-EMC. Nature **495:**251-
388 254.

389 10. **Delmas B, Gelfi J, Lharidon R, Vogel LK, Sjostrom H, Noren O, Laude H.**
390 1992. AMINOPEPTIDASE-N IS A MAJOR RECEPTOR FOR THE
391 ENTEROPATHOGENIC CORONAVIRUS TGEV. Nature **357:**417-420.

392 11. **Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S.**
393 2005. Human coronavirus NL63 employs the severe acute respiratory syndrome
394 coronavirus receptor for cellular entry. Proceedings of the National Academy of
395 Sciences of the United States of America **102:**7988-7993.

Accepted Manuscript Posted Online

12. **Reguera J, Santiago C, Mudgal G, Ordono D, Enjuanes L, Casasnovas JM.** 2012. Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. Plos Pathogens **8**.

13. **Chen L, Lin YL, Peng G, Li F.** 2012. Structural basis for multifunctional roles of mammalian aminopeptidase N. Proc Natl Acad Sci U S A **109:**17966-17971.

14. **Wu K, Li W, Peng G, Li F.** 2009. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. Proc Natl Acad Sci U S A **106:**19970-19974.

15. **Li F, Li WH, Farzan M, Harrison SC.** 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science **309:**1864-1868.

16. **Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, Zhang Y, Zhang W, Yuan Y, Bao J, Zhang B, Shi Y, Yan J, Gao GF.** 2013. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature **500:**227-231.

17. **Peng GQ, Sun DW, Rajashankar KR, Qian ZH, Holmes KV, Li F.** 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. Proceedings of the National Academy of Sciences of the United States of America **108:**10696-10701.

18. **Peng GQ, Xu LQ, Lin YL, Chen L, Pasquarella JR, Holmes KV, Li F.** 2012. Crystal Structure of Bovine Coronavirus Spike Protein Lectin Domain. Journal of Biological Chemistry **287:**41931-41938.

19. **Tan KM, Zelus BD, Meijers R, Liu JH, Bergelson JM, Duke N, Zhang R, Joachimiak A, Holmes KV, Wang JH.** 2002. Crystal structure of murine sCEACAM1a 1,4 : a coronavirus receptor in the CEA family. Embo Journal **21:**2076-2086.

20. **Dveksler GS, Pensiero MN, Cardellichio CB, Williams RK, Jiang GS, Holmes KV, Dieffenbach CW.** 1991. Cloning of the Mouse Hepatitis-Virus (Mhv) Receptor - Expression in Human and Hamster-Cell Lines Confers Susceptibility to Mhv. Journal of Virology **65:**6881-6891.

21. **Williams RK, Jiang GS, Holmes KV.** 1991. Receptor for Mouse Hepatitis-Virus Is a Member of the Carcinoembryonic Antigen Family of Glycoproteins. Proceedings of the National Academy of Sciences of the United States of America **88:**5533-5536.

22. **Schultze B, Gross HJ, Brossmer R, Herrler G.** 1991. THE S-PROTEIN OF BOVINE CORONAVIRUS IS A HEMAGGLUTININ RECOGNIZING 9-O-ACETYLATED SIALIC-ACID AS A RECEPTOR DETERMINANT. Journal of Virology **65:**6232-6237.

Journal of Virology
JVI
Journal of Virology
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

23. **Schultze B, Krempl C, Ballesteros ML, Shaw L, Schauer R, Enjuanes L, Herrler G.** 1996. Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (N-glycolylneuraminic acid) binding activity. Journal of Virology **70:**5634-5637.

24. **Harrison SC.** 2015. Viral membrane fusion. Virology **479-480:**498-507.

25. **Bosch BJ, van der Zee R, de Haan CAM, Rottier PJM.** 2003. The coronavirus spike protein is a class I virus fusion protein: Structural and functional characterization of the fusion core complex. Journal of Virology **77:**8801-8811.

26. **Skehel JJ, Wiley DC.** 2000. Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. Annual Review of Biochemistry **69:**531-569.

27. **Colman PM, Lawrence MC.** 2003. The structural biology of type I viral membrane fusion. Nat Rev Mol Cell Biol **4:**309-319.

28. **Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X, Gao GF.** 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nat Commun **8:**15092.

29. **Walls AC, Tortorici MA, Frenz B, Snijder J, Li W, Rey FA, DiMaio F, Bosch BJ, Veesler D.** 2016. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat Struct Mol Biol **23:**899-905.

30. **Walls AC, Tortorici MA, Bosch BJ, Frenz B, Rottier PJ, DiMaio F, Rey FA, Veesler D.** 2016. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature **531:**114-117.

31. **Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Corbett KS, Graham BS, McLellan JS, Ward AB.** 2016. Pre-fusion structure of a human coronavirus spike protein. Nature **531:**118-121.

32. **Duquerroy S, Vigouroux A, Rottier PJ, Rey FA, Bosch BJ.** 2005. Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. Virology **335:**276-285.

33. **Xu YH, Lou ZY, Liu YW, Pang H, Tien P, Gao GF, Rao ZH.** 2004. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. Journal of Biological Chemistry **279:**49414-49419.

34. **Supekar VM, Bruckmann C, Ingallinella P, Bianchi E, Pessi A, Carfi A.** 2004. Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proceedings of the National Academy of Sciences of the United States of America **101:**17958-17963.

Journal of Virology
JVI
Journal of Virology
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

35. **Ou X, Zheng W, Shan Y, Mu Z, Dominguez SR, Holmes KV, Qian Z.** 2016. Identification of the Fusion Peptide-Containing Region in Betacoronavirus Spike Glycoproteins. J Virol **90:**5586-5600.

36. **Basso LG, Vicente EF, Crusca E, Jr., Cilli EM, Costa-Filho AJ.** 2016. SARS-CoV fusion peptides induce membrane surface ordering and curvature. Sci Rep **6:**37131.

37. **Belouzard S, Chu VC, Whittaker GR.** 2009. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proceedings of the National Academy of Sciences of the United States of America **106:**5871-5876.

38. **Heald-Sargent T, Gallagher T.** 2012. Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses **4:**557-580.

39. **Ma Y, Zhang Y, Liang X, Lou F, Oglesbee M, Krakowka S, Li J.** 2015. Origin, evolution, and virulence of porcine deltacoronaviruses in the United States. MBio **6:**e00064.

40. **Hu H, Jung K, Vlasova AN, Chepngeno J, Lu Z, Wang Q, Saif LJ.** 2015. Isolation and characterization of porcine deltacoronavirus from pigs with diarrhea in the United States. J Clin Microbiol **53:**1537-1548.

41. **Dong N, Fang L, Yang H, Liu H, Du T, Fang P, Wang D, Chen H, Xiao S.** 2016. Isolation, genomic characterization, and pathogenicity of a Chinese porcine deltacoronavirus strain CHN-HN-2014. Vet Microbiol **196:**98-106.

42. **Mahajan M, Bhattacharjya S.** 2015. NMR structures and localization of the potential fusion peptides and the pre-transmembrane region of SARS-CoV: Implications in membrane fusion. Biochim Biophys Acta **1848:**721-730.

43. **Seetharaman J, Kanigsberg A, Slaaby R, Leffler H, Barondes SH, Rini JM.** 1998. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-angstrom resolution. Journal of Biological Chemistry **273:**13047-13052.

44. **Millet JK, Whittaker GR.** 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc Natl Acad Sci U S A **111:**15214-15219.

45. **Li F, Berardi M, Li WH, Farzan M, Dormitzer PR, Harrison SC.** 2006. Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain. Journal of Virology **80:**6794-6800.

46. **Carragher B, Kisseberth N, Kriegman D, Milligan RA, Potter CS, Pulokas J, Reilein A.** 2000. Leginon: an automated system for acquisition of images from vitreous ice specimens. J Struct Biol **132:**33-45.

Journal of Virology
JVI
Journal of Virology
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

47. **Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y.** 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods **10:**584-590.

48. **Zhang K.** 2016. Gctf: Real-time CTF determination and correction. J Struct Biol **193:**1-12.

49. **Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SH, Henderson R.** 2013. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy **135:**24-35.

50. **Goddard TD, Huang CC, Ferrin TE.** 2007. Visualizing density maps with UCSF Chimera. J Struct Biol **157:**281-287.

51. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr **66:**486-501.

52. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr **66:**213-221.

53. **Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC.** 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr **66:**12-21.

54. **Barad BA, Echols N, Wang RY, Cheng Y, DiMaio F, Adams PD, Fraser JS.** 2015. EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. Nat Methods **12:**943-946.

55. **Liu C, Tang J, Ma Y, Liang X, Yang Y, Peng G, Qi Q, Jiang S, Li J, Du L, Li F.** 2015. Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. J Virol **89:**6121-6125.

56. **Du L, Tai W, Yang Y, Zhao G, Zhu Q, Sun S, Liu C, Tao X, Tseng CK, Perlman S, Jiang S, Zhou Y, Li F.** 2016. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. Nat Commun **7:**13473.

57. **Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE.** 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem **25:**1605-1612.

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Journal of Virology JVI Journal of Virology

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

542 **Materials and Methods**

543 *Expression, purification, and treatment of PdCoV spike ectodomain*

544 PdCoV spike ectodomain (S-e) (residues 18-1077) was cloned into pFastBac
545 vector (Life Technologies Inc.) with a N-terminal honeybee melittin signal peptide and
546 C-terminal GCN4 and His_6 tags. It was expressed in sf9 insect cells using the Bac-to-Bac
547 system (Life Technologies Inc.) and purified as previously described (15). Briefly, the
548 protein was harvested from cell culture medium, and purified sequentially on Ni-NTA
549 column and Superdex200 gel filtration column (GE Healthcare). Because we showed
550 earlier that low pH could facilitate trimer formation (45), we incubated PdCoV S-e in
551 buffer containing 0.1 M sodium citrate (pH 5.6) at room temperature for 1 hour, and then
552 re-purified it on Superdex200 gel filtration column in buffer containing 20 mM Tris
553 pH7.2 and 200 mM NaCl.

554 *Cryo-electron microscopy*

555 For sample preparation, aliquots of PdCoV S-e (3 µl, 0.35 mg/ml, in buffer
556 containing 2 mM Tris pH7.2 and 20 mM NaCl) were applied to glow-discharged CF-2/1-
557 4C C-flat grids (Protochips). The grids were then plunge-frozen in liquid ethane using a
558 FEI MarkIII Vitrobot system (FEI Company).

559 For data collection, images were recorded using a Gatan K2 Summit direct
560 electron detector in the direct electron counting mode (Gatan), attached to a Titan-Krios
561 TEM (FEI Company), at Purdue University. The automated software Leginon (46) was
562 used to collect ~2,100 movies at 22,500x magnification and at a defocus range of
563 between 0.5 and 3 µm. Each movie had a total accumulated exposure of 52 $e/Å^2$

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

fractionated in 55 frames of 200 ms exposure. Data collection statistics are summarized in Table 1.

For data processing, the recorded movies were corrected for beam-induced motion using MotionCor2 (47). The final image was bin-averaged to give the pixel size to be 1.3Å. The parameters of the microscope contrast transfer function were estimated for each micrograph using GCTF (48). Particles were automatically picked and extracted using RELION 2.0 on a GPU workstation with a box size of 256 pixels. Initially, particles were subjected to 2D alignment and clustering using RELION 2.0, and the best classes were selected for an additional 2D alignment. Some of the particles on 2D class averages appear to have a tail (Fig. S1A), which may correspond to HR-C. Nevertheless, the weak density of the tail region suggests that this region is poorly ordered, and hence this region was not included in subsequent map calculation and model building. All of the particles, with or without the tail, were subjected to 3D auto-refine with a mask covering the overall shape of the particles (excluding the tail region) to yield the map. The orientations of the particles used in the final reconstruction map sufficiently covered the whole sphere in the Fourier space to allow calculation of a 3D map with isotropic resolution. The map was sharpened with modulation transfer function of K2 operated at 300kV using RELION 2.0 post processing. Reported resolution was based on the gold-standard Fourier shell correlation (FSC) = 0.143 criterion, and Fourier shell correction curves were corrected for the effects of soft masking by high-resolution noise substitution (49). Data processing statistics are summarized in Table 1.

Model building and refinement

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

For atomic model building, the cryo-EM structure of HCoV-NL63 spike (PDB: 5SZS) were divided into 7 parts (S1-NTD, SD2', SD1', S1-CTD, SD1", SD2" and S2), and fitted into the cryo-EM map of PdCoV S-e individually using UCSF Chimera (50) and *Coot (51)*. Model rebuilding was performed manually in *Coot* based on the well-defined continuous density of the main chain, and sequence register assignment was guided mainly by the density of N-linked glycans and of bulky amino acid residues. The structural model was refined using Phenix (52) with geometry restrains and three-fold noncrystallographic symmetry constraints. Refinement and manual model correction in *Coot* were carried out iteratively until there was no more improvement in geometry parameters. The quality of the final model was analyzed with MolProbity (53) and EMRinger (54). The validation statistics of the structural model are summarized in Table 1.

ELISA sugar-binding assay

PdCoV S1-NTD containing a C-terminal His_6 tag was expressed and purified in the same way as PdCoV S-e, and assayed for its sugar-binding capability using an ELISA assay as previously described (18). Briefly, ELISA plates were pre-coated with bovine mucin (1 mg/ml) at 37 °C for 1 hour. After blocking with 1% BSA at 37 °C for 1 hour, PdCoV S1-NTD (1 μg/ml) was added to the plates and incubated with mucin at 37 °C for 1 hour. After washes with PBS buffer, the plates were incubated with anti-His_6 antibody (Santa Cruz) at 37 °C for 1 hour. Then the plates were washed with PBS and incubated with HRP-conjugated goat anti-mouse IgG antibody (1:5,000) at 37 °C for 1 hour. After more washes with PBS, enzymatic reaction was carried out using ELISA substrate (Life Technologies Inc.) and stopped with 1 M H_2SO4. Absorbance at 450 nm (A450) was

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

609 measured using Tecan Infinite M1000 PRO Microplate Reader (Tecan Group Ltd.). Five
610 replicates were done for each sample. Porcine epidemic diarrhea virus (PEDV) S1 and
611 SARS-CoV S1-CTD were prepared as previously described (15, 55), and PdCoV S1-
612 CTD was prepared as described below; these three proteins were used in the assay as
613 controls.
614
615 *Dot-blot receptor-binding assay*

616 PdCoV S1-CTD containing a C-terminal His_6 tag was expressed and purified in
617 the same way as PdCoV S-e, and assayed for its receptor-binding capability using a dot-
618 blot receptor-binding assay as previously described (55). Briefly, 5 μM receptor (human
619 ACE2 or porcine APN) was dotted onto nitrocellulose membranes. The membranes were
620 dried and blocked with 1% BSA, and then incubated with 1 μM PdCoV S1-CTD at 4 °C
621 for 2 hours. After washes with PBS buffer, the membranes were incubated with anti-His_6
622 antibody (Life Technologies Inc.) at 4 °C for 2 hours, washed with PBS, incubated with
623 HRP-conjugated goat anti-mouse IgG antibody (1:5,000) at 4 °C for 2 hours, and washed
624 with PBS. Finally, the receptor-bound proteins were detected using a chemiluminescence
625 reagent (ECL plus, GE Healthcare). Recombinant human ACE2 and porcine APN were
626 prepared as previously described (13, 15).
627 *Flow cytometry cell-binding assay*

628 PdCoV S1-CTD containing a C-terminal Fc tag was expressed, purified, and
629 assayed for its cell-binding capability by flow cymotetry as previously described (56).
630 Briefly, human (HeLa and A549) and pig (ST and PK15) cells were incubated with
631 PdCoV S1-CTD-Fc (40 μg/ml), or human IgG-Fc control, at room temperature for 30
632 min, followed by incubation with fluorescein isothiocyanate (FITC)-labeled anti-human

Accepted Manuscript Posted Online

633 IgG-Fc antibody for 30 min. The cells were then analyzed for the binding using flow
634 cytometry.
635

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Journal of Virology JVI Journal of Virology

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

636 **Table 1 Data and model statistics**
637

Data collection	
Microscope	Titan Krios
Voltage (keV)	300
Defocus range (μm)	1.0 to 4.0
Movies	2168
Frames per movie	55
Dose rate ($e^-/A^2/s$)	4.7
Total dose per movie (e^-/A^2)	51.7
Data processing	
Particles	87,002
Symmetry	C3
Provided B-factor ($Å^2$)	-150
Map resolution (Å)	3.3
Model Validation	
UCSF Chimera CC(57)	0.865
EMRinger Score(54)	2.77
MolProbity Score(53)	1.91
All-atom clashscore(53)	5.48
Rotamers outliers (%)	0.78
Ramachandran allowed (%)	99.59
Ramachandran outliers (%)	0.41
R.m.s deviations	
Bond length (Å)	0.009
Bond angles (°)	1.437

639

640

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

641 **Figure Legends:**

642 **Figure 1. Overall structure of PdCoV S-e in the pre-fusion conformation.** (A)
643 Schematic drawing of PdCoV S-e (spike ectodomain). S1: receptor-binding subunit. S2:
644 membrane-fusion subunit. GCN4-His_6: GCN4 trimerization tag followed by His_6 tag. S1-
645 NTD: N-terminal domain of S1. S1-CTD: C-terminal domain of S1. CH-N and CH-C:
646 central helices N and C. FP: fusion peptide. HR-N and HR-C: heptad repeats N and C.
647 Residues in shaded regions (N-terminus, GCN4 tag, and His_6 tag) were not traced in the
648 structure. (B) Cryo-EM maps of PdCoV S-e with atomic model fitted in. The maps have
649 a contour of 6.6 σ. (C) Cryo-EM structure of pre-fusion PdCoV S-e. Each of the
650 monomeric subunits is colored differently. (D) Structure of a monomeric subunit in the
651 pre-fusion conformation. The structural elements are colored in the same way as in panel
652 (A).

653

654 **Figure 2. Cryo-EM data analysis of PdCoV S-e.** (A) Representative micrographs of
655 frozen-hydrated PdCoV S-e particles and representative 2D class averages in different
656 orientations. Arrow indicates a poorly ordered tail region in some of the particles. (B)
657 Gold-standard Fourier shell correlation (FSC) curves. The resolution was determined to
658 be 3.3 Å. The 0.143 and 0.5 cut-off values are indicated by horizontal grey bars. (C) Final
659 cryo-EM map of PdCoV S-e colored according to the local resolution.

660

661 **Figure 3. Structure of PdCoV S1.** (A) Schematic drawing of PdCoV S1. SD1:
662 subdomain 1. SD2: subdomain 2. SD1 consists of two discontinuous regions SD1' and
663 SD1". SD2 consists of two discontinuous regions SD2' and SD". (B) Structure of

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

664 monomeric S1. Domains and subdomains are colored in the same way as in panel (A).
665 Residue ranges for each of the domains and subdomains are labeled. (C) Structure of
666 trimeric S1, viewed from the side. Each of the monomeric subunits is colored differently.
667 The empty space under S1 is occupied by S2, which is not shown here. (D) Structure of
668 trimeric S1, viewed from the top. Each of the monomeric subunits is colored differently.
669
670 **Figure 4. Structural alignments of PdCoV spike with the spikes from other**
671 **coronavirus genera.** (A) Alignment of PdCoV and β-genus MHV spikes. PdCoV spike
672 is colored in magenta. MHV spike (PDB ID: 3JCL) is colored in cyan. (B) Alignment of
673 PdCoV and α-genus HCoV-NL63 spikes. PdCoV spike is colored in magenta. HCoV-
674 NL63 spike (PDB ID: 5SZS) is colored in green. Each subunit of PdCoV S1 contains
675 only one S1-NTD, whereas each subunit of HCoV-NL63 S1 contains two.
676
677 **Figure 5. Structure and function of PdCoV S1-NTD.** (A) Structure of PdCoV S1-
678 NTD. The putative sugar-binding site is indicated by the question mark. (B) Structure of
679 α-genus HCoV-NL63 S1-NTD (PDB ID: 5SZS). (C) Structure of β-genus BCoV S1-
680 NTD (PDB ID: 4H14). (D) ELISA sugar-binding assay for PdCoV S1-NTD. Here the
681 ELISA plates were pre-coated with sugar-rich mucin, and then PdCoV S1-NTD was
682 added and incubated with mucin. Mucin-bound S1-NTD was detected using antibodies
683 recognizing its C-terminal His_6 tag. Porcine epidemic diarrhea virus (PEDV) S1 was used
684 as the positive control; PdCoV S1-CTD, SARS-CoV S1-CTD, and BSA were used as
685 negative controls. Plate without mucin was used as an additional negative control.

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

686 Statistic analyses were performed using two-tailed t-test. Error bars indicate S.E.M.
687 (n=5). *** P<0.001.
688
689 **Figure 6. Structure and function of PdCoV S1-CTD.** (A) Structure of PdCoV S1-
690 CTD. The putative RBM loops are indicated by the question mark. (B) Structure of α-
691 genus HCoV-NL63 S1-CTD (PDB ID: 3KBH). (C) Structure of β-genus SARS-CoV S1-
692 CTD (PDB ID: 2AJF). (D) Flow cytometry assay for the binding of PdCoV S1-CTD to
693 the surface of mammalian cells. Cell-bound PdCoV S1-CTD was detected using
694 antibodies recognizing its C-terminal Fc tag. Fc or cells only were used as negative
695 controls. Statistic analyses were performed using two-tailed t-test. Error bars indicate
696 S.E.M. (n=4). *** P<0.001. (E) Dot-blot receptor-binding assay for PdCoV S1-CTD.
697 Here the receptor (either APN or ACE2) was first dotted onto a membrane. Subsequently,
698 PdCoV S1-CTD was dotted and incubated with the receptor. Receptor-bound S1-CTD
699 was detected using antibodies recognizing its C-terminal His_6 tag. TGEV and SARS-CoV
700 S1-CTDs were used as positive controls. PBS buffer was used as a negative control.
701
702 **Figure 7. Structure and function of PdCoV S2.** (A) Structure of the pre-fusion
703 monomeric PdCoV S2 only including CH-C, HR-N and FP. Arrow indicates the direction
704 in which HR-N would need to extend to reach the post-fusion conformation. Question
705 mark indicates part of CH-C that likely is part of HR-N. Residue ranges for each of the
706 structural elements are labeled. (B) S1-CTD and SD1 from a different subunit stack with
707 HR-N and FP, respectively, preventing them from switching to their post-fusion
708 conformation. Scissor indicates the proteolysis sites to the N-terminus of FP. (C)

Accepted Manuscript Posted Online

Structures of influenza HA2 in the pre-fusion and post-fusion conformations (PDB IDs: 2YPG and 1QU1). Arrow indicates the direction in which HR-N would need to extend to reach the post-fusion conformation. Scissor indicates the proteolysis sites to the N-terminus of FP.

Figure 8. Glycosylation sites on the surface of PdCoV spike. (A) Distribution of N-linked glycosylation sites on the one-dimensional structure of PdCoV spike. Ψ indicates N-linked glycosylate site. Those on the top indicate glycans observed in the structure. Those at the bottom indicate predicted, but not observed, glycosylate sites. Predicted glycosylation sites in the N-terminal region and HR-C were not included because these two regions were not traced in the structure. (B) Distribution of N-linked glycosylation sites on the three-dimensional structure of PdCoV spike. Observed glycans are in dark blue. Predicted, but not observed, glycosylation sites are in light blue. (C) Distribution of N-linked glycosylation sites in monomeric S1. Question marks indicate the putative sugar-binding site in S1-NTD and putative RBMs in S1-CTD, respectively.

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Journal of Virology JVI Journal of Virology

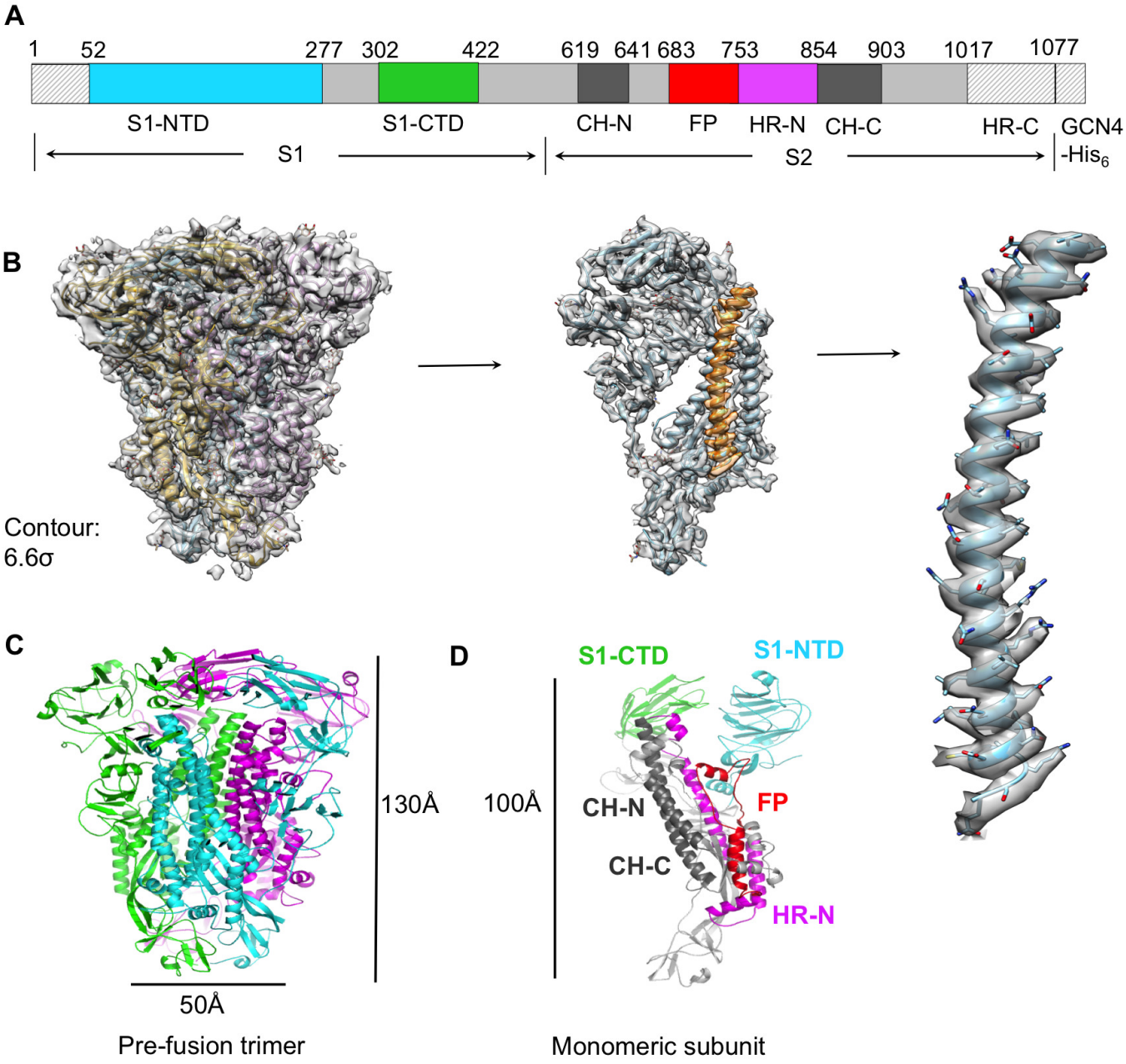
Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

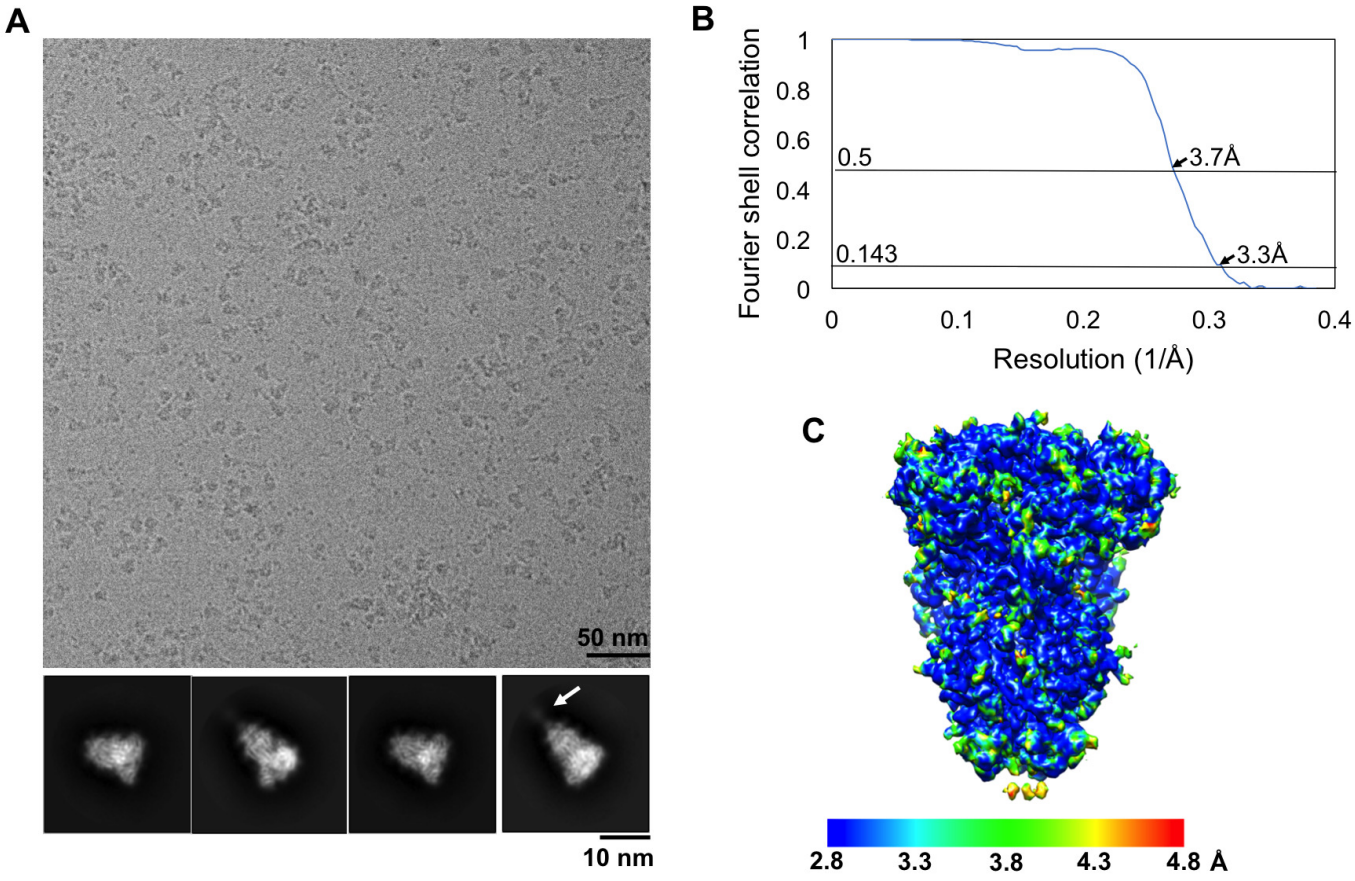


Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



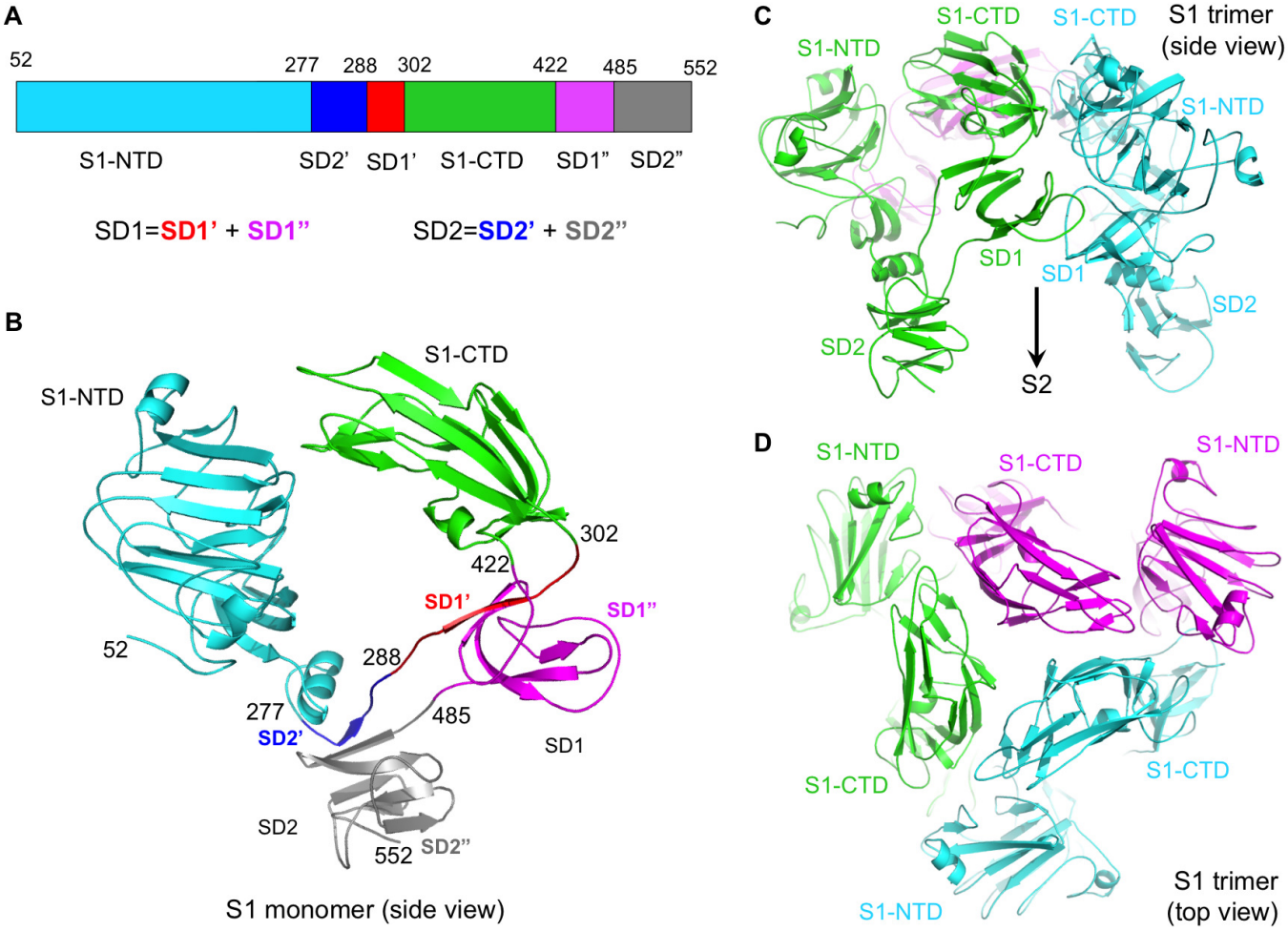
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



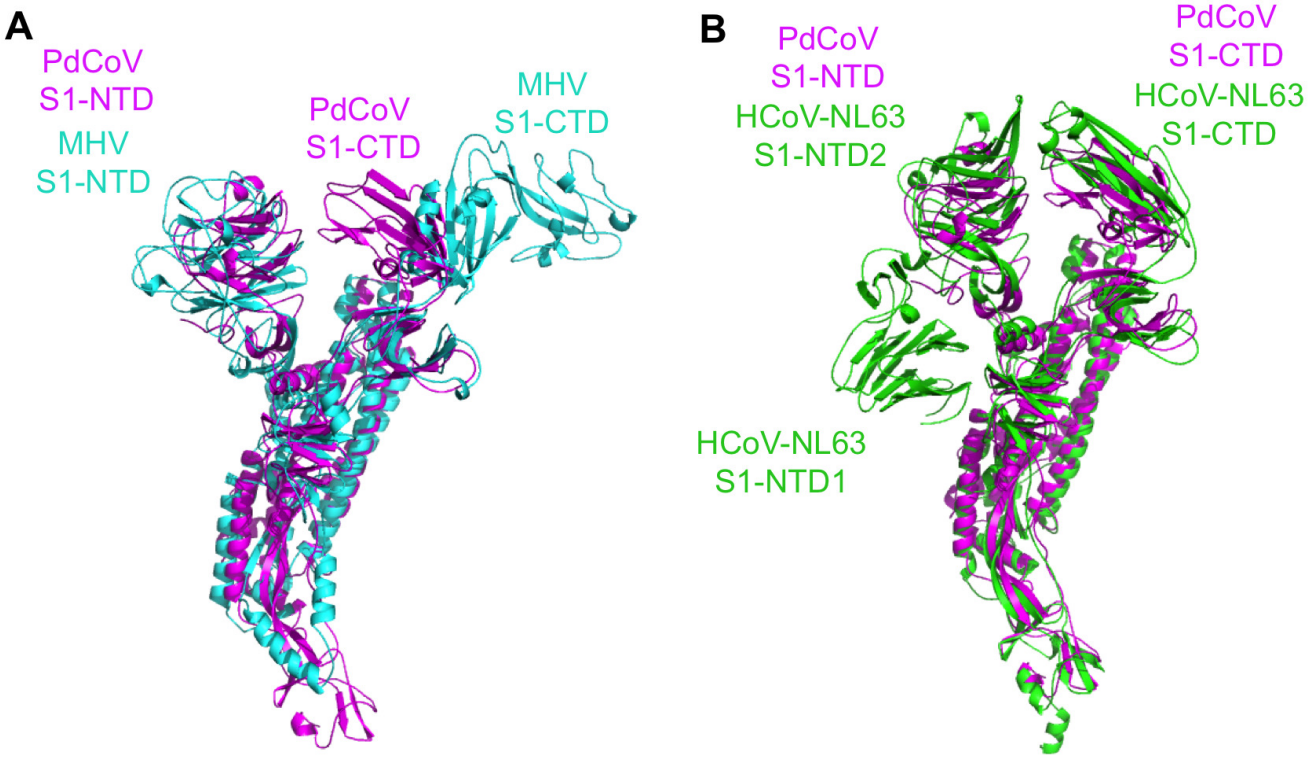
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



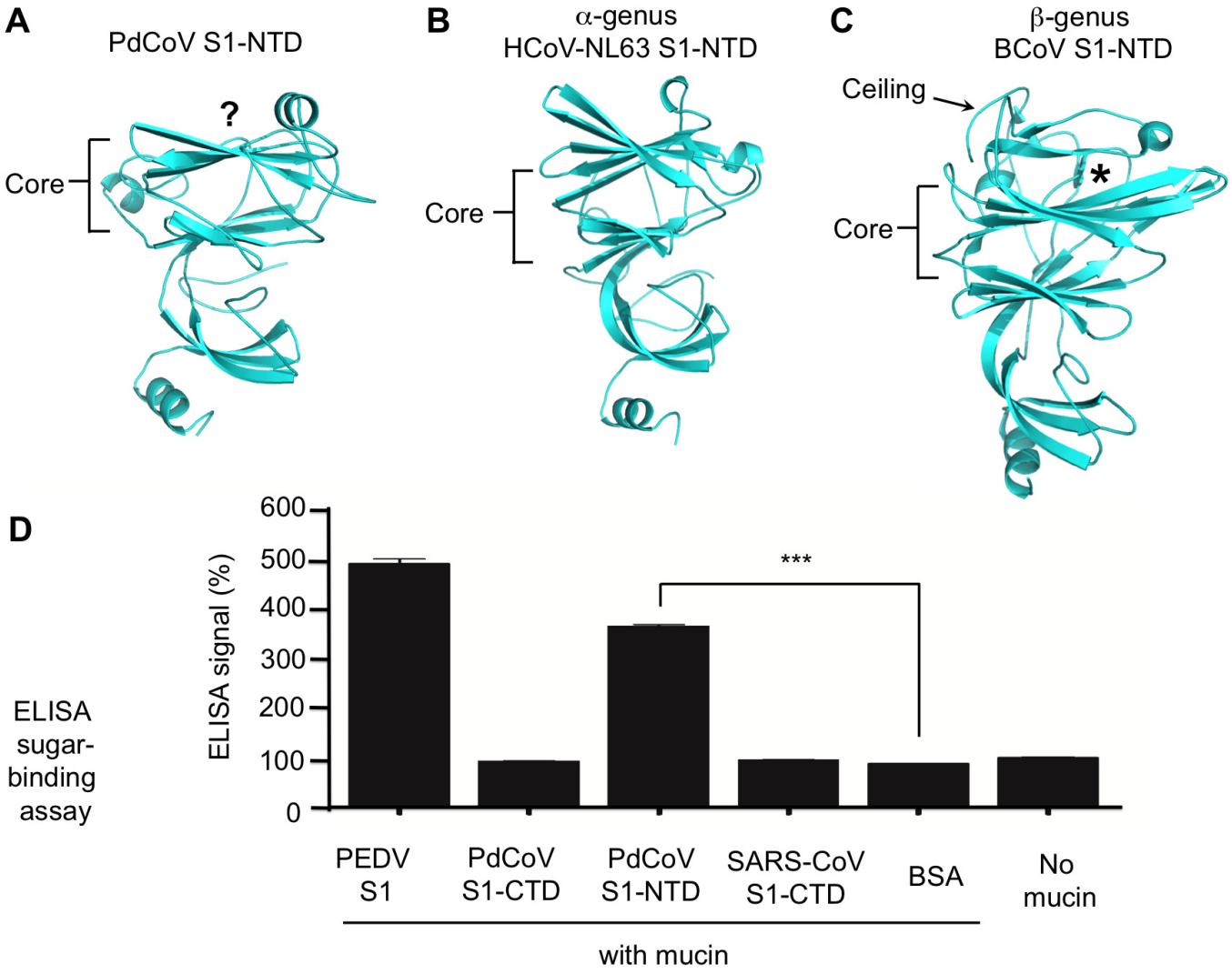
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



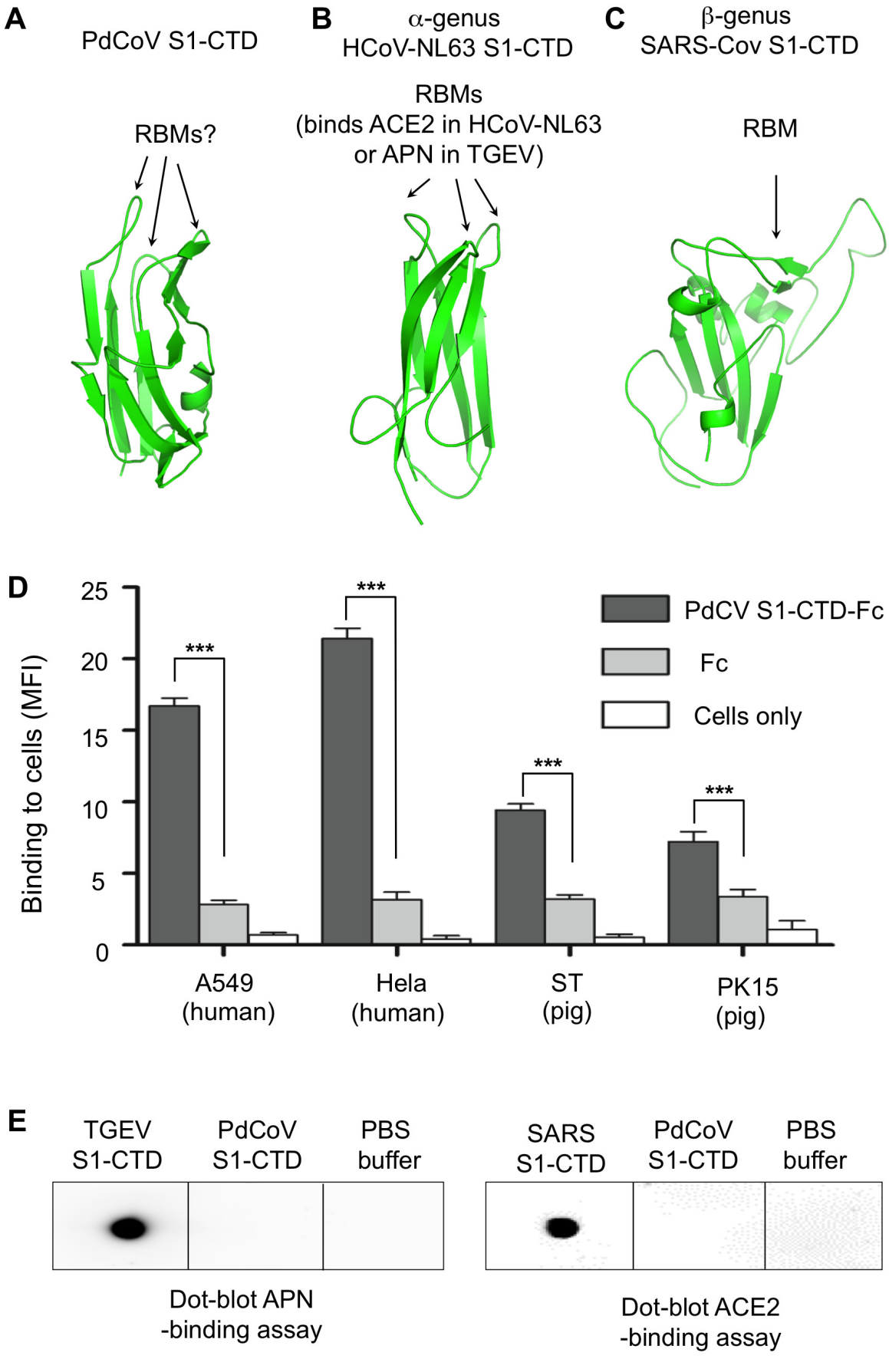
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



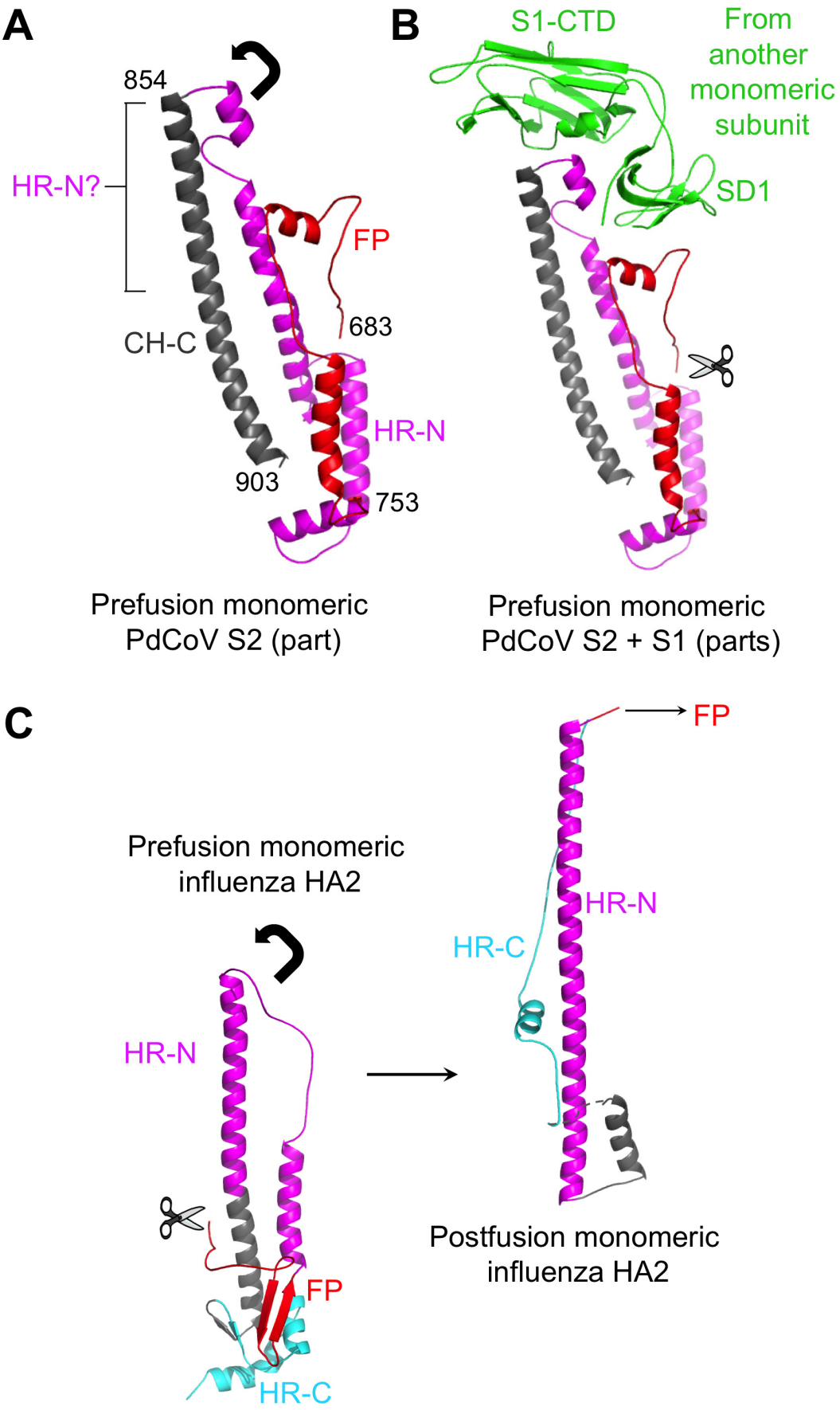
Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology



Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY

Accepted Manuscript Posted Online

Journal of Virology

JVI

Journal of Virology

Downloaded from http://jvi.asm.org/ on October 31, 2017 by FUDAN UNIVERSITY