Proteolytic activation of the SARS-CoV-2 spike S1/S2 site: a re-evaluation of furin cleavage

Tiffany Tang[1], Javier A. Jaimes[2], Miya K. Bidon[1], Marco R. Straus[2], Susan Daniel[1*], Gary R. Whittaker[2*]

[1]Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY, 14853, USA

[2]Department of Microbiology and Immunology, Cornell University, Ithaca, NY, 14853, USA

*Corresponding authors:

Gary Whittaker: grw7@cornell.edu

Susan Daniel: sd386@cornell.edu

Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) uses its spike (S) protein to mediate viral entry into host cells. Cleavage of the S protein at the S1/S2 and/or S2' site is known to activate the S protein for viral entry, which can occur at either the cell plasma membrane or the endosomal membrane. Previous studies show that SARS-CoV-2 has a unique insert at the S1/S2 site that can be cleaved by furin, which expands viral tropism to lung cells. Here, we analyze the presence of a furin S1/S2 site in related CoVs and offer thoughts on the implications of SARS-CoV-2's unique insert on its origin. We also utilized viral pseudoparticles to study the impact of the S1/S2 cleavage on infectivity. Our results demonstrate that S1/S2 pre-cleavage is essential for plasma membrane entry into Calu-3 cells, a model lung epithelial cell line, but not for endosomal entry Vero E6 cells, a model cell culture line, and that other proteases in addition to furin are responsible for processing SARS-CoV-2 S1/S2.

Introduction

The 21st century has seen the rise of pathogenic strains of human coronaviruses (CoVs) causing major public health concerns, first with the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002[1], then the Middle East respiratory syndrome coronavirus (MERS-CoV) first emerged in 2012[2], and now, with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 causes the disease syndrome known as COVID-19[3], now classified as a pandemic with global reach and devastation.

CoV host cell entry is mediated by its spike (S) glycoprotein, a large transmembrane protein that decorates the virus particle[4]. The S protein is demarcated into two domains, the S1 domain, which is the receptor binding domain, and the S2 domain, which contains the membrane fusion machinery. There are two cleavage events associated with S-mediated membrane fusion[5]. The first is a priming cleavage that occurs

at the interface of the S1/S2 region (S1/S2) for some coronaviruses, and the second is the obligatory triggering cleavage that occurs within the S2 region (S2')[5]. The priming cleavage generally converts the S protein into a fusion competent form, by enabling the S protein to better bind receptors or expose hidden cleavage sites[5]. The triggering cleavage initiates a series of conformational changes that enable the S protein to harpoon into the host membrane for membrane fusion[5].

There are a variety of proteases capable of priming and triggering CoV S proteins. Depending on which protease are available, CoV can fuse with either the plasma membrane or the endosomal membrane[6]. SARS-CoV was found to utilize the transmembrane protease TMPRSS2 to fuse at the plasma membrane surface[7]. However, TMPRSS2 expression is limited to respiratory cell lines, and in TMPRSS2 negative cell lines, SARS-CoV utilized endosomal cathepsin L to fuse to the endosomal membrane[8]. MERS-CoV was also found to utilize TMPRSS2 and cathepsin L with one major difference. The MERS-CoV S1/S2 boundary contains an RSVR insert that can be recognized by furin or related proprotein convertases (PC), proteases commonly found in the secretory pathway of most cell lines[9,10]. During the S maturation process, the S protein can be cleaved by furin/PCs[9,10]. Thus, MERS-CoV particles harbor cleaved S protein and it was observed that the S1/S2 pre-cleavage was crucial for MERS-CoV[10], but not SARS-CoV[11], to infect via the plasma membrane route. However, the S1/S2 cleavage was not a requirement for MERS-CoV endosomal pathway infection[10].

For SARS-CoV-2, early studies showed that the S1/S2 junction contained an insert with two additional basic residues, P-R-R-A (R – arginine, A- alanine), that was not present in SARS-CoV or its closest bat ancestor viruses[12,13]. This insert forms a P-R-R-A-R sequence and while it does contain the minimum furin recognition motif, R-X-X-R, it is unusual and diverges from the preferred and canonical R-X-K/R-R motif. Compared to the canonical motif, the residues at the P2 and P3 location for SARS-CoV-2 S1/S2 are reversed, with R at the P3 instead of the P2 location, and an A at the P2 instead of the P3 location[14,15]

(**Figure 1**). Intriguingly, the only other known example of this insert on FurinDB[16], a database of furin substrates, is found in proaerolysin, a bacterial toxin, and was determined to be activated by furin[17].

Indeed, early studies suggest that the SARS-CoV-2 is also processed by furin since SARS-CoV-2 harbors a cleaved S protein, likely due to furin processing at the S1/S2 site[18–20]. Similar to what was observed for MERS-CoV, this S1/S2 cleavage was determined to be a prerequisite for TMPRSS2 activation at the S2' for SARS-CoV-2 infection in respiratory cell lines, such as Calu-3[21,22]. Furthermore, like MERS-CoV, SARS-CoV-2 can also utilize cathepsin L in the endosomal pathway in TMPRSS2-negative cell lines, such as Vero E6[20].

In this study, we sought to better characterize the S1/S2 cleavage site of SARS-CoV-2 S and raise some intriguing possibilities for how the site might have emerged. Using pseudoparticles, we investigated the impact of this S1/S2 cleavage for successful infection of cells via the plasma membrane and endosomal route. Our data suggests that the S1/S2 cleavage is essential for TMPRSS2 mediated plasma membrane entry, but not for cathepsin L mediated endosomal entry, and that furin may not be the only protease responsible for the S1/S2 cleavage event.

Results and Discussion

Furin-cleavage predictions of the SARS-CoV-2 S1/S2 site. We used the PiTou[23] and ProP[24] cleavage prediction tools to analyze the likelihood that the S1/S2 site is processed by furin, with positive scores (PiTou) and 0.5 (ProP) indicating furin cleavage. ProP predicts furin cleavage sites based on networks derived from experimental data, whereas PiTou uses a combination of a hidden Markov model and biological knowledge-based cumulative probability score functions to characterize a 20 amino acid motif from P14 to P6' that reflects the binding strength and accessibility of the motif in the furin binding pocket. Thus, the PiTou algorithm has been reported to be more sensitive and specific than ProP for predicting

furin cleavage[23]. Both algorithms agree with each other in predicting furin cleavage, strengthening the predictions (**Figure 2**), but since the PiTou algorithm is more sensitive and specific, we decided to focus on the PiTou values for further analysis.

The PiTou algorithm predicts that the SARS-CoV-2 S1/S2 site (score: 9.2) can be cleaved by furin, whereas other lineage B CoVs that have been proposed as SARS-CoV-2 precursors, such as SARS-CoV, RaTG13, ZC45, ZXC21 and RmYN02, cannot be cleaved by furin. These predictions have been confirmed experimentally as recombinant furin was shown to process the purified SARS-CoV-2 S protein, generating appropriate cleavage products[25], whereas no cleavage products were detected when processing the SARS-CoV S.

Notably, the PiTou scores of traditionally accepted furin cleavage sites, such as those found in influenza H5N1 (score: 13.6) and HCoV-HKU1 (score: 14.6), are much higher than for SARS-CoV-2, suggesting that while SARS-CoV-2 S1/S2 site has increased furin recognition when compared to other lineage B-betaCoVs, the site is not optimal for furin cleavage. The modest SARS-CoV-2 PiTou score raises question about the origin and evolution of the virus, possibly suggesting that it does not represent an “insertion” into the viral genome as is typically assumed[26].

In this scenario, it is worth noting that in betaCoV lineage C, MERS-CoV (score: 5.2) is barely above threshold. In fact, within the betaCoV lineage C, there is a range of PiTou scores for the closest known bat virus ancestors to MERS-CoV; i.e. BatCoV-HKU4 (score: <0) to BatCoV-HKU5 (score: 10.3). With this in mind, the MERS-CoV S1/S2 site can be seen as either a gain-of-furin-cleavage mutation from BatCoV-HKU4, or equally a loss-of-furin-cleavage from BatCoV-HKU5. Thus, for SARS-CoV-2, a simple gain-of-furin-cleavage may not necessarily be the cause of increased disease, leaving open the potential for a yet-to-be discovered ancestor virus in bats with a robust furin cleavage site (i.e. score approx. 13-15), but one that may not have gained the capacity to bind a human receptor and for which a robust furin cleavage site may be been down-regulated or expanded to other proteases with less specific recognition motifs. These

suggested parallels are informed by betaCoV lineage C where MERS-CoV (score: 5.2) and BatCoV-HKU4 bind hDPP4, but with BatCoV-HKU5 (score: 10.3) unable to bind this receptor[27]. Overall, while it is generally believed that the SARS-CoV-2 S1/S2 site is a gained furin insert, we consider it is important to consider alternative explanations—especially as the search continues to determine the origin of SARS-CoV-2.

MLV pseudoparticles as a system to study SARS-CoV-2 entry. To assess the functional importance of the S1/S2 site for SARS-CoV-2 entry, we utilized viral pseudoparticles. These particles consist of a murine leukemia virus (MLV) core and are decorated with the viral envelope protein to accurately recapitulate the entry steps of their native counterpart[28]. These particles also contain a luciferase reporter that integrates into the host cell genome upon successful infection and drives the cells to produce luciferase, which is quantifiable. We and others have used MLV-based pseudoparticles widely to study CoV entry pathways[9,19,29]. MLV pseudoparticles exhibiting the SARS-CoV-2 S protein (SARS-CoV-2pp), or the SARS-CoV S protein (SARS-CoVpp) were generated alongside positive control particles containing the vesicular stomatitis virus G protein (VSVpp) or negative control particles (Δenvpp) lacking envelope proteins.

Since coronaviruses can enter via the plasma membrane or via endosomes, we chose to infect cell lines representative of each pathway, as the entry mechanism can be highly cell-type dependent[30]. We utilized the Calu-3 and the Vero E6 cell lines for these studies, as they are commonly used cell lines for studying SARS-CoV and SARS-CoV-2 plasma membrane and endosomal entry, respectively. As expected, VSVpp (positive control) infected Vero E6 and Calu-3 cells with several orders of magnitude higher luciferase units than the values reported with Δenvpp infection (negative control) (**Figure 3A, 3B**). This confirms that the envelope protein is driving infection, and not the particle itself. In the case of SARS-CoVpp and SARS-CoV-2pp, both particles are infectious as they drive luciferase production several orders of magnitude higher than Δenvpp. (**Figure 3A, 3B**).

It is noted that SARS-CoVpp are more infectious than SARS-CoV-2pp. To determine the cause of this difference, we analyzed the particles via Western blot to visualize S content. For the SARS-CoV-2pp, we detected a band at 85 kDa and for SARS-CoVpp, a strong band at 185 kDa (**Figure 3C**)**.** The 85 kDa band corresponds to the S2 segment of the S protein following S1/S2 cleavage. The 185 kDa band for the SARS-CoVpp corresponds to uncleaved S protein. We observe that the SARS-CoV S band is more intense than the SARS-CoV-2 S2 band, despite both particles showing similar intensities for the MLVp30 loading control band. As a result, we infer that the SARS-CoVpp incorporated more S protein than SARS-CoV-2pp, resulting in the higher infectivity observed for SARS-CoVpp transduction.

Use of the dec-RVKR-CMK protease inhibitor to produce SARS-2pp with uncleaved S. To examine the functional role of S priming by furin/PC proteases, we needed to produce SARS-CoV-2pp expressing uncleaved S protein. This can be accomplished by adding in an appropriate protease inhibitor to producer cells to prevent S1/S2 cleavage during biogenesis. We chose dec-RVKR-CMK because it has been shown to inhibit furin/PCs, preventing S1/S2 cleavage[21,29,31]. Indeed, we previously showed that addition of dec-RVKR-CMK to producer cells generates MLV pseudotyped particles that harbor full-length MERS-CoV S protein (MERS-CoVpp)[29] and we observe a similar trend with SARS-CoV-2pp generated from cells treated with dec-RVKR-CMK (**Figure 4C,** lanes 1 and 4).

The dec-RVKR-CMK-treated (uncleaved) particles were used to transduce either the Vero E6 or Calu-3 cells to observe the impact of the S1/S2 cleavage on infection. In Vero E6 cells, the uncleaved SARS-CoV-2pp were 30-fold more infectious than their cleaved counterparts, suggesting that uncleaved particles result in more infectious particles, and that the S1/S2 pre-cleavage is not required or hinders infection via the endosomal pathway (**Figure 4A**). However, in Calu-3 cells, the opposite trend was observed; the uncleaved particles were significantly less infectious than their cleaved counterparts (**Figure 4B**)**,** suggesting that

S1/S2 pre-cleavage is essential[21,22]for plasma membrane entry and increased uncleaved S incorporation cannot compensate.

Thus, we were curious if we could restore the infection of the uncleaved particles by treating them with exogeneous furin, which has been previously shown to process purified SARS-CoV-2 S protein[25]. In Calu-3, we observed a 2-fold increase in infection when uncleaved particles were treated with exogeneous furin, but with cleaved particles, we did not observe any statistically significant increase with furin treatment (**Figure 4B**). In Vero E6 cells, for both types of particles, we did not observe a statistically significant increase in infection with furin treatment. (**Figure 4A**).

Since furin treatment only modestly increased infectivity in all cases observed, we wanted to visualize how efficiently furin was processing the SARS-CoV-2 S. For cleaved particles, exogeneous furin was able to fully reduce the faint full-length S band into S2 (**Figure 4C**, lanes 1 and 2). For uncleaved particles, furin treatment had no observable impact, as the intensity of the full-length S bands are the same and there are no additional bands shown between furin treated and non-treated particles (**Figure 4C**, lanes 3 and 4). Additional bands at >185 kDa likely corresponding to dimeric and trimeric S[20].

In addition, the intensity of the uncleaved SARS-CoV-2pp S bands are greater than the intensity of the cleaved SARS-CoV-2pp S bands and this trend was observed with MERS-CoVpp S[29]. This likely suggests that the increased infectivity we observed with the uncleaved SARS-CoV-2pp is due to greater S incorporation, though the S is uncleaved.

Due to the limited ability of exogenous furin to rescue and cleave SARS-CoV-2 S, we re-evaluated the role of furin in processing the S1/S2 site. Since dec-RVKR-CMK can also inhibit a number of furin related PCs, the effects that have been observed with dec-RVKR-CMK particles could have resulted from inhibition of other proteases and not furin. Therefore, we produced particles in the presence of alpha1-PDX, a potent and more selective furin inhibitor[32] than dec-RVKR-CMK. Infectivity results of these particles in Vero E6

cells that a highly furin specific inhibitor failed to recapitulate the high level of enhancement provided by dec-RVK-CMK at all tested inhibitor concentrations (**Figure 5**). This indicates that furin itself may not in fact be the only active protease processing the S1/S2 site, as is generally assumed. Other PCs within the secretory pathway may also assist with the S1/S2 cleavage and have been observed to cleave SARS-CoV-2 S1/S2 peptide[33].

Lastly, we wanted to investigate if dec-RVKR-CMK inhibition affects viral proteins that do not feature a furin S1/S2 site, such as SARS-CoV. dec-RVKR-CMK treatment had no significant impact on SARS-CoV S mediated infection of Vero E6 and Calu-3 cells (**Figure 6A and 6B**), suggesting that dec-RVKR-CMK impacts on SARS-CoV-2 S is due to inhibiting the S1/S2 pre-cleavage and not due to some general effect on protein expression. Treatment of SARS-CoVpp with exogeneous furin also yielded no difference in protein conformation (**Figure 6C**).

Overall, the results support observations[21,34] that the role of the SARS-CoV-2 S1/S2 site is to expand viral tropism to lung cells. Cleaved S1/S2 site is crucial for the SARS-CoV-2 to be subsequently cleaved at the S2' location by TMPRSS2 for immediate plasma membrane entry in respiratory cells. Without the S1/S2 pre-cleavage, SARS-CoV-2 would be endocytosed, and due to low cathepsin L expression in respiratory endosomes[10], coupled with expression of antiviral restriction factors in endosomes[34], SARS-CoV-2 would not effectively infect via respiratory endosome. If SARS-CoV-2 is infecting TMPRSS2 negative cells, it can utilize endosomal cathepsin L, an ubiquitous protease generally found throughout mammalian cells[35], to activate the S protein[20], though at undetermined sites. However, for cathepsin L activation, S1/S2 pre-cleavage is not required, and our results indicate that preventing this cleavage increases infectivity of SARS-CoV-2. This may be connected to recent work showing that S1/S2 pre-cleavage reduces S thermal stability[25]. Interestingly, the S1/S2 site activation appears to also have a role in the immune response

against SARS-CoV-2, as it has been shown that SARS-CoV-2 S with a deleted S1/S2 loop can provide better protective immune response than with the S1/S2 loop[36].

The role of furin in activating the S1/S2 site was also investigated. As protease inhibitors commonly employed in cell entry studies, such as dec-RVKR-CMK, can also inhibit other PCs in addition to furin[31], it is difficult to ascertain the impact of individual proteases. While furin likely can process the SARS-CoV-2 site, our data suggests that other PCs are also involved, since Western blots show poor processing of uncleaved S upon addition of purified furin and also the use of a highly selective furin inhibitor has a modest impact on infectivity. As we consider protease-based inhibitors in treating COVID-19, especially those targeting furin[37], it is important to thoroughly consider evaluate the role of these proteases to be certain that the relevant proteases are targeted.

Acknowledgements

This work was funded by the National Institute of Health research grant R01AI35270 and Fast Grant, Mercatus Center. TT is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441 and the Samuel C. Fleming Family Graduate Fellowship. We would like to thank Hector Aguilar-Carreno for helpful input and thank members of the Daniel and Whittaker groups as well as the Eliezer and Weinstein groups at Weill Cornell for helping with discussions.

Methods and Materials

Predicted structural modeling. S protein models were built using UCSF Chimera (v.1.14, University of California) through modeler homology tool of the Modeller extension (v.9.25, University of California) as

described[13]. SARS-CoV and SARS-CoV--2 S models were built based on the SARS-CoV S structure (PDB No. 5X58).

Cells, plasmids, and reagents. HEK293T (accession no: CRL-11268, ATCC), Vero E6 (accession no: CRL-1586, ATCC) , and Calu-3 (accession no: HTB-55) cells were maintained at 37°C and 5% CO_2 incubator and cultivated in Dulbecco's Modified Eagle Medium (Cellgro) supplemented with 10% Hyclone FetalClone II (GE) and 10 mM HEPES (Cellgro) (DMEMc). Cells were passaged using Dulbecco's Phosphate-Buffered Saline (DPBS) and Trypsin EDTA 1x (Cellgro). The pCMV-MLV-gagpol murine leukemia virus (MLV) packaging construct, the pTG-luc luciferase reporter, pCAGGS/VSV-G, pCAGGS, and pcDNA/SARS-S plasmids are as previously described [28]. The pcDNA/SARS2-S was a generous gift from Veesler lab[19].

Protease inhibitor decanoyl-RVKR-CMK (dec-RVKR-CMK) was purchased from Tocris and resuspended in sterile water to 10 mM. Recombinant furin was purchased from New England Biolabs. Human alpha-1 PDX (alpha1-PDX) recombinant protein was purchased from ThermoFisher Scientific.

Furin prediction calculations. Prop: CoV sequences were inputted into the ProP 1.0 Server hosted here: cbs.dtu.dk/services/ProP/ Pitou: CoV sequences were analyzed using the PiTou V3 software freely available via request here[23].

Pseudoparticle production. Pseudotyped virus were produced from published protocols with minor modifications [28]. Briefly, HEK293T cells were seeded to 30% confluency. Cells were transfected with pTG-luc (600 ng), pCMV-MLV-gagpol (800 ng), and the respective viral envelope protein (600 ng) using polyethylenimine for 48 hours. Supernatants were then harvested, centrifuged, clarified, and stored in -80°C aliquots.

For + dec-RVKR-CMK particles, 7.5 µL of dec-RVKR-CMK was added to cells immediately after transfection and boosted with an additional 7.5 µL 24 hours later (final concentration: 75 µM).

For alpha1-PDX particles, indicated concentrations were added to cells immediately after transfection.

Pseudoparticle assays. Infection assays were as previously described with minor modifications[28]. Briefly, target cells were seeded to confluency. Cells were washed with DPBS infected with particles on a rocker in an incubator with rocker for 1.5 hours. Cells were supplemented with DMEMc and incubated for 72 hours. Cells were lysed and luciferase activity was assessed using the Luciferase Assay System and the Glomax 20/20 luminometer (Promega).

Exogeneous protease treatment. 1 mL pseudoparticles were pelleted using a TLA-55 rotor with an Optima-MAX-E ultracentrifuge (Beckman Coulter) for 2 hours at 42,000 rpm at 4°C. Particles were resuspended in 30 µL DPBS buffer supplemented with 20 mM HEPES, 0.2 mM Cacl2, 0.2 mM β-mercaptoethanol. Particles were treated with 6 U recombinant furin for 3 hours at 37°C.

Western blot analysis of pseudoparticles. 2 mL pseudoparticles were pelleted and resuspended as described above, but in 20 uL. Sodium dodecyl sulfate (SDS) loading buffer and DTT were added to samples and heated at 65°C for 20 minutes. Samples were separated on NuPAGE Bis-Tris gel (Invitrogen) and transferred on polyvinylidene difluoride membranes (GE). SARS-CoV-2 and SARS-CoV S was detected using a rabbit polyclonal antibody that recognizes the S2 domain (Cat: 40590-T62, Sinobiological) and an AlexaFluor 488 anti-rabbit secondary antibody. MLV protein was detected using a mouse monoclonal antibody that recognizes the MLV p30 protein (Cat: ab130757, abcam) and an AlexaFluor 488 anti-mouse secondary antibody. Bands were detected using the ChemiDoc Imaging software (Bio-Rad).

References

1. Zhong, N. *et al.* Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* **362**, 1353–1358 (2003).

2. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E. & Fouchier, R. A. M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *N. Engl. J. Med.* **367**, 1814–1820 (2012).

3. Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).

4. Belouzard, S. *et al.* Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **4**, 1011–1033 (2012).

5. White, J. M. & Whittaker, G. R. Fusion of Enveloped Viruses in Endosomes. *Traffic* **17**, 593–614 (2016).

6. Heald-Sargent, T. & Gallagher, T. Ready, Set, Fuse! The Coronavirus Spike Protein and Acquisition of Fusion Competence. *Viruses* **4**, 557–580 (2012).

7. Matsuyama, S. *et al.* Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).

8. Simmons, G. *et al. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry*. *Proceedings of the National Academy of Sciences of the United States of America* vol. 102 www.pnas.orgcgidoi10.1073pnas.0505577102 (2005).

9. Millet, J. K. & Whittaker, G. R. 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 (2014).

10. Park, J. E. *et al.* Proteolytic processing of middle east respiratory syndrome coronavirus spikes expands virus tropism. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12262–12267 (2016).

11. Reinke, L. M. *et al.* Different residues in the SARS-CoV spike protein determine cleavage and activation by the host cell protease TMPRSS2. *PLoS One* **12**, (2017).

12. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).

13. Jaimes, J. A., André, N. M., Chappie, J. S., Millet, J. K. & Whittaker, G. R. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. *J. Mol. Biol.* **432**, 3309–3325 (2020).

14. Henrich, S., Lindberg, I., Bode, W. & Than, M. E. Proprotein convertase models based on the crystal structures of furin and kexin: Explanation of their specificity. *J. Mol. Biol.* **345**, 211–227 (2005).

15. Henrich, S. *et al.* The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. *Nat. Struct. Biol.* **10**, 520–526 (2003).

16. Tian, S., Huang, Q., Fang, Y. & Wu, J. FurinDB: A Database of 20-residue furin cleavage site motifs, substrates and their associated drugs. *Int. J. Mol. Sci.* **12**, 1060–1065 (2011).

17. Abrami, L. *et al.* The pore-forming toxin proaerolysin is activated by furin. *J. Biol. Chem.* **273**, 32656–32661 (1998).

18. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor Article SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2

and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 1–10 (2020).

19. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* (2020) doi:10.1016/j.cell.2020.02.058.

20. Ou, X. *et al.* Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620 (2020).

21. Hoffmann, M., Kleine-Weber, H. & Pöhlmann, S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol. Cell* **78**, 779-784.e5 (2020).

22. Bestle, D. *et al.* TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. alliance* **3**, (2020).

23. Tian, S., Huajun, W. & Wu, J. Computational prediction of furin cleavage sites by a hybrid method and understanding mechanism underlying diseases. *Sci. Rep.* **2**, (2012).

24. Duckert, P., Brunak, S. & Blom, N. Prediction of proprotein convertase cleavage sites. *Protein Eng. Des. Sel.* **17**, 107–112 (2004).

25. Wrobel, A. G. *et al.* SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Biol.* **27**, 763–767 (2020).

26. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **89**, 44–48 (2020).

27. Yang, Y. *et al.* Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12516–12521 (2014).

28. Millet, J. K. *et al.* Production of pseudotyped particles to study highly pathogenic coronaviruses in a biosafety level 2 setting. *J. Vis. Exp.* **2019**, e59010 (2019).

29. Straus, M. R. *et al.* Ca 2+ ions promote fusion of Middle East respiratory syndrome coronavirus with host cells and increase infectivity . *J. Virol.* **94**, 426–446 (2020).

30. Tang, T., Bidon, M., Jaimes, J. A., Whittaker, G. R. & Daniel, S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Res.* **178**, 104792 (2020).

31. Matsuyama, S. *et al.* Middle East Respiratory Syndrome Coronavirus Spike Protein Is Not Activated Directly by Cellular Furin during Viral Entry into Target Cells. *J. Virol.* **92**, (2018).

32. Jean, F. *et al.* α1-Antitrypsin Portland, a bioengineered serpin highly selective for furin: Application as an antipathogenic agent. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7293–7298 (1998).

33. Jaimes, J. A., Millet, J. K. & Whittaker, G. R. Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site. *iScience* **23**, (2020).

34. Peacock, T. P. *et al.* The furin cleavage site of SARS-CoV-2 spike protein is a key determinant for transmission due to enhanced replication in airway cells. *bioRxiv* 2020.09.30.318311 (2020) doi:10.1101/2020.09.30.318311.

35. Mcgrath, M. E. *THE LYSOSOMAL CYSTEINE PROTEASES*. *Annu. Rev. Biophys. Biomol. Struct* vol. 28 www.annualreviews.org (1999).

36. Dorottya Laczkó, A. *et al.* A Single Immunization with Nucleoside-Modified mRNA Vaccines Elicits Strong Cellular and Humoral Immune Responses against SARS-CoV-2 in Mice ll A Single Immunization with Nucleoside-Modified mRNA Vaccines Elicits Strong Cellular and Humoral Immune Responses against SARS-CoV-2 in Mice. *Immunity* **53**, (2020).

37. Cheng, Y.-W. *et al.* Furin Inhibitors Block SARS-CoV-2 Spike Protein Cleavage to Suppress Virus Production and Cytopathic Effects. *Cell Rep.* **0**, 108254 (2020).

Figure 1

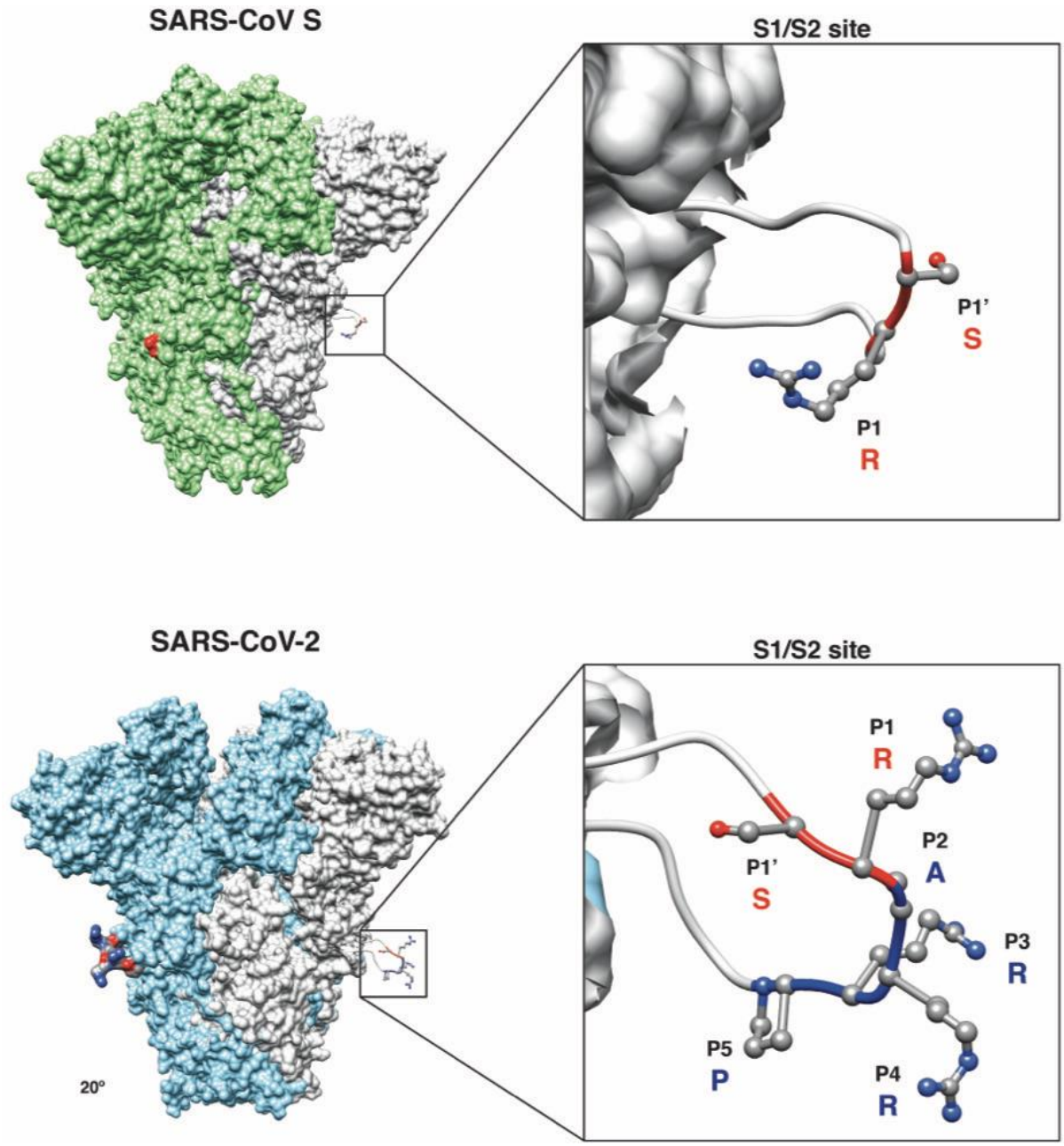

Figure 1: Predicted structural model of the SARS-CoV and SARS-CoV-2 S proteins. (**Inset**) Magnification of S1/S2 site with conserved R and S residues (red ribbon) and the unique four amino acid insertion P-R-R-A for SARS-CoV-2 (blue ribbon) are shown. The P's denote the position of that amino acid from the S1/S2 cleavage site, with P1-P5 referring to amino acids before the cleavage site and P1' referring to amino acids after the cleavage site.

Figure 2

Virus	S1/S2 Sequence	Furin Score PiTou	Furin Score ProP
Betacoronavirus			
Lineage A			
HCoV-HKU1	747-YNSPSSSSSRRKRR|SISASY-766	**+14.634**	**0.918**
Lineage B			
SARS-CoV	654-AGICASYHTVSLLR|STSQKS-673	-5.167	0.123
SARS-CoV-2	672-ASYQTQTNSPRRAR|SVASQS-691	**+9.196**	**0.620**
Bat-RmYN02*	667-GAGVCASYNSPAAR|VGTNSI-686	-5.085	0.111
Bat-CoV RaTG13	668-AGICASYQTQTNSR|SVASQS-687	-4.672	0.151
Bat-CoV ZC45	645-AGICASYHTASILR|STSQKA-664	-5.333	0.153
Bat-CoV ZXC21	644-AGICASYHTASILR|STGQKA-663	-5.333	0.170
Lineage C			
MERS-CoV	738-LPDTPSTLTPRSVR|SVPGEM-757	**+5.155**	**0.563**
BatCoV-HKU4	736-GQSLCAVPPVSTFR|SYSASQ-755	-5.14	0.229
BatCoV-HKU5	732-LCAIPPTTSSRVRR|ATSGAS-751	**+10.259**	**0.822**
Influenza A			
H5N1	322-LRNTPQRERRRKKR|GLFGAI-341	**+13.59**	**0.808**

Figure 2. **Furin cleavage score analysis of CoV S1/S2 cleavage sites.** CoV S sequences were analyzed using the ProP[1] 1.0 and PiTou[2] 3.0 furin prediction algorithm, generating a score with green numbers indicating predicted furin cleavage and red numbers indicating no predicted furin cleavage. Purple lines denote the position of the predicted S1/S2 cleavage site. Basic arginine (R) and lysine (K) residues are highlighted in blue. Sequence numbers refer to position of amino acids within the spike protein. *For Bat-RmYN02, sequence number was determined from S alignment with SARS-CoV-2 S using Geneious. Sequences corresponding to the S1/S2 region of HCoV-HKU1 (AAT98580.1), SARS-CoV (AAT74874.1), SARS-CoV-2 (QHD43416.1), Bat-CoVRaTG13 (QHR63300.2), Bat-SL-CoVZC45 (AVP78031.1), Bat-SL-CoVZXC21 (AVP78042.1), MERS-CoV (AFS88936.1), BatCoV-HKU4 (YP_001039953.1), BatCoV-HKU5 (YP_001039962.1), and Influenza A/Chicken/Hong Kong/822.1/01/H5N1 (AF509026.2) were obtained from GenBank. Sequences corresponding to the S1/S2 and S2' region of RmYN02 (EPI_ISL_412977) were obtained from GISAID.

Figure 3

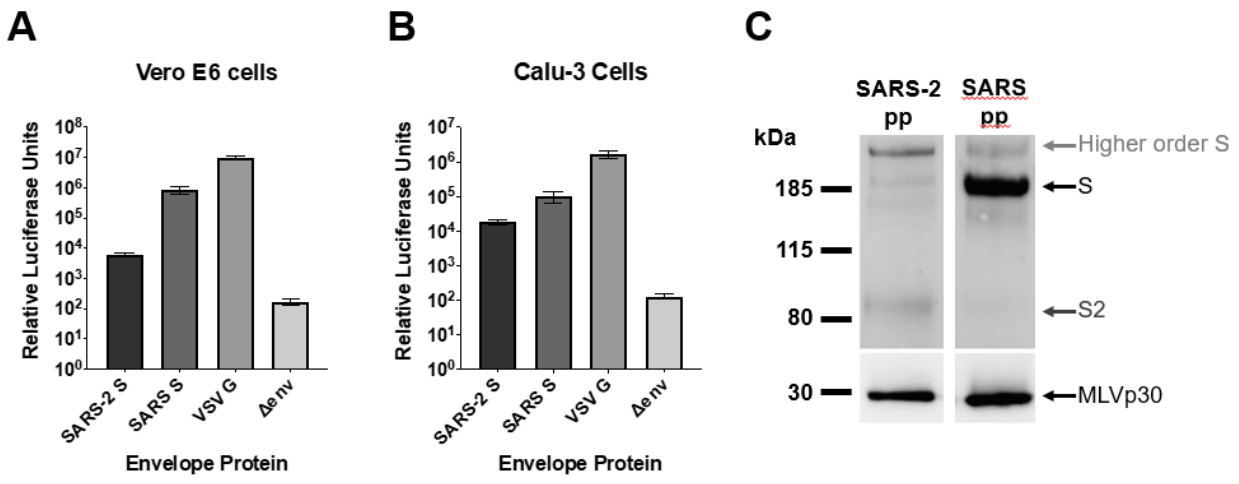

Figure 3: Characterization of MLVpp system. (**A and B**) MLVpp infectivity in Vero E6 and Calu-3 cells. Cells were infected with MLVpps exhibiting the SARS-CoV-2 S, SARS-CoV S, VSV G or no envelope protein and assessed for luciferase activity. Error bars represent the standard error measurements of three biological replicates (n=3). (**C**) Western blot analysis of SARS-2pp and SARSpp S content. S was detected using a rabbit antibody against the SARS-CoV-2 S2 region that cross reacts against the SARS-CoV S. MLV content was detected using a mouse antibody against MLVp30. Image cropped from a singular Western blot and lanes are reshown in Figure 3C and 5C.

Figure 4

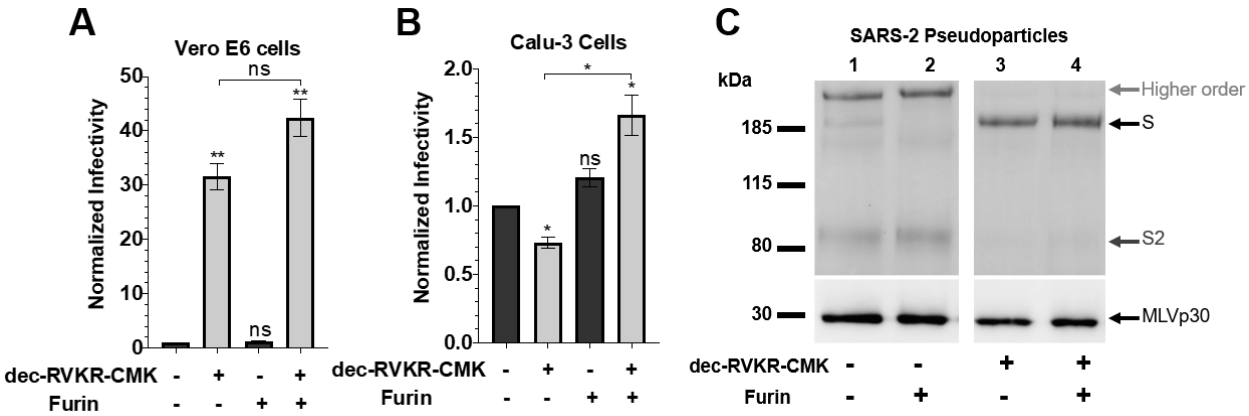

Figure 4: Impact of dec-RVKR-CMK on MLVpps. + dec refers to MLVpp produced in HEK293T treated with 75 µM of dec-RVKR-CMK at the time of transfection. + Fur refers to particles treated with 6 U of recombinant furin for 3 hours at 37°C. (**A and B**) +/- dec-RVKR-CMK and +/- Fur SARS-CoV-2pp were used to infect Vero E6 and Calu-3 cells. Infectivity was normalized to the – dec – furin condition in each cell line. Error bars represent the standard error measurements of three biological replicates (n=3). Statistical analysis was performed using an unpaired student's t test to the – dec – furin condition, unless otherwise noted. *, P <0.5, **, P <0.05, ns, non-significant, P >0.5. (**C**) Western blot analysis of SARS-CoV-2pp produced in +/- dec for protein content to complement the infection conditions. Image cropped from a singular Western blot.

Figure 5

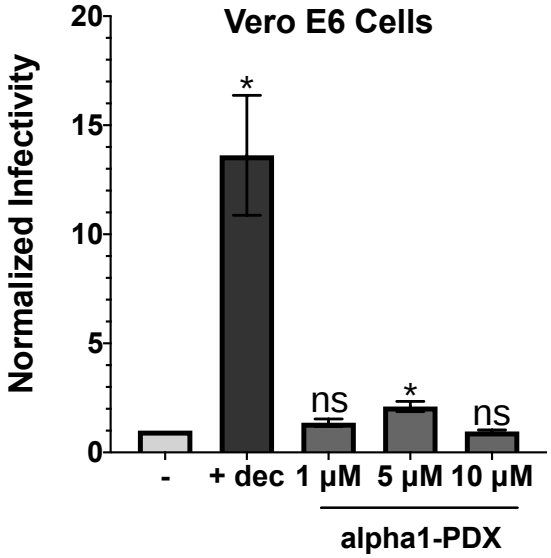

Figure 5: Impact of furin specific inhibitor alpha1-PDX on particle production. Inhibitor was applied at the indicated concentration to producer HEK293T cells. Particles were used to infect Vero E6 cells. Error bars represent the standard error measurements of four biological replicates (n=4). Statistical analysis was performed using an unpaired student's t test compared to - condition. *, P <0.5, ns, non-significant, P >0.5.

Figure 6

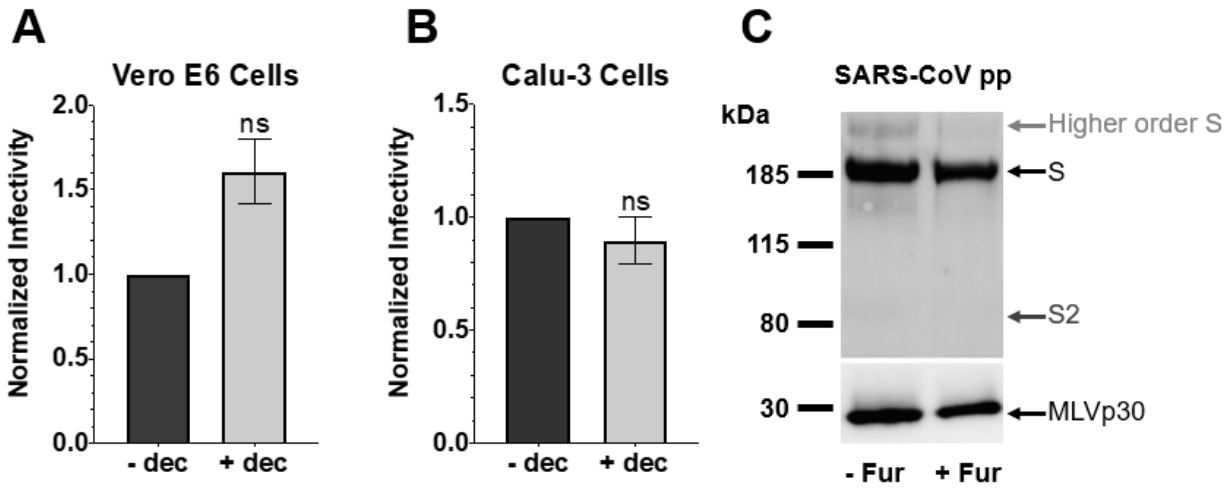

Figure 6: Impact of dec-RVKR-CMK inhibitor on SARS-CoV production, which lacks the furin S1/S2 site. + dec refers to MLVpp produced in HEK293T that were treated with 75 µM of dec-RVKR-CMK at the time of transfection. + Fur refers to particles treated with 6 U of recombinant furin for 3 hours at 37°C. (**A and B**) Particles were used to infect Vero and Calu-3 cells and infectivity is normalized to the – dec condition. Error bars represent the standard error measurements of three biological replicates (n=3). Statistical analysis was performed using an unpaired Student's t test. ns, non-significant, P >0.5. (**C**) Western blot analysis of incorporated S in SARS-CoVpp.