

Copyright  
by  
Michelle N. Vu  
2023

**The Dissertation Committee for Michelle N. Vu Certifies that this is the approved  
version of the following dissertation:**

**Spike Processing and Protease Usage Effects  
on SARS-CoV-2 Pathogenesis**

**Committee:**

---

Vineet D. Menachery, PhD,  
Mentor

---

Scott C. Weaver, PhD

---

Shinji Makino, DVM, PhD

---

Andrew L. Routh, PhD

---

Chad E. Mire, PhD

Seema S. Lakdawala, PhD

---

Dean, Graduate School

**Spike Processing and Protease Usage Effects  
on SARS-CoV-2 Pathogenesis**

**by**

**Michelle N. Vu, BS**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy in Microbiology and Immunology**

**The University of Texas Medical Branch**

**March, 2023**

## **Dedication**

To Charles, my everything who I don't thank enough. You've taken on so much throughout this whole process. I don't know how I would have done this without you.

## **Acknowledgements**

To Dr. Vineet Menachery, I couldn't have asked for a better PhD experience. I can say proudly that I became confident and competent in my abilities under your mentorship. You have always given me excellent advice on experiments, career, and life in general. Thank you for the opportunity to be your graduate student. You are special.

Thank you to all of the members of the Menachery lab: Dr. Kumari Lokugamage, Dr. Bryan Johnson, Dr. Craig Schindewolf, R. Elias Alvarado, Dr. Birte Kalveram, Leah Estes, Alyssa McLellan, Jordon Murray, and my rotation students Yani Ahearn, William Meyers and Angelica Morgan. Your support, advice, and general lab fun have made my experience never boring. I especially want give special thanks to Dr. Kumari Lokugamage, who I cannot thank enough for her support in generating viruses and troubleshooting.

Thank you to my committee: Dr. Scott Weaver, Dr. Shinji Makino, Dr. Andrew Routh, Dr. Chad Mire, and Dr. Seema Lakdawala. Your collective experience in virology, molecular biology, animal models, and transmission were invaluable in the success of this dissertation. Thank you to Dr. Lynn Soong for your support in the Microbiology and Immunology program. Thank you to Aneth Zertuche who has gone above and beyond to help us graduate students. Thank you to Dr. Kenneth Plante, Dr. Jessica Plate, Dionna Scharton and the rest of the WRCEVA for their help in animal experiments and training me. Thank you to Dr. Andrew Routh and his lab for lending their expertise on NGS. Thank you to Dr. William Russell and the mass spectrometry core for lending their expertise. Thank you to Dr. David Walker and Dr. Patricia Croquet-Valdes for providing histopathology analysis. Thank Dr. Krishna Narayanan for his scientific advice.

Thank you to my former mentors Dr. Bert Semler, Dr. Alexis Bouin, Hung Nguyen and My Phuong Tran for teaching me the basics of molecular virology and getting my career started.

Thank you to my friends Dorothea Morris, Jacob Stockton, and Mandi Feinberg who have been with me through the ups and downs and forever reminding me of land gulls and other silly moments. I can't imagine going through this with any other people. Thank you to Charles Luu and my friends and family back home for their constant love and support.

# **Spike Processing and Protease Usage Effects on SARS-CoV-2 Pathogenesis**

Publication No. \_\_\_\_\_

Michelle N. Vu, PhD

The University of Texas Medical Branch, 2023

Supervisor: Vineet D. Menachery

The emergence of SARS-CoV-2 and its subsequent variants have ignited a multitude of studies on the spike protein and its domains. While most studies focus on the impact of receptor binding domain changes, mutations in the C-terminus of S1 (CTS1) have largely been overlooked. The SARS-CoV-2 CTS1 contains various features that are unusual to sarbecoviruses – the furin cleavage site (FCS) and the QTQTN motif. In this dissertation, we demonstrate that the CTS1 is more complex than previously thought. We have established the aspects of cleavage site efficiency, differential protease usage, loop length, and post translational modifications that all contribute to SARS-CoV-2 pathogenesis. Using our reverse genetics system, we generated several infectious clones to examine the various features of the CTS1:  $\Delta$ QTQTN demonstrated the importance of loop length and glycosylation. PQQA established that the integrity of the FCS is necessary. Omicron CTS1 mutants YKH (H655Y, N679K, and P681H) and N679K demonstrated the interplay between mutations with different contributions and expanded on the impact of post translational modifications. These studies show that alterations to any of these aspects

greatly affect the pathogenicity of the virus. Together, we demonstrate that the mutations in the CTS1 are key determinants of coronavirus pathogenicity and should be especially considered in surveillance for the next coronavirus outbreak.

This dissertation is based in part on the previously published manuscripts listed below. Reuse of works listed below is with permission from authors and publishers under the Creative Commons license. Figures from the listed manuscripts that have been used in full or part in this dissertation include a citation within the text. Dissertation sections which have been based on the works listed below include:

Chapter 1: Introduction – Coronavirus Spike Protein

Chapter 2: Beyond the Furin Cleavage Site

Chapter 4: Discussion and Future Directions

1. Vu, M. N. & Menachery, V. D. Binding and entering: COVID finds a new home. *PLOS Pathogens* 17, e1009857 (2021). <https://doi.org:10.1371/journal.ppat.1009857>
2. Vu, M. N. et al. QTQTN motif upstream of the furin-cleavage site plays a key role in SARS-CoV-2 infection and pathogenesis. *Proceedings of the National Academy of Sciences* 119, e2205690119 (2022). <https://doi.org:doi:10.1073/pnas.2205690119>

# TABLE OF CONTENTS

List of Tables .....	xiii
List of Figures .....	xiv
List of Abbreviations .....	xviii
Chapter 1: Introduction – SARS-CoV-2 Spike Protein .....	20
Spike Protein and Viral Entry .....	20
S2 Subunit.....	20
Receptor Binding Domain .....	21
N-terminal Domain .....	21
C-terminal of the S1 Subunit .....	22
Balancing Transmissibility and Pathogenicity .....	23
Location, location, location.....	25
Finding the key. ....	26
Cutting the lock.....	27
Making itself at home. ....	28
Bringing it home. ....	30
Complexities of the CTS1.....	30
Chapter 2: Beyond the Furin Cleavage Site.....	35
Introduction.....	35
Results.....	36
$\Delta$ QTQTN attenuates viral replication <i>in vitro</i> . ....	36
$\Delta$ QTQTN attenuates disease, but not replication <i>in vivo</i> .....	37
$\Delta$ QTQTN reduces spike processing and entry.....	39
Glycosylation of the QTQTN motif contributes to spike processing. ....	40
Integrity of the FCS sequence is required by SARS-CoV-2.....	42
Discussion.....	43
Methods .....	45
Cell culture.....	45
Viruses .....	45
<i>In vitro</i> infection .....	45

Protease inhibitor treatment .....	46
Competition assay and next generation sequencing analysis. ....	46
Virion purification and western blotting.....	46
Hamster infection studies.....	47
Histology.....	48
Structural modeling.....	48
Transcriptomics.....	48
Nanoflow-LC-MS/MS glycopeptide mapping .....	49
 Chapter 3: Interplay between Omicron CTS1 mutations.....	 78
Introduction.....	78
Results.....	79
H655Y, N679K, and P681H together increase viral replication and spike processing. ....	79
N679K mutation attenuates SARS-CoV-2 infection. ....	80
N679K mutation results in decreased spike protein expression. ....	81
Discussion.....	83
Methods .....	86
Cell Culture.....	86
Viruses .....	86
<i>In vitro</i> Infection .....	87
Plaque Assay.....	87
Focus Forming Assay .....	87
Hamster Infection.....	88
Virion Purification .....	88
Western Blot .....	89
Spike HexaPro Cloning and Transfection.....	90
Structural Modeling .....	90
 Chapter 4: Discussion and Future Directions .....	 105
Impact of the CTS1 .....	105
Spike Cleavage Efficiency.....	105
Protease Usage.....	107

Spike Stability.....	109
Transmission.....	110
CTS1 in Emerging Zoonotic Coronaviruses.....	111
References.....	113
VITA.....	122
Education .....	123
Publications.....	123

## List of Tables

Table 2.1. Quantification of Nanoflow-LC-MS/MS analysis of QTQTN glycosylation mutants.....	70
---------------------------------------------------------------------------------------------	----

**No table of figures entries found.**

## List of Figures

Figure 1.1. Coronavirus spike.....	31
Figure 1.2. Coronavirus entry mechanisms. ....	32
Figure 1.3. Transmission versus pathogenesis.....	33
Figure 1.4. Comparison of coronavirus tropism. ....	34
Figure 2.1. Coronavirus S1/S2 cleavage site. ....	51
Figure 2.2. SARS-CoV-2 $\Delta$ QTQTN cloning schematic.....	52
Figure 2.3. Structural modeling of SARS-CoV-2 $\Delta$ QTQTN.....	53
Figure 2.4. Plaque morphology of $\Delta$ QTQTN SARS-CoV-2.....	54
Figure 2.5. <i>In vitro</i> replication of $\Delta$ QTQTN.....	55
Figure. 2.6. Vero E6 competition assay WT: $\Delta$ QTQTN .....	56
Figure 2.7. <i>In vivo</i> characterization of SARS-CoV-2 $\Delta$ QTQTN in golden Syrian hamsters. ....	57
Figure 2.8. Histopathology of $\Delta$ QTQTN infected hamsters.....	58
Figure 2.9. Viral replication of $\Delta$ QTQTN infected hamsters. ....	59
Figure 2.10. Clustering analysis of $\Delta$ QTQTN SARS-CoV-2 infected hamster transcriptomics.....	60

Figure 2.11. Comparison of gene regulation in $\Delta$ QTQTN SARS-CoV-2 infected hamster transcriptomics. ....	61
Figure 2.12. Spike processing of $\Delta$ QTQTN. ....	62
Figure 2.13. Inhibition of SARS-CoV-2 entry.....	63
Figure 2.14. $\Delta$ QTQTN inefficiently uses TMPRSS2. ....	64
Figure 2.15. $\Delta$ QTQTN SARS-CoV-2 replication in TMPRSS2-expressing Vero E6. ....	65
Figure 2.16. O-linked glycosylation of SARS-CoV-2 loop.....	66
Figure 2.17. Cloning schematic of QTQTN glycosylation mutants. ....	67
Figure 2.18. Detection of glycosylation on QTQTN. ....	68
Figure 2.19. Nanoflow-LC-MS/MS analysis of QTQTN motif. ....	69
Table 2.1. Quantification of Nanoflow-LC-MS/MS analysis of QTQTN glycosylation mutants.....	70
Figure 2.20. Plaque morphology of QTQTN glycosylation mutants.....	71
Figure 2.21. <i>In vitro</i> replication of QTQTN glycosylation mutants. ....	72
Figure 2.22. Spike processing of QTQTN glycosylation mutants.....	73
Figure 2.23. Glycosylation of QTQTN motif influences protease usage. ....	74
Figure 2.24. <i>In vitro</i> replication of PQQA.....	75

Figure 2.25. Spike processing of PQQA.....	76
Figure 2.26. Characteristics of SARS-CoV-2 S1/S2 cleavage site for efficient infection. ....	77
Figure 2.27. Comparison of CTS1 across SARS-CoV-2 variants. ....	91
Figure 2.28. Structural modeling of spike CTS1 residues mutated in Omicron. ....	92
Figure 2.29. Cloning schematic of SARS-CoV-2 YKH. ....	93
Figure 2.30. Plaque morphology of YKH.....	94
Figure 2.31. <i>In vitro</i> replication of YKH. ....	95
Figure 2.32. Spike processing of YKH purified virions. ....	96
Figure 2.33. Structural modeling of spike CTS1 residues mutated in Omicron adjacent to O-linked glycosylation. ....	97
Figure 2.34. Cloning schematic of SARS-CoV-2 N679K.....	98
Figure. 2.35. Plaque morphology of N679K.....	99
Figure 2.36. <i>In vitro</i> replication of N679K. ....	100
Figure 2.37. <i>In vivo</i> characterization of SARS-CoV-2 N679K in golden Syrian hamsters. ....	101
Figure 2.38. Spike processing and expression in N679K purified virions. ....	102
Figure 2.39. Intracellular spike expression of N679K.....	103

Figure 2.40. Expression of exogenous N679K Spike HexaPro.....104

## List of Abbreviations

ACE2	Angiotensin converting enzyme 2
APN	Aminopeptidase N
CFR	Case fatality rate
COVID-19	Coronavirus disease 2019
CTS1	C-terminal of the S1 subunit
CoV	Coronavirus
DDP4	Dipeptidyl peptidase-4
DPI	Day post infection
FCS	Furin cleavage site
hCoV	Human coronavirus
HR	Heptad repeat region
HPI	Hours post infection
HPT	Hours post transfection
MERS-CoV	Middle Eastern respiratory syndrome coronavirus
NTD	N-terminal domain
PFU	Plaque forming unit
RBD	Receptor binding domain
SA	Sialic acid
SARS-CoV	Severe acute respiratory syndrome coronavirus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
S/N ratio	Spike to nucleocapsid ratio
TMPRSS2	Transmembrane serine protease 2

VOC Variant of concern  
WT Wild type

# Chapter 1:Introduction – SARS-CoV-2 Spike Protein

## SPIKE PROTEIN AND VIRAL ENTRY

The coronavirus (CoV) spike protein is a vital protein for establishing infection in a host. Existing as a trimer on the virion surface, the spike is composed of two regions: globular heads responsible for receptor binding and a stalk used in membrane fusion (**Figure 1.1**)<sup>1</sup>. The spike protein itself is divided into the S1 and S2 subunits. The S1 subunit is largely responsible for attaching and binding to the host receptor, containing the N-terminal domain (NTD) and receptor binding domain (RBD). Additionally, the S1 subunit has the C-terminus of the S1 subunit (CTS1) that is involved in spike activation. The S2 subunit harbors the conserved fusion machinery necessary for viral entry.

Upon attachment and binding to the receptor, the spike is processed at two cleavage sites by host proteases. The first cleavage event occurs at the S1/S2 cleavage site located in the CTS1 and is thought to remove steric hindrance to the second cleavage site, priming the spike for membrane fusion. The second cleavage event at the S2' cleavage site reveals the fusion peptide in the S2 subunit and triggers membrane fusion. Depending on the entry route, the spike protein will utilize different host proteases for processing: serine proteases at the cell surface or cathepsins in endosomes (**Figure 1.2**). Once membrane fusion occurs, the viral genetic material is released into the cytosol to be used in translation of viral proteins or to generate more genomic material for packaging into progeny virions.

## S2 SUBUNIT

The CoV spike is a class 1 fusion protein whose S2 subunit harbors the fusion machinery: the fusion peptide, two heptad repeat regions (HR1 and HR2) and the

transmembrane domain. Due to this, the S2 subunit is the most conserved domain of the spike protein, making it popular target for vaccine design. Few mutations have risen in this subunit with only 4 mutations occurring commonly before Omicron – T716I, D950N, S982A, and D1118H <sup>2</sup>. Structural analyses of these mutations suggest that they may contribute to the stability of the “up” and “down” RBD conformations and of the spike trimer as a whole; however, they may also be offsetting mutations that cause instability <sup>3-5</sup>. Altogether, mutations in the S2 subunit may contribute to maintaining the balance between spike stabilization and the ability to change conformations.

### **RECEPTOR BINDING DOMAIN**

As the name entails, the RBD is involved in receptor recognition, attachment, and binding. The RBD compatibility to host receptors has been considered as the primary barrier to emergence for zoonotic CoVs with adaptation required for infection of mice by most strains <sup>6-11</sup>. It is therefore unsurprising for frequent mutations to occur in this region with each passing SARS-CoV-2 variant. RBD mutations may increase binding to the human ACE2 receptor, increasing infectivity. The RBD is also the target of vaccine designs due to its involvement in receptor recognition and binding. Mutations may alter recognition by the immune system as most neutralizing antibodies target the RBD. In the case of Omicron, which emerged with more than 30 spike mutations, 15 of those mutations are located in the RBD leading to increased neutralizing titers and vaccine breakthrough <sup>12,13</sup>. Collectively, the RBD is vital to the virus with its role in receptor and immune recognition.

### **N-TERMINAL DOMAIN**

Although its functions have not been clearly defined, the NTD has been associated

with receptor binding and immune evasion. As glycoproteins, CoV spikes have been found to interact with glycans like sialic acids either through the RBD or the NTD<sup>14-16</sup>. Mutations in the sialic acid binding site of the NTD have been found to enhance or reduce virion attachment by modulating sialylated glycan interaction and viral surfing across the membrane<sup>16,17</sup>. This modulation may contribute to tropism changes in cell type and regions (i.e., upper versus lower airways)<sup>16,18</sup>. Regarding immunity, about ~20% of neutralizing antibodies are NTD-specific that are either highly potent against infectious virus or are glycan-dependent<sup>5,19,20</sup>. The SARS-CoV-2 NTD has been found to harbor an antigenic supersite to which the majority of NTD-specific antibodies bind<sup>21</sup>. As such, mutations frequently occur in and surrounding the supersite, contributing to immune evasion<sup>5,21</sup>. Shifting away from receptor binding and immune evasion, a frequent NTD mutation  $\Delta$ H69/V70 has been shown to facilitate spike cleavage and incorporation as well as cell-to-cell spread, demonstrating another application of the NTD<sup>22</sup>. Taken together, the NTD is a multifaceted domain with contributions to CoV infectivity and immune evasion.

### **C-TERMINAL OF THE S1 SUBUNIT**

The function of the CTS1 has been primarily attributed to protease usage to process the spike at the S1/S2 cleavage site. While the RBD has been considered as the primary barrier to zoonotic emergence, chimeric viruses exchanging the SARS-CoV spike for the closely related bat CoVs WIV-1 and SHC014 spikes were capable of establishing infection in primary human epithelial cells, despite receptor compatibility predictions<sup>9,23</sup>. However, although the SHC014-spike SARS-CoV chimera replicated similarly to SARS-CoV *in vitro*, the virus resulted in attenuated disease *in vivo*<sup>9</sup>. The capacity of bat CoV spikes to readily infect human cells without adaptation but still result in attenuated disease

demonstrates that receptor compatibility is not the primary indicator for potential epidemic strains. Studies exchanging the spike domains between SARS-CoV and SHC014-CoV revealed that the CTS1 chimera attenuated disease *in vivo* whereas the NTD-RBD chimera was comparable to SARS-CoV (data not shown). Additionally, the SHC014-SARS chimera demonstrated changes in protease usage, a function inherent to the CTS1 (data not shown). Although the CTS1 is more conserved compared to the NTD and RBD, the slight changes to the CTS1 have impactful effects on CoV pathogenesis.

### **BALANCING TRANSMISSIBILITY AND PATHOGENICITY**

A novel coronavirus, subsequently named severe acute respiratory coronavirus 2 (SARS-CoV-2), emerged in December 2019 and has caused ~177 million cases and ~3.8 million deaths worldwide (global case fatality rate [CFR] of 2.2% varying in individual countries)<sup>24</sup>. SARS-CoV-2 is not the first coronavirus to have caused a pandemic, as both severe acute respiratory coronavirus (SARS-CoV) and Middle Eastern respiratory coronavirus (MERS-CoV) have done so previously in 2002 and 2012, respectively<sup>25</sup>. While the CFRs of SARS-CoV and MERS-CoV are greater than SARS-CoV-2 (11% and 34%, respectively), the earlier endemics did not reach the same scale (8400 and ~2500 cases, respectively). In contrast, human coronaviruses (hCoV-229E, -NL63, -OC43, and -HKU1) produce limited, common cold symptoms, yet have been endemic globally for decades<sup>26</sup>. Upon examination, there appears to be an inverse relationship between virus transmissibility and pathogenicity – the less pathogenic hCoVs have robust transmission, the highly pathogenic SARS-CoV and MERS-CoV have limited transmission, and the current SARS-CoV-2 appears to have established a rare balance between the two.

A similar trend between transmissibility and pathogenicity can be seen with influenza viruses. Like hCoVs, seasonal influenza viruses are common globally, characterized by high transmissibility, low pathogenicity with ~20% of the population infected, and a CFR less than 1% annually <sup>27</sup>. Analogous to SARS-CoV and MERS-CoV, high pathogenicity influenza strains such as H5N1 and H7N9 influenza viruses have high CFR (53% and 39%, respectively), but low transmission with ~1550 and ~850 cases since their introduction <sup>28,29</sup>. The 1918 H1N1 influenza virus is comparable to SARS-CoV-2 as it achieved a balance between transmissibility and pathogenicity, with 1/3 of the world's population infected and ~50 million deaths (CFR  $\approx$  2.5%) <sup>30</sup>. More recently, the H1N1 subtype caused another pandemic in 2009 with relatively robust transmissibility, although with lower pathogenicity (CFR < 0.02%) <sup>31</sup>.

The idea of a trade-off between virulence and transmission of pathogens has long been proposed with little empirical supporting evidence <sup>32,33</sup>. However, such a relationship is clear when studying coronaviruses and influenza viruses (**Figure 1.3**). Both groups of viruses have strains capable of great harm (SARS-CoV, MERS-CoV, H5N1, H7N9) yet are limited in human transmission. In contrast, strains that cause mild disease (hCoVs & seasonal influenza) have been endemic globally. Then, there are viruses (SARS-CoV-2 and 1918 H1N1) that have established a key balance, capable of achieving devastating pathogenicity, but minimal loss in transmissibility. While only a few viruses have attained this balance, it is imperative to understand how these viruses emerged in order to head off the next pandemic. Here, we focus on the start of infection, examining the impact of receptor binding and protease activation on virus tropism, transmissibility, and pathogenicity.

### **Location, location, location.**

The location in which the virus replicates along the respiratory tract guides how the virus is able to transmit and its pathogenicity. Traditionally, viruses that bind in the upper airways exhibit lower pathogenicity compensated with high transmissibility; the inverse is true for those that bind lower in the airways. For example, coronaviruses associated with the common cold (hCoV-229E, -NL63, -OC43 and -HKU1) replicate primarily in the upper respiratory tract including the nasal and oropharyngeal areas <sup>34</sup>. Testing for these less virulent hCoVs typically finds positives from upper airways samples and rarely finds lower airway infection. In turn, upper airway replication may correspond with greater transmissibility. In contrast, both SARS-CoV and MERS-CoV rarely induce positive tests from upper airway samples <sup>35,36</sup>. Instead, during the SARS-CoV outbreak and the later MERS-CoV emergence, samples from deep in the respiratory tract (bronchiolar lavage and sputum) were critical for diagnosis. These results suggest that the more virulent CoVs induce more replication and damage deep in the lung. The corresponding loss in lung function directly plays a role in the increased lethality associated with each. It also may explain relatively low transmission outside hospital settings and super-spreaders.

For SARS-CoV-2, a balance exists with the infection of both the upper and lower airways. Early on, testing of nasopharyngeal swabs confirmed upper airway infection with SARS-CoV-2; subsequent work found positive samples in saliva and nasal swabs suggesting replication throughout the upper airways <sup>37</sup>. However, lung infection as evidenced by bronchiolar alveolar lavage fluid, sputum, and severe disease demonstrate robust replication in the lung in a large subset of patients <sup>37</sup>. This even distribution in the upper and lower airways highlight and correspond to the transmissibility and pathogenicity

relationship highlighted above (**Figure 1.3**) and may offer a primer for why SARS-CoV-2 and COVID-19 has had such a broad impact on society.

### **Finding the key.**

Location of virus replication is in part determined by receptor recognition and binding. For influenza, the viruses' interactions between hemagglutinin and sialylated glycan receptors have been well characterized<sup>38</sup>. Briefly, sialic acid is linked to galactose through either an  $\alpha 2,3$  (S $\alpha 2,3$ ) or  $\alpha 2,6$  (S $\alpha 2,6$ ) linkage, which are expressed in an inverse gradient along the human respiratory tract – S $\alpha 2,6$  more highly expressed in the upper airways and S $\alpha 2,3$  in the lower airways<sup>38,39</sup>. The hemagglutinin of high pathogenicity avian influenza viruses (e.g. H7N9 and H5N1) preferentially bind to S $\alpha 2,3$ , directing their replication to the lower airways, limiting their transmission, but inducing more severe disease<sup>39-41</sup>. Conversely, human influenza viruses (e.g. H1N1 and H3N2) typically do not cause severe pathogenicity and have been observed to bind more to S $\alpha 2,6$  in the upper airways, therefore affording them greater transmission opportunities<sup>39,41</sup>. Indeed, influenza viruses that are more transmissible via air preferentially replicate in the upper airways<sup>42</sup>.

Coronaviruses bind to their receptors through their spike protein, with different coronavirus spikes being similar at their core, but different enough in their globular head to utilize a variety of receptors. Most coronaviruses bind to membrane ectopeptidases – hCoV-229E to APN, MERS-CoV to DPP4, and hCoV-NL63, SARS-CoV and SARS-CoV-2 to ACE2<sup>43,44</sup>. The receptors for hCoV-OC43 and -HKU1 are unknown, although they interact with acetylated sialic acids<sup>45</sup>. Interestingly, hCoV-NL63, SARS-CoV, and SARS-CoV-2 share the same receptor yet display distinct infection and virulence. A key to this discrepancy may be the receptor binding affinity of their spike protein. Compared

to the SARS-CoV spike protein, hCoV-NL63 spike protein has less affinity to human ACE2 while SARS-CoV-2 spike protein has comparable, or slightly greater affinity<sup>46,47</sup>. ACE2 expression follows a gradient along the respiratory tract with greater expression in the upper airways and decreasing further down<sup>48</sup>. One hypothesis argues that the strength receptor affinity may correlate with receptor abundance in an inverse relationship. On one end of the spectrum, the lower receptor affinity of hCoVs may hamper potential lower respiratory infection where ACE2 is less abundant, resulting in replication primarily in the upper airways where the greater ACE2 expression offers more opportunities for binding. On the opposite end, SARS-CoV replication is predominantly observed in the lower airways, perhaps due to its strong receptor affinity allowing binding to fewer ACE2 when swept deeper. Interestingly, although the SARS-CoV-2 spike protein has similar receptor affinity as that of the SARS-CoV, its infection is observed throughout the respiratory tract and with more replication in the upper airways<sup>48</sup>. Receptor affinity is therefore not solely responsible for where the virus replicates and subsequent effects on transmission and virulence. The spike protein interacts with other host proteins, specifically proteases, to establish successful entry and infection.

### **Cutting the lock.**

While a virus may be able to bind to its receptor, other factors are needed to establish replication. Upon receptor binding, both coronaviruses and influenza viruses undergo an uncoating and fusion process to release the viral genomes. Both the coronavirus spike protein and the influenza virus hemagglutinin are type I fusion proteins; following receptor binding, the proteins are cleaved by host proteases to prime fusion and entry with the host cell. Host proteases, such as trypsin, TMPRSS2, furin and cathepsin, are

ubiquitously expressed throughout the respiratory tract. However, due to variations in the cleavage sequences, not all proteases are compatible. As such, the proteases used by a particular virus for cell entry may be key to its tropism.

Revisiting the previous comparison, hCoV-NL63, SARS-CoV and SARS-CoV-2 use different host proteases despite sharing the hACE2 receptor. While hCoV-NL63 spike protein cleavage has not been studied in depth, TMPRSS2 is involved in its infection as well as SARS-CoV and SARS-CoV-2 infection<sup>49,50</sup>. SARS-CoV is also capable of utilizing cathepsin while hCoV-NL63 cannot, which may explain differences in tropism<sup>51</sup>. Interestingly, SARS-CoV-2 spike protein has a polybasic furin cleavage site, PRRA, directly upstream of the S1/S2 cleavage site, that is not found in other group 2B coronaviruses; however, furin cleavage sites are found in MERS-CoV, hCoV-OC43 and hCoV-HKU1<sup>47,52</sup>. This PRRA site has been found to be necessary for efficient spike cleavage, cell entry, cell-to-cell fusion and overall viral replication and pathogenesis of SARS-CoV-2, with its deletion attenuating replication and disease<sup>53</sup>. The insertion of a furin cleavage site at the S1/S2 cleavage site may contribute to the wider infection range of SARS-CoV-2 compared to SARS-CoV by expanding the pool of proteases priming entry. Similarly, influenza viruses with multibasic cleavage sites, mainly those in the H5 and H7 subtypes, have increased pathogenicity compared to the H1, H2, H3, and H9 subtypes with monobasic sites<sup>54</sup>.

### **Making itself at home.**

Receptor affinity and protease usage are crucial in determining tropism (**Figure 1.4**). Any change affecting those factors would alter how well a virus transmits and its pathogenicity. Importantly, SARS-CoV-2 is making itself more at home in the host as

mutations impacting both receptor binding and proteolytic cleavage have been observed in SARS-CoV-2 variants <sup>55</sup>. While research focus has centered on changes in the receptor-binding domain (RBD), variant mutations outside the RBD have been shown to be important. For example, the D614G variant quickly overtook the original SARS-CoV-2 strain <sup>56-59</sup>; positioned in the C-terminal domain of the spike protein but outside the RBD, D614G influences receptor interaction through the orientation of the spike in the up conformation <sup>57</sup>. Studies using both pseudotyped virus and engineered infectious SARS-CoV-2 harboring D614G have shown that the D614G mutation increases virus infectivity, which in turn would lead to increased transmission <sup>57-60</sup>. The D614G mutation coupled with other RBD changes show that variant mutations impact receptor binding and affinity.

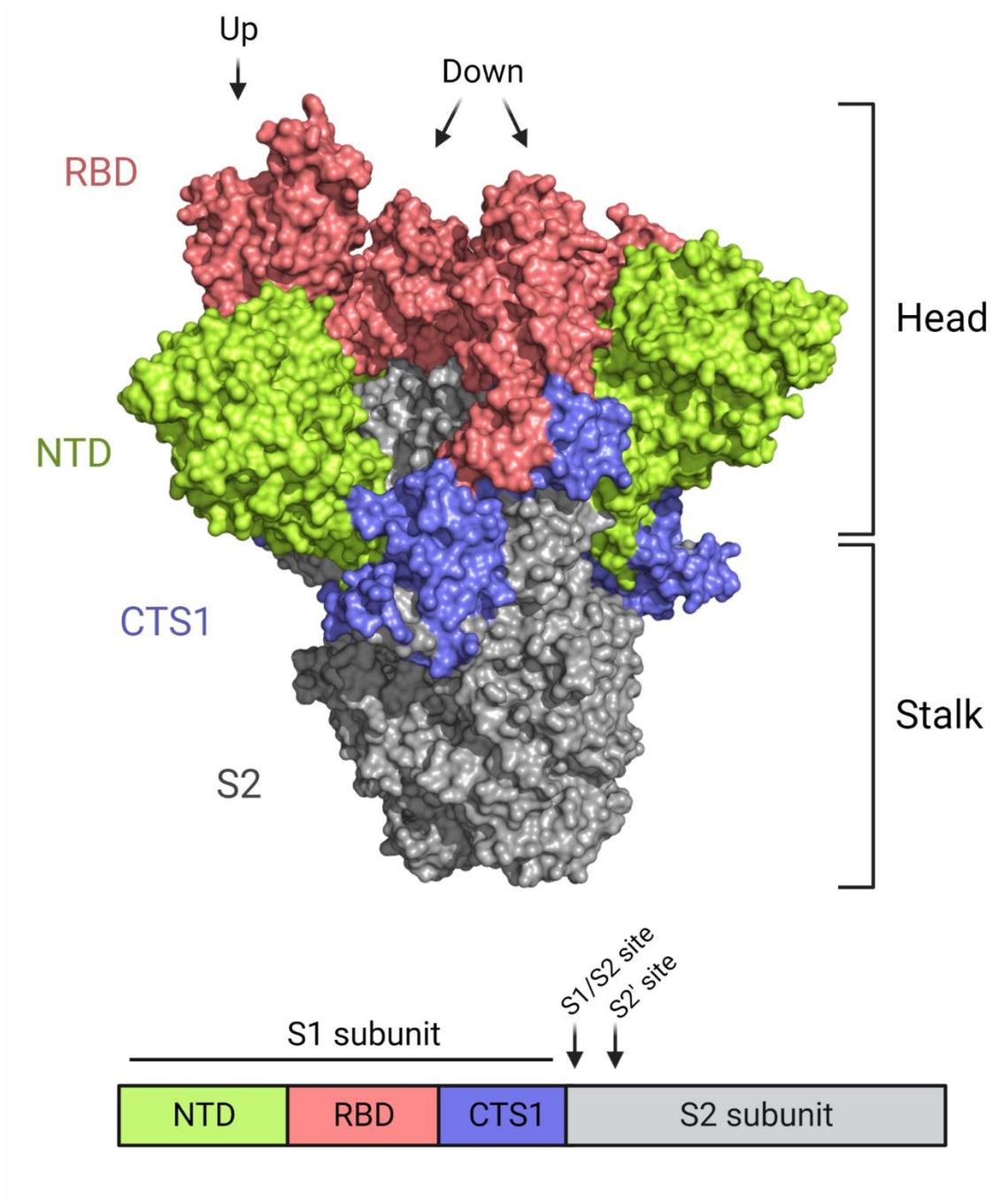
The impact of variant mutations is not limited to receptor binding and affinity. Two variants of concerns, Alpha (B.1.1.7) and Delta (B.1.617), have mutations within the furin cleavage site (P681H and P681R) found in the spike protein of SARS-CoV-2. Both SARS-CoV-2 variants have been observed to have increased transmission based on human cases <sup>61,62</sup>. In addition, our group has demonstrated that B.1.1.7 has a replication and transmission advantage in human airway cells and the hamster models of infection, respectively <sup>63</sup>. Importantly, P681H and P681R in the furin cleavage site of the variant spike proteins may contribute to these differences. Initial findings with the delta variant show increased processing of the spike protein as compared to the original WT SARS-CoV-2 <sup>64</sup>. Converse to studies that removed the furin cleavage site <sup>53</sup>, the presence of P681R mutations shifts processing more clearly to the S1/S2 cleavage product. This may contribute to changes in transmission and pathogenicity. The results also suggest an importance for proteolytic activity in SARS-CoV-2 pathogenesis and transmission.

### **Bringing it home.**

The emergence and continued circulation of SARS-CoV-2 highlight the ongoing threat posed by emerging respiratory viruses. The location, receptor binding, and proteolytic activation of coronaviruses dictates critical elements of transmissibility and pathogenicity. In this regard, SARS-CoV-2 has achieved an unfortunate balance allowing for a world-wide outbreak with parallels to the 1918 H1N1 outbreak from a century earlier. Importantly, understanding how replication location, receptor interactions, and protease activation drive virus emergence and spread is paramount in disrupting and possibly preventing future emergence events.

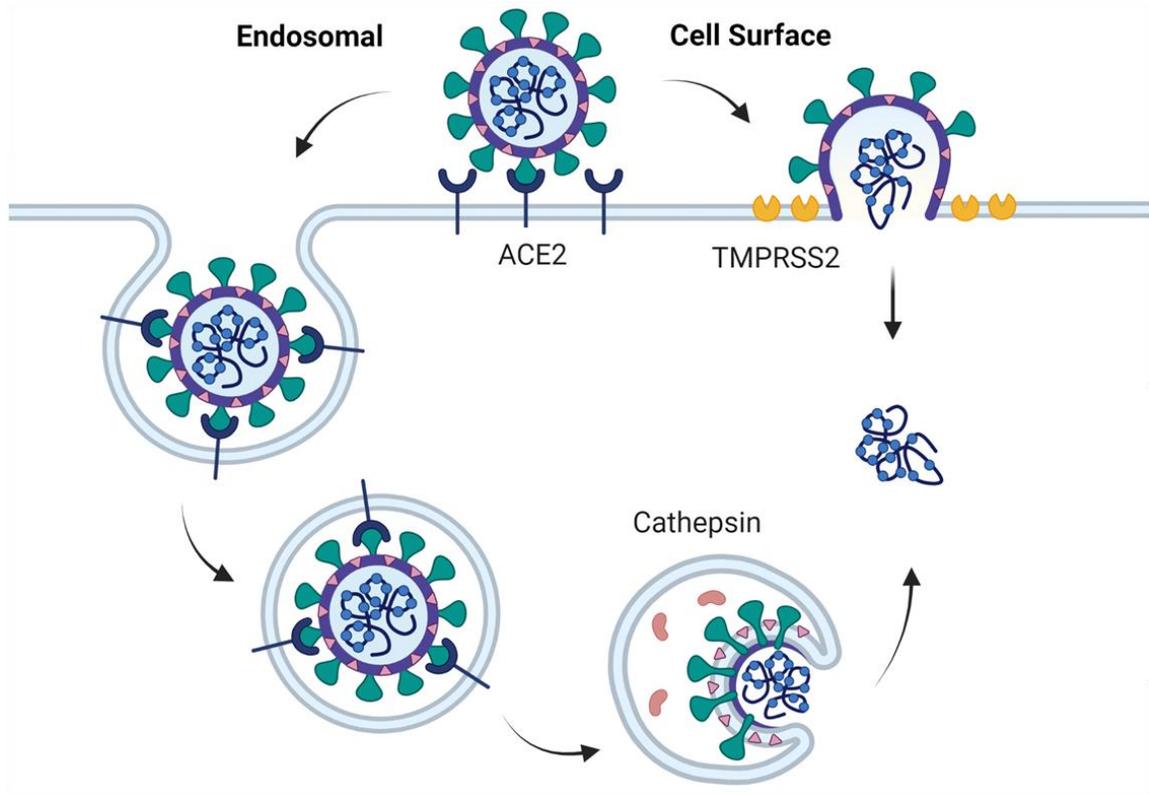
### **COMPLEXITIES OF THE CTS1**

For CoVs, the spike protein has been identified as the key to overcoming the barriers to zoonotic transmission. The RBD and its ability to bind to human receptors as opposed to the animal reservoir receptors have been thought of as primary barriers to zoonotic transmission. However, as another critical step to viral entry, the ability to interact and use host proteases for spike processing at the CTS1 presents as an additional barrier for CoVs to overcome for zoonotic emergence. Therefore, the overall goal of this dissertation is to demonstrate that spike processing and protease usage are key determinants of coronavirus emergence and pathogenesis. This dissertation sets out to demonstrate this through studying crucial mutations in the SARS-CoV-2 CTS1 that contribute to pathogenesis through alterations in spike cleavability, structural changes and post translational modifications. Through these studies, we also demonstrate spike processing and protease usage are distinct and independent processes.



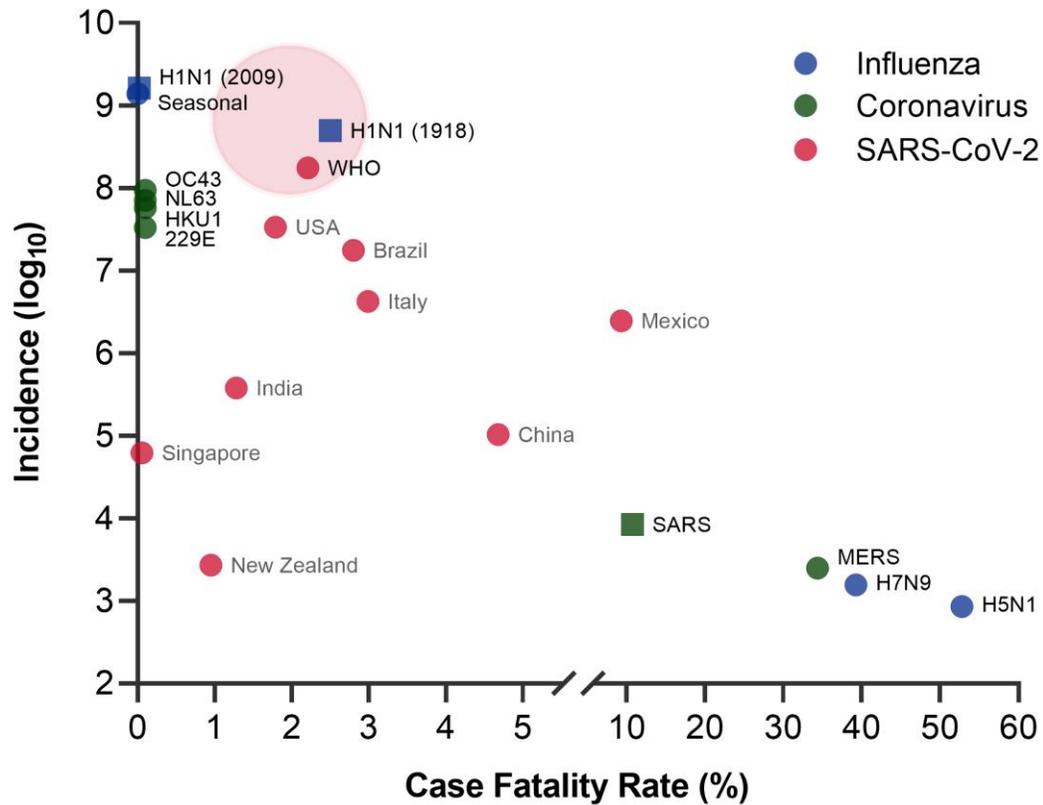
**Figure 1.1. Coronavirus spike.**

The coronavirus spike exists as a trimer on the virion surface with globular heads involved in receptor binding and a stalk harboring the fusion machinery. The spike protein itself is divided into two subunits – the S1 and S2 subunits. The S1 subunit is primarily in the head region, consisting of the N-terminal domain (NTD) (green), receptor binding domain (RBD) (pink), and the C-terminal of the S1 (CTS1) (purple). The S2 subunit (grey) makes up the entirety of the stalk. The spike is processed by host proteases first at the S1/S2 site to prime and then at the S2' site, triggering and revealing the fusion peptide for membrane fusion.



**Figure 1.2. Coronavirus entry mechanisms.**

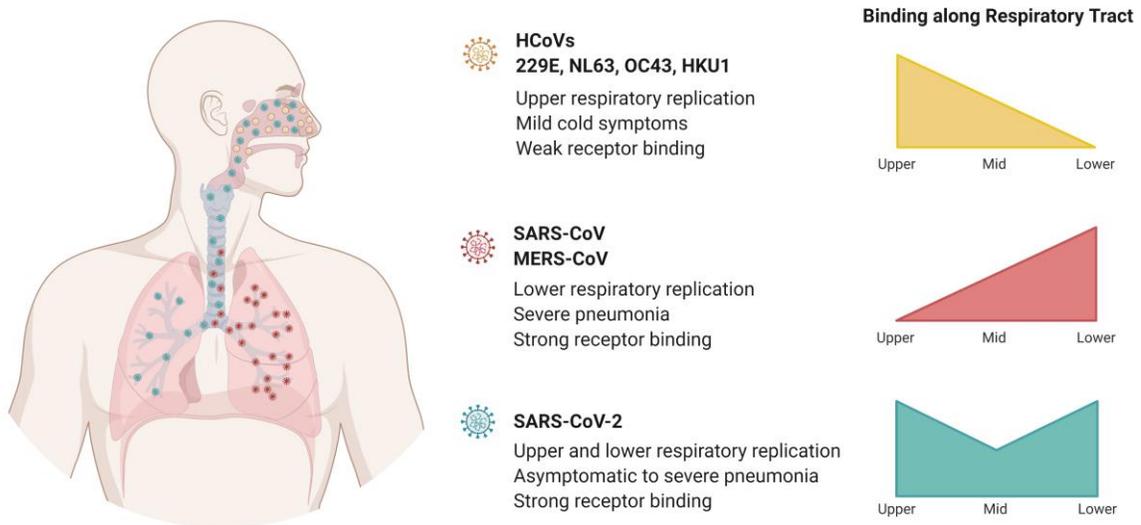
Coronaviruses enter cells through two mechanisms – the cell surface or endosomal route – where the spike will be cleaved by different host proteases to initiate membrane fusion. The spike is processed by serine proteases like TMPRSS2 in the cell surface route and cathepsins in the endosomal route. The virus will then undergo fusion and release its genetic material into the cytosol for replication and translation.



**Figure 1.3. Transmission versus pathogenesis.**

An inverse relationship exists between transmission (incidence) and pathogenesis (case fatality rate, CFR) as the more transmissible viruses are less pathogenic (e.g. hCoVs and seasonal influenza) and vice versa (e.g. MERS and H5N1 influenza). SARS-CoV-2 has established a balance displaying high transmissibility and moderate pathogenesis, similar to the 1918 H1N1 influenza. The CFR of SARS-CoV-2 vary among individual countries with the World Health Organization (WHO) representing the global values. The red cloud represents an estimated range of the final SARS-CoV-2 incidence and CFR. ● represents ongoing or seasonal values; ■ represents historical values.

Reproduced from Vu and Menachery (2021) *PLOS Pathogens* with permission<sup>65</sup>.



**Figure 1.4. Comparison of coronavirus tropism.**

Less pathogenic human coronaviruses (hCoVs) replicate in the upper respiratory tract while highly pathogenic SARS-CoV and MERS-CoV are found to replicate in the lower respiratory tract. Viral replication is not typically observed in the mid airways for both groups. Although concentrated in the upper and lower airways, viral replication of SARS-CoV-2 has been detected throughout the respiratory tract. The ability of SARS-CoV-2 to infect and replicate throughout the airways allows for greater range in transmissibility and pathogenicity.

Reproduced from Vu and Menachery (2021) *PLOS Pathogens* with permission <sup>65</sup>.

## Chapter 2: Beyond the Furin Cleavage Site

### INTRODUCTION

SARS-CoV-2 emerged in late 2019 and has caused the largest pandemic since the 1918 influenza outbreak <sup>66</sup>. An unusual feature of SARS-CoV-2 is the presence of a furin cleavage site in its spike protein <sup>67</sup>. The CoV spike is a trimer of spike proteins composed of the S1 and S2 subunits, responsible for receptor binding and membrane fusion, respectively <sup>66</sup>. After receptor binding, the spike protein is proteolytically cleaved at the S1/S2 and S2' sites to activate the fusion machinery. For SARS-CoV-2, the spike protein contains a novel cleavage motif recognized by the host cell furin protease (PRRAR) directly upstream of the S1/S2 cleavage site that facilitates cleavage prior to virion release from the producer cell. This furin cleavage site (FCS), not found in other group 2B CoVs, plays a key role in spike processing, infectivity, and pathogenesis as shown by our group and others <sup>53,68</sup>.

Importantly, another novel amino acid motif, QTQTN, is found directly upstream of the FCS. This QTQTN motif, also absent in other group 2B CoVs, is often deleted and has been pervasive in cultured virus stocks of the alpha, beta, and delta variants <sup>69-72</sup>. In addition, the QTQTN deletion has been observed in a small subset of patient samples as well <sup>73-75</sup>. Because this deletion has been frequently identified, we set out to characterize it and determine whether it has consequences for viral replication and virulence. Using our infectious clone <sup>76,77</sup>, we demonstrated that the loss of this motif attenuates SARS-CoV-2 replication in respiratory cells *in vitro* and pathogenesis in hamsters. The QTQTN deletion results in reduced spike cleavage and diminished capacity to use serine proteases on the cell surface for entry. Possible structural elements that contribute to attenuation by

$\Delta$ QTQTN include changes to loop length and glycosylation. Importantly, mutations of glycosylation-enabling residues in the QTQTN motif results in similar replication attenuation despite intact spike processing. Evaluations of loop length importance returned to  $\Delta$ PRRA; a new infectious clone PQQA was generated to determine if shortening the loop by deleting PRRA without FCS disruption resulted in the previously observed attenuation. PQQA produced attenuated replication and decreased spike processing, confirming integrity of the FCS is necessary for SARS-CoV-2. Together, our results highlight elements in the SARS-CoV-2 spike in addition to the furin cleavage site that contribute to increased replication and pathogenesis.

## RESULTS

### **$\Delta$ QTQTN attenuates viral replication *in vitro*.**

In addition to the furin cleavage site, comparison of group 2B coronavirus sequences also revealed the presence of an upstream QTQTN motif directly in the SARS-CoV-2 spike protein; this motif is absent in most other coronaviruses except for the closely related RaTG13 bat coronavirus (**Figure 2.1**). Importantly, this QTQTN motif is often deleted in SARS-CoV-2 strains propagated in Vero E6 cells<sup>69-72</sup>. To explore the role of the QTQTN motif in SARS-CoV-2 infection and pathogenesis, we generated a mutant in the WA-1 background (early U.S. case from 2020) by deleting QTQTN ( $\Delta$ QTQTN) using our reverse genetics system<sup>76,77</sup> (**Figure 2.2**). Examining the deletion on the SARS-CoV-2 spike structure, our modeling suggested that the  $\Delta$ QTQTN mutant forms a stable  $\alpha$ -helix in the loop containing the S1/S2 cleavage site (**Figure 2.3**). While the mutant retains the furin cleavage motif, its  $\alpha$ -helix is predicted to make the loop less flexible and reduce access to the proteolytic cleavage site.

The deletion of QTQTN motif did not affect virus replication in Vero E6 cells (African green monkey kidney cells) with the rescue stock titer comparable to wild-type WA-1 (WT) in yield; yet, the  $\Delta$ QTQTN mutant produced a large plaque morphology (**Figures 2.4A-B**), as seen with a FCS knockout mutant ( $\Delta$ PRRA)<sup>53</sup>. We then evaluated replication kinetics of  $\Delta$ QTQTN in Vero E6 cells and found no difference between WT and  $\Delta$ QTQTN (**Figure 2.5A**). However, following direct 1:1 competition infection, the  $\Delta$ QTQTN mutant had a significant advantage relative to WT SARS-CoV-2 in Vero E6 cells (**Figure 2.6A**). This fitness advantage for  $\Delta$ QTQTN likely explains the accumulation of this mutation in Vero E6-amplified virus stocks, as we also observed emergence of this mutation in Vero E6 cells infected with WT alone (**Figures 2.6B-C**). Notably, in Calu-3 2B4 cells, a human respiratory cell line, we observed a  $\sim$ 2.5 log reduction in  $\Delta$ QTQTN replication at both 24 and 48 hours post infection (hpi) (**Figure 2.5B**). Together, the results indicate that  $\Delta$ QTQTN mutant is attenuated in respiratory cells and has a fitness advantage in Vero E6 cells; these results are similar findings to those we reported for the SARS-CoV-2 FCS knockout virus<sup>53</sup>.

#### **$\Delta$ QTQTN attenuates disease, but not replication *in vivo*.**

We next evaluated the role of  $\Delta$ QTQTN on virulence in an *in vivo* model. Three-to-four-week-old male golden Syrian hamsters, which develop disease similar to that seen in humans<sup>78</sup>, were intranasally inoculated with  $10^5$  plaque-forming units (pfu) of WT SARS-CoV-2 or  $\Delta$ QTQTN mutant and monitored for 7 days post infection (dpi) (**Figure 2.7A**). Hamsters infected with WT steadily lost weight from 2 dpi with average peak weight loss of  $\sim$ 10% before beginning to recover at 5 dpi and regaining their starting weight by 7 dpi (**Figure 2.7B**). The disease score peaks corresponded with maximum weight loss,

with hamsters exhibiting ruffled fur, hunched posture, and/or reduced activity requiring additional welfare checks (**Figure 2.7C**). In contrast, hamsters infected with  $\Delta$ QTQTN experienced minimal weight loss, instead gaining weight over the course of the infection (**Figure 2.7B**). Similarly, no obvious disease was observed in  $\Delta$ QTQTN infected animals (**Figure 2.7C**). Hamsters infected with  $\Delta$ QTQTN developed pulmonary lesions that were less extensive than those in hamsters infected with WT SARS CoV-2, involving smaller portions of the infected lungs on both days 2 and 4 after intranasal inoculation (**Figure 2.8**). All of the lesions were similar, with interstitial pneumonia, peribronchitis, peribronchiolitis, and vasculitis with predominantly subendothelial and perivascular infiltration by lymphocytes, and perivascular edema. Characteristic cytopathologic effects were observed in alveolar pneumocytes and bronchiolar epithelium, including cellular enlargement, binucleation and multinucleation, and prominent nucleoli. Together, the results demonstrate that the deletion of QTQTN motif attenuates SARS-CoV-2 disease *in vivo*.

Despite clear attenuation in disease,  $\Delta$ QTQTN viral replication *in vivo* was not compromised compared to WT SARS-CoV-2. Indeed,  $\Delta$ QTQTN viral titers were greater than those of WT SARS-CoV-2 with a 10-fold increase in nasal wash titers at 1, 2, and 4 dpi (**Figure 2.9A**). Similar titer increases were observed in the trachea of infected hamsters at 2 dpi with equivalent titers at 4 dpi (**Figure 2.9B**). Notably, viral titers were equivalent in the lungs for both 2 and 4 dpi (**Figure 2.9C**). RNA expression data from hamster lung samples revealed clustering of WT and  $\Delta$ QTQTN at 2 and 4 dpi (**Figure 2.10**). Of note, although more variability was present at 2 dpi,  $\Delta$ QTQTN was slightly closer to mock samples at both time points. However, upregulated genes were similar between WT and

$\Delta$ QTQTN in comparison to mock at both time points (**Figures 2.11A-B**). Together, these results indicate that attenuation of  $\Delta$ QTQTN *in vivo* is not due to a change in replication capacity. In addition, these data are consistent with *in vivo* results with the FCS knockout virus<sup>53</sup>.

### **$\Delta$ QTQTN reduces spike processing and entry.**

To examine the role of the QTQTN motif in spike processing, Vero E6 and Calu3-2B4 cells were infected with WT or  $\Delta$ QTQTN and supernatants harvested at 24 hpi. Virions were then purified through sucrose cushion ultracentrifugation. Western blotting of the purified virions revealed reduced spike processing at the S1/S2 cleavage site for  $\Delta$ QTQTN compared to WT in Vero E6 cells (**Figures 2.12A-B**). Loss of the QTQTN motif resulted in little S1/S2 cleavage product and a significant increase in full-length spike compared to WT control. A similar reduction in spike processing was seen in Calu3-2B4 cells, although with more processing overall compared to in Vero E6 (**Figures 2.12C-D**). Thus, deletion of the QTQTN motif impairs spike cleavage at the S1/S2 site, similar to findings with the SARS-CoV-2 mutants lacking the FCS<sup>53,68</sup>.

After receptor binding, the CoV spike protein is cleaved by a host protease as part of the virus entry process. Different proteases can be utilized to activate the spike fusion machinery: serine proteases like TMPRSS2 at the cell surface or cathepsins within endosomes (**Figure 2.13**). Our prior work found that the absence of TMPRSS2 in Vero E6 cells plays a role in selection of SARS-CoV-2 strains with FCS deletions<sup>53</sup>. Calu-3 2B4 cells also express high levels of TMPRSS2. We therefore hypothesized that the absence of TMPRSS2 activity contributes to  $\Delta$ QTQTN selection in Vero E6 and attenuation in Calu3 2B-4 cells. To test this hypothesis, Vero E6 cells expressing TMPRSS2 were pretreated

with cathepsin inhibitor K11777 before infection with WT or  $\Delta$ QTQTN, and viral titers were measured at 24 hpi. With cathepsin inhibited and TMPRSS2 activity intact, a significant,  $\sim$ 1.5 log reduction in viral titer was observed for  $\Delta$ QTQTN compared to WT over a dose range of K11777, mirroring the attenuation observed in the Calu-3 2B4 cells (**Figure 2.14A**). Infection of untreated TMPRSS2-expressing Vero E6 revealed no difference in replication between WT and  $\Delta$ QTQTN (**Figure 2.15**). When Calu-3 2B4 cells were pretreated with the serine protease inhibitor camostat mesylate, WT SARS-CoV-2 titers were reduced and equivalent to  $\Delta$ QTQTN (**Figure 2.14B**). Together, these data indicate that the loss of the QTQTN motif reduces the capacity of the virus to use TMPRSS2 for entry via the cell surface.

#### **Glycosylation of the QTQTN motif contributes to spike processing.**

As the absence of the QTQTN motif attenuates SARS-CoV-2, we set out to determine if the QTQTN motif itself has a significant role during infection. Notably, the second threonine, T678, of the motif has been previously shown to be O-linked glycosylated<sup>79,80</sup>. Structurally, the QTQTN site resides on an exterior loop of the spike and is capable of accommodating large glycans, which may contribute to interactions with proteases like TMPRSS2 (**Figure 2.16**). To determine the role of glycosylation in spike processing, we generated mutants abolishing the glycosylated T678 (QTQVN) alone or together with the first threonine T676 (QVQVN) to exclude possible compensatory glycosylation (**Figure 2.17**). To confirm that the mutations disrupted glycosylation of the QTQTN motif, virions from Vero E6-produced supernatant were purified through a sucrose cushion and prepared for LC-MS/MS analysis. Peptide fragments were analyzed for post translational modifications, including N-linked and O-linked glycosylation.

Similar to other studies, T678 of WT spike was detected to have 10% occupancy with O-glycan HexNAc(1)Hex(1) (**Figures 2.18 and 2.19A-B**). In contrast, no O-linked glycosylation or other modification was detected across the S1/S2 cleavage site on this peptide of either the QTQVN or QVQVN mutants (**Figures 2.18 and 2.19C-D**). However, further analysis of the data detected that an adjacent, overlapping peptide had an O-linked glycosylation in the QTQVN mutant (**Table 2.1**); this modification is absent in both the WT and QVQVN mutant, suggesting a potential compensatory glycosylation in QTQVN.

Similar to  $\Delta$ QTQTN, the glycosylation mutations did not affect virus yield with titers comparable to WT (**Figure 2.20A**). However, plaque morphologies of QTQVN and QVQVN were more similar to WT than to the  $\Delta$ QTQTN mutant (**Figure 2.20B**). Viral replication in Vero E6 cells was not affected for QTQVN and QVQVN mutants; however, both were attenuated at 24 hpi in Calu-3 2B4 cells, mirroring what was observed with the  $\Delta$ QTQTN mutant, before reaching an endpoint titer equivalent to WT by 48 hpi (**Figures 2.21A-B**). Together, these results indicate that the loss of glycosylation sites in the QTQTN motif attenuates replication in Calu-3 2B4 cells.

We next examined if the glycosylation-defective mutants had altered spike processing similar to  $\Delta$ QTQTN. Western blotting of virions purified from Calu-3 2B4 cells by sucrose cushion ultracentrifugation revealed intact spike processing at the S1/S2 site with both glycosylation mutants (**Figures 2.22A-B**). Contrasting with  $\Delta$ QTQTN, the QTQVN and QVQVN mutants had significant S1/S2 spike cleavage with levels similar to WT SARS-CoV-2. Likewise, spike processing of QTQVN and QVQVN in normal and TMPRSS2-expressing Vero E6 cells was similar to that of WT (**Figures 2.22C-F**).

Together, these results suggest that the loss of glycosylated residues does not impact spike processing of SARS-CoV-2.

To further understand the role of glycosylation on the QTQTN motif, we next examined if glycosylation is involved in protease interaction and usage. Using TMPRSS2-expressing Vero E6, we pretreated cells with K11777 to disrupt cathepsin activity, infected with WT,  $\Delta$ QTQTN, QTQVN, or QVQVN, and examined viral titers at 24 hpi (**Figure 2.23A**). While the  $\Delta$ QTQTN titer was reduced compared to WT as observed before, the QTQVN titer was equivalent to WT (**Figure 2.23A**). In contrast, disruption of both glycosylation residues with the QVQVN mutant resulted in attenuation equivalent to that of  $\Delta$ QTQTN, suggesting that abolishing both O-linked glycosylation sites disrupted TMPRSS2 utilization (**Figure 2.23A**). We subsequently pretreated Calu-3 2B4 cells with camostat mesylate to disrupt serine protease activity and infected with the glycosylation mutants (**Figure 2.23B**). Interestingly, while treatment with camostat in Calu-3 2B4 cells reduced WT titer to equivalent levels as  $\Delta$ QTQTN, viral titers of QTQVN and QVQVN mutants were even lower. However, the overall differences in titer between WT and the glycosylation mutants was reduced, suggesting that glycosylation is important for TMPRSS2 utilization and entry (**Figure 2.23B**). Overall, these results argue that glycosylation of the QTQTN motif is important to protease interactions with spike and SARS-CoV-2 infection.

### **Integrity of the FCS sequence is required by SARS-CoV-2.**

As shown by structural modeling, one of the possible attenuation mechanisms for  $\Delta$ QTQTN is a shortened loop at the site of the S1/S2 cleavage site due to the deletion. Considering this, we questioned whether the phenotype observed previously by  $\Delta$ PRRA

was due to shortening the loop instead of removing the furin recognition site itself. Using our reverse genetics system, we generated an infectious clone in the WA1 background that substituted the arginines in the FCS with glutamines (PQQA), disrupting the FCS sequence while maintaining similar loop length. Vero E6 and Calu-3 2B4 cells were infected with WT or PQQA, and viral titer was monitored over 48 hpi (**Figure 2.24**). While PQQA displayed comparable replication kinetics as WT in Vero E6 cells (**Figure 2.24A**), PQQA was attenuated at all time points in Calu-3 2B4 (**Figure 2.24B**), similar to  $\Delta$ PRRA. This suggests that the FCS sequence itself is the contributing factor leading to attenuated viral replication, not the shortened loop caused by the deletion.

We next evaluated whether mutating the FCS but maintaining loop length alters spike processing. Calu-3 2B4 cells were infected with WT or PQQA, and virions were sucrose purified from 24 hpi supernatant. Western blotting of the spike protein revealed that while ~80% of WT spike was processed into the S1/S2 cleavage product, only ~3% of PQQA was processed (**Figures 2.25A-B**), similar to  $\Delta$ PRRA. From these results, an intact FCS sequence is indeed required for efficient spike processing.

## **DISCUSSION**

The presence of the FCS in SARS-CoV-2 plays a critical role in infection and pathogenesis by facilitating an increase spike processing upon nascent virion release <sup>53</sup>. This FCS, unusual to SARS-like coronaviruses, has been highlighted as a potential “smoking gun” for an engineered virus <sup>81</sup>. Yet, the FCS alone is insufficient to drive infection and pathogenesis. The upstream QTQTN motif adds two distinct elements that contribute to this capacity and virulence. The loss of the QTQTN motif produces a shorter, more rigid exterior loop in the spike, likely reducing access to the FCS. The result is a

significant reduction in spike processing and attenuation of the  $\Delta$ QTQTN mutant both *in vitro* and *in vivo*. Additionally, while mutations of the glycosylation residues in the QTQTN motif do not change overall spike processing, the modification of the motif attenuates virus replication in a TMPRSS2-dependent manner. Similar TMPRSS2-dependent attenuation has been observed in pseudoviruses displaying  $\Delta$ QTQTN SARS-CoV-2 spike protein<sup>82</sup>. Further studies with PQQA will also assess protease interactions; should PQQA have unaltered protease interaction, unlike the QTQTN glycosylation mutants, it would demonstrate that efficient spike processing is not dependent on protease interaction.

Concerning loop length, the distinct difference in spike processing between  $\Delta$ QTQTN and the glycosylation mutants suggests that the shortened loop affected spike processing more so than the lack of glycosylation. However, reexamination of the FCS with PQQA confirms that the FCS sequence is the primary driver for efficient SARS-CoV-2 spike processing. That is not to say loop length is discounted in its contributions to coronavirus pathogenesis. In a fatal case associated with hCoV-OC43, which typically results in mild cold symptoms, the virus was found with an additional four amino acids adjacent to the S1/S2 cleavage site, extending the loop<sup>83</sup>. Altogether, alterations to the loop length and thus access to the S1/S2 cleavage site should be considered an indicator of coronavirus pathogenesis.

Overall, our results argue that the FCS, the length/composition of the exterior loop, and glycosylation of the QTQTN motif are all needed for efficient infection and pathogenesis (**Figure 2.26**). Disruption of any of these three elements attenuates SARS-

CoV-2, highlighting the complexity of spike activation beyond the simple presence of a furin cleavage site.

## **METHODS**

### **Cell culture**

Vero E6 cells were grown in DMEM (Gibco #11965-092) supplemented with 10% fetal bovine serum (FBS) (HyClone #SH30071.03) and 1% antibiotic-antimycotic (Gibco #5240062). Calu-3 2B4 cells were grown in DMEM supplemented with 10% defined FBS (HyClone #SH30070.03), 1% antibiotic-antimycotic, and 1 mg/ml sodium pyruvate. Vero E6 expressing TMPRSS2 cells were grown in DMEM (Gibco #11885-084) supplemented with 10% FBS and 1 mg/ml geneticin (Gibco #10131035).

### **Viruses**

The recombinant wild-type (WT) and mutant SARS-CoV-2 virus sequences are based on the USA-WA1/2020 isolate sequence provided by the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA), which was originally obtained from the USA Centers for Disease Control and Prevention (CDC) as previously described<sup>84</sup>. The mutant viruses ( $\Delta$ QTQTN, QTQVN, and QVQVN) were generated using standard cloning techniques and reverse genetics system as previously described<sup>53,76</sup>. Mutations were verified through Sanger sequencing of cDNA produced from viral RNA extract. Standard plaque assays were used for virus titer.

### ***In vitro* infection**

Viral infections in Vero E6, Calu-3 2B4 and TMPRSS2-expressing Vero E6 cells were performed as previously described<sup>77,85</sup>. Briefly, cells were washed with PBS and infected with WT or mutant SARS-CoV-2 at an MOI of 0.01 for 45 min at 37°C. Following absorption, cells were washed three times with PBS and fresh growth media was added. Three or more biological replicates were collected at each time point.

### **Protease inhibitor treatment**

TMPRSS2-expressing Vero E6 or Calu-3 2B4 cells were pretreated with 50-100 µM of K11777 (AdipoGen #AG-CR1-0158-M005) or 50-100 µM of camostat mesylate (Sigma-Aldrich #SML0057-10MG), respectively, in 1 ml growth medium for 1 hr at 37°C. Cells were subsequently washed with PBS and infected with WT or mutant SARS-CoV-2 at an MOI of 0.01 as described in '*In vitro* infection'.

### **Competition assay and next generation sequencing analysis.**

Ratios (1:0, 1:1, and 0:1 WT:ΔQTQTN) for the competition assays were determined by pfu of virus stock. Vero E6 cells were infected with a total MOI of 0.01 (WT alone, 1:1 WT:ΔQTQTN, or ΔQTQTN alone) as described in '*In vitro* infection'. RNA was collected from cell lysate with Trizol reagent (Invitrogen #15596018) and extracted with Direct-zol RNA Miniprep Plus kit (Zymo #R2072). RNA libraries were prepared by ClickSeq and sequenced as previously described<sup>53,86</sup>.

### **Virion purification and western blotting**

Vero E6, Calu-3 2B4 and TMPRSS2-expressing Vero E6 cells were infected with WT or mutant SARS-CoV-2 at an MOI of 0.01. Culture supernatant was harvested 24 hpi and clarified by low-speed centrifugation. Virus particles were then pelleted by

ultracentrifugation through a 20% sucrose cushion at 26,000 rpm for 3 hr using a Beckman SW28 rotor. Pellets were resuspended in 2x Laemmli buffer to obtain protein lysates. Relative viral protein levels were determined by SDS-PAGE followed by western blot analysis as previously described<sup>7,53,84,87</sup>. In brief, sucrose-purified WT and mutant SARS-CoV-2 virions were inactivated by boiling in Laemmli buffer. Samples were loaded in equal volumes into 4–20% Mini-PROTEAN TGX Gels (Bio-Rad #4561093) and electrophoresed by SDS-PAGE. Protein was transferred to polyvinylidene difluoride (PVDF) membranes. Membranes were probed with SARS-CoV S-specific antibodies (Novus Biologicals #NB100-56578) and followed with horseradish peroxidase (HRP)-conjugated anti-rabbit antibody (Cell Signaling Technology #7074S). Membranes were stripped and reprobed with SARS-CoV N-specific antibodies (provided by S. Makino) and the HRP-conjugated anti-rabbit secondary IgG to measure loading. Signal was developed using Clarity Western ECL substrate (Bio-Rad #1705060) or Clarity Max Western ECL substrate (Bio-Rad #1705062) and imaging on a ChemiDoc MP System (Bio-Rad #12003154). Densitometry was performed using ImageLab 6.0.1 (Bio-Rad #2012931).

### **Hamster infection studies**

Male golden Syrian hamsters (3-4 weeks old) were purchased from Envigo. All studies were conducted under a protocol approved by the UTMB Institutional Animal Care and Use Committee and complied with USDA guidelines in a laboratory accredited by the Association for Assessment and Accreditation of Laboratory Animal Care. Procedures involving infectious SARS-CoV-2 were performed in the Galveston National Laboratory ABSL3 facility. Animals were housed in groups of five and intranasally inoculated with 10<sup>5</sup> pfu of WT or  $\Delta$ QTQTN SARS-CoV-2. Animals were monitored daily for weight loss

and development of clinical disease through the course of the study. Hamsters were anesthetized with isoflurane (Henry Schein Animal Health) for viral infection and nasal washes.

### **Histology**

Left lungs were harvested from hamsters and fixed in 10% buffered formalin solution for at least 7 days. Fixed tissue was then embedded in paraffin, cut into 5  $\mu$ M sections, and stained with hematoxylin and eosin (H&E) on a SAKURA VIP6 processor by the University of Texas Medical Branch Surgical Pathology Laboratory.

### **Structural modeling**

Structural models were generated using SWISS-Model to generate homology models for WT,  $\Delta$ QTQTN, and glycosylated QTQTN SARS-CoV-2 spike protein on the basis of the SARS-CoV-1 trimer structure (Protein Data Bank code 6ACD). Homology models were visualized and manipulated in PyMOL (version 2.4).

### **Transcriptomics**

Hamster lungs were homogenized in Trizol reagent (Thermo Fisher), and RNA was extracted with Direct-zol RNA Miniprep Plus kit (Zymo #R2072). The short-read sequencing libraries were generated from extracted RNA using Poly-A Click-Seq (PAC-Seq)<sup>88,89</sup>. Briefly, RNAs containing poly(A) tails were selectively reverse transcribed with oligo(dT) primers and stochastically terminated with azido-NTPS. Libraries were then gel purified (200-400 bp) and sequenced using Illumina platform (NextSeq550). Differential Poly-A Cluster (DPAC) was used to identify changes in overall expression<sup>88,89</sup>. An p-adjusted value (p-adj) of <0.1 and an absolute value of log2 fold change ( $|\log_2FC|$ ) greater

than 0.58 (or minimum of 50% increase/decrease) was used to filter results. The command ran for this data set was `~/DPAC -p PMCDB -t 4 -x [flattened_annotations] -y [reference_names] -g [genome] -n 6 -v golden_hamster,Mesaur [metadata_file] [index] [experiment name] [output_directory]`, where `-p` indicates parameters used, in this case P (perform data pre-processing), M (map data), C (force new PAS cluster generation), D (perform differential APA analysis), and B (make individual bedgraphs), `-t` indicates how many threads were to be used, and `-n` indicates number of replicates. Annotations, gene names, index were used for Syrian golden hamster and mapped to the Syrian golden hamster genome (Mesaur). Data quality was determined to be sufficient by generating and loading bedgraph files into the UCSC Genome Browser <sup>90</sup>.

### **Nanoflow-LC-MS/MS glycopeptide mapping**

Sucrose cushion-purified SARS-CoV-2 virion samples were prepared for LC-MS/MS analysis as previously described <sup>91</sup>. Samples were analyzed by nanoLC-MS/MS (nanoRSLC, ThermoFisher) using an Aurora series (Ion Opticks) reversed phase HPLC column (25 cm length x 75  $\mu$ m inner diameter) directly injected to an Orbitrap Eclipse using a 120 min gradient (mobile phase A = 0.1% formic acid (Thermo Fisher), mobile phase B = 99.9% acetonitrile with 0.1% formic acid (Thermo Fisher); hold 12% B for 5 min, 2-6% B in 0.1 min, 6-25% in 100 min, 25-50% in 15 min) at a flow rate of 350 nL/min. Eluted peptide ions were analyzed using a data-dependent acquisition (DDA) method with resolution settings of 120,000 and 15,000 (at m/z 200) for MS1 and MS2 scans, respectively. DDA-selected peptides were fragmented using stepped high energy collisional dissociation (27, 32, 37%). Tandem mass spectra were analyzed according to a label-free proteomic strategy using Proteome Discoverer (version 2.5.0.400,

ThermoFisher) with the Byonic (version 4.1.10, Protein Metrics) and Minora nodes using specific single protein databases for each sample, respective to the individual Spike WT and mutant amino acid sequences<sup>92,93</sup>. Mass tolerances of 10 ppm and 20 ppm were used for matching parent and fragment masses, respectively. Mass spectra were searched with a fixed modification of carbamidomethyl (C), and up to 2 common variable modifications of deamidation (N,Q), oxidation (M), phosphorylation (S, T, Y), and glycan modifications using an N-glycan database of 182 human N-glycans, 1 maximum allowed; 70 human O-glycans, 1 maximum allowed. Peptide spectral matches were filtered for quality: PEP2D < 0.01, Byonic score > 100<sup>94</sup>.

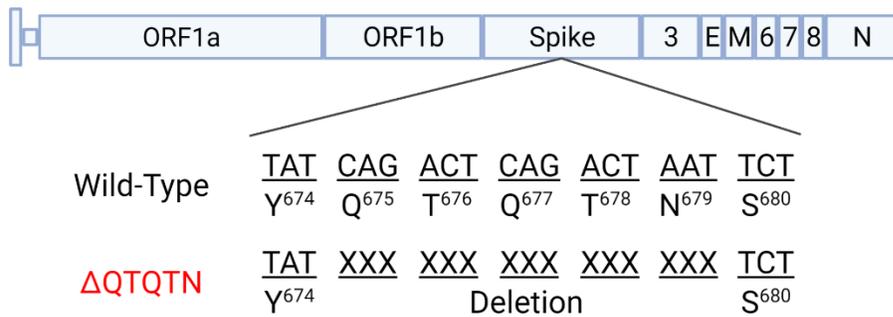
	S	Y	Q	T	Q	T	N	S	P	R	R	A	R	S	V
<b>SARS-2</b>	S	Y	Q	T	Q	T	N	S	P	R	R	A	R	S	V
<b>RaTG13</b>	S	Y	Q	T	Q	T	N	S	-	-	-	-	R	S	V
<b>HKU3</b>	S	Y	H	T	A	S	V	L	-	-	-	-	R	S	T
<b>WIV1</b>	S	Y	H	T	V	S	S	L	-	-	-	-	R	S	T
<b>SHC014</b>	S	Y	H	T	V	S	S	L	-	-	-	-	R	S	T
<b>WIV16</b>	S	Y	H	T	V	S	S	L	-	-	-	-	R	S	T
<b>SARS</b>	S	Y	H	T	V	S	S	L	-	-	-	-	R	S	T

S1 Cleavage Site

**Figure 2.1. Coronavirus S1/S2 cleavage site.**

Comparison of S1/S2 cleavage site across SARS-CoV, SARS-CoV-2 and related bat CoVs.

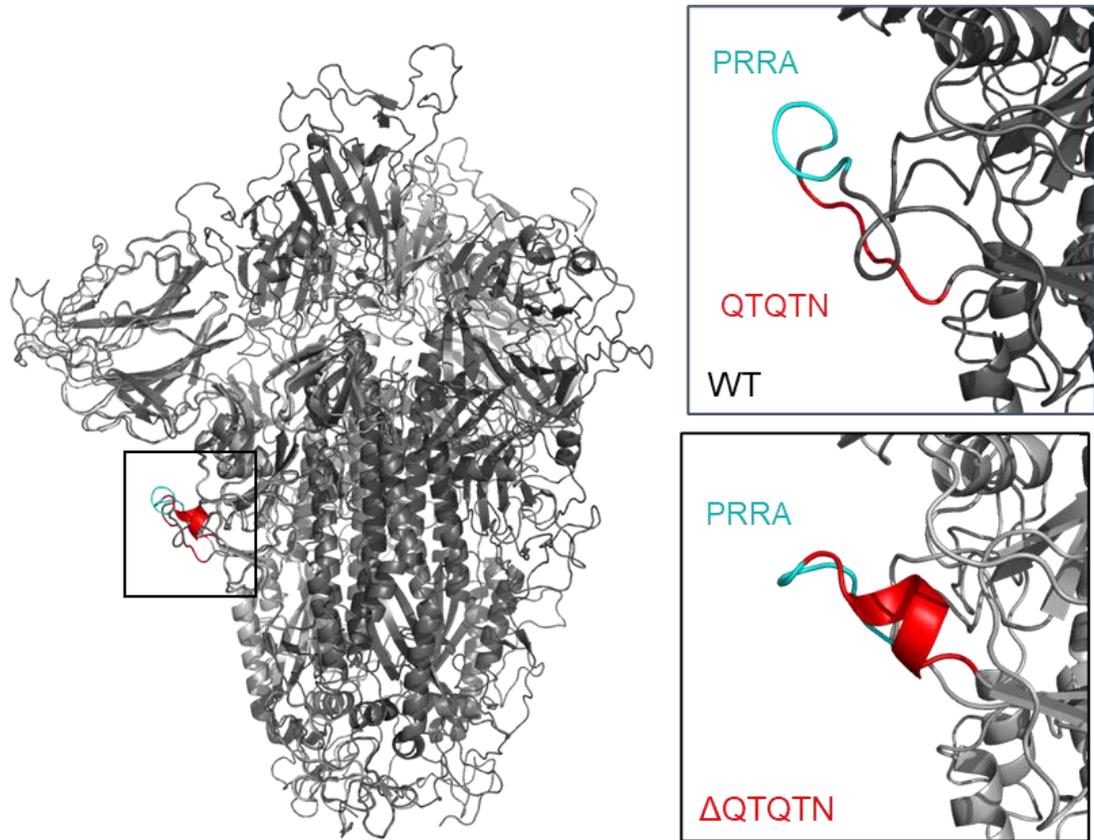
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.2. SARS-CoV-2 ΔQTQTN cloning schematic.**

Schematic of SARS-CoV-2 genome with deletion of QTQTN.

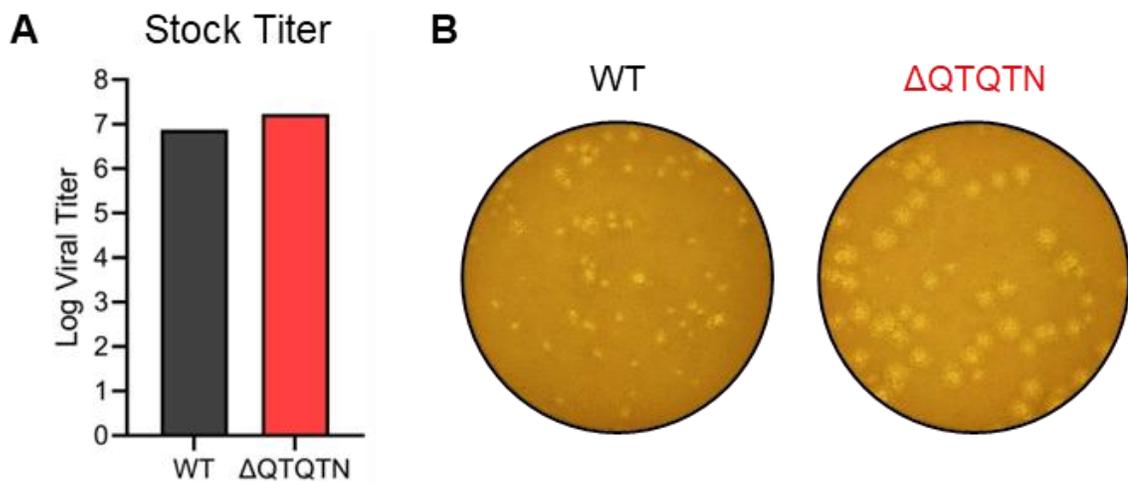
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.3. Structural modeling of SARS-CoV-2  $\Delta$ QTQTN.**

SARS-CoV-2 spike trimer (grey) with WT (upper) and  $\Delta$ QTQTN (lower) overlaid. PRRA (cyan) is exposed with QTQTN (red) extending the loop (upper). An  $\alpha$ -helix is formed with deletion of QTQTN (red) and PRRA (cyan) is exposed (lower).

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

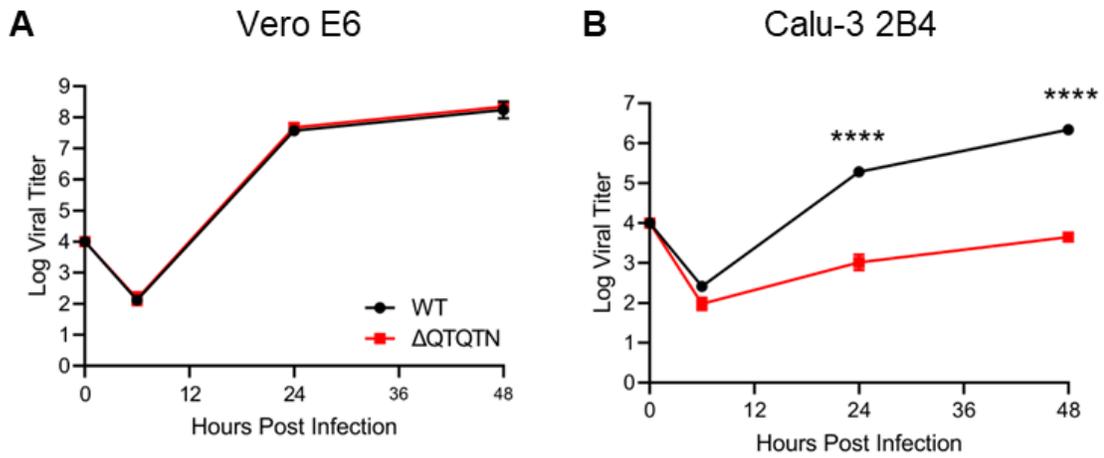


**Figure 2.4. Plaque morphology of  $\Delta$ QTQTN SARS-CoV-2.**

(A) Virus stock titer of WT and  $\Delta$ QTQTN SARS-CoV-2 from Vero E6.

(B) Plaque morphology of WT and  $\Delta$ QTQTN in Vero E6

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



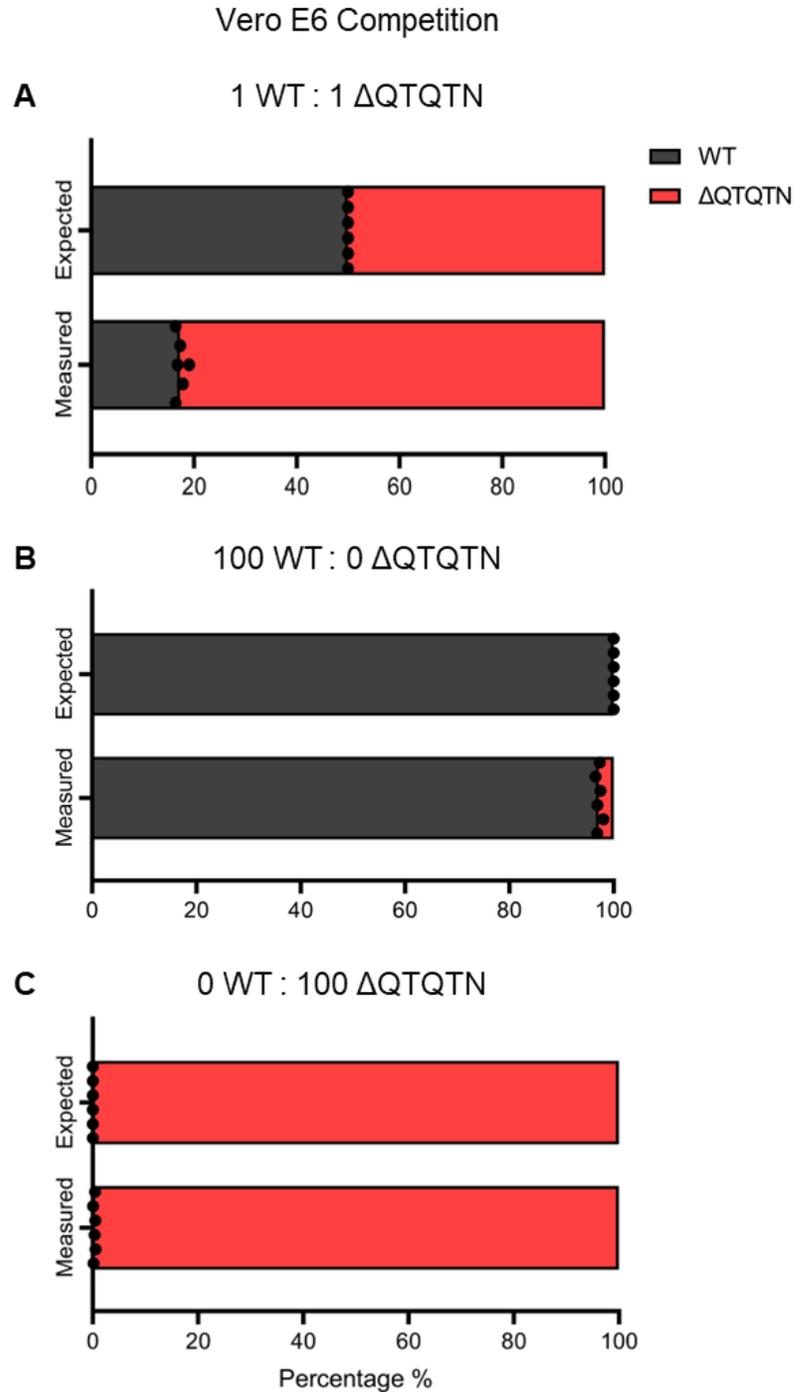
**Figure 2.5. *In vitro* replication of  $\Delta$ QTQTN.**

**(A)** Viral titer from Vero E6 infected with WT (black) or  $\Delta$ QTQTN (red) SARS-CoV-2 at an MOI of 0.01 (n=3).

**(B)** Viral titer from Calu-3 2B4 infected with WT or  $\Delta$ QTQTN SARS-CoV-2 at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d.

Statistical analysis measured by two-tailed Student's t-test. \*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ; \*\*\*\*,  $p \leq 0.0001$ .

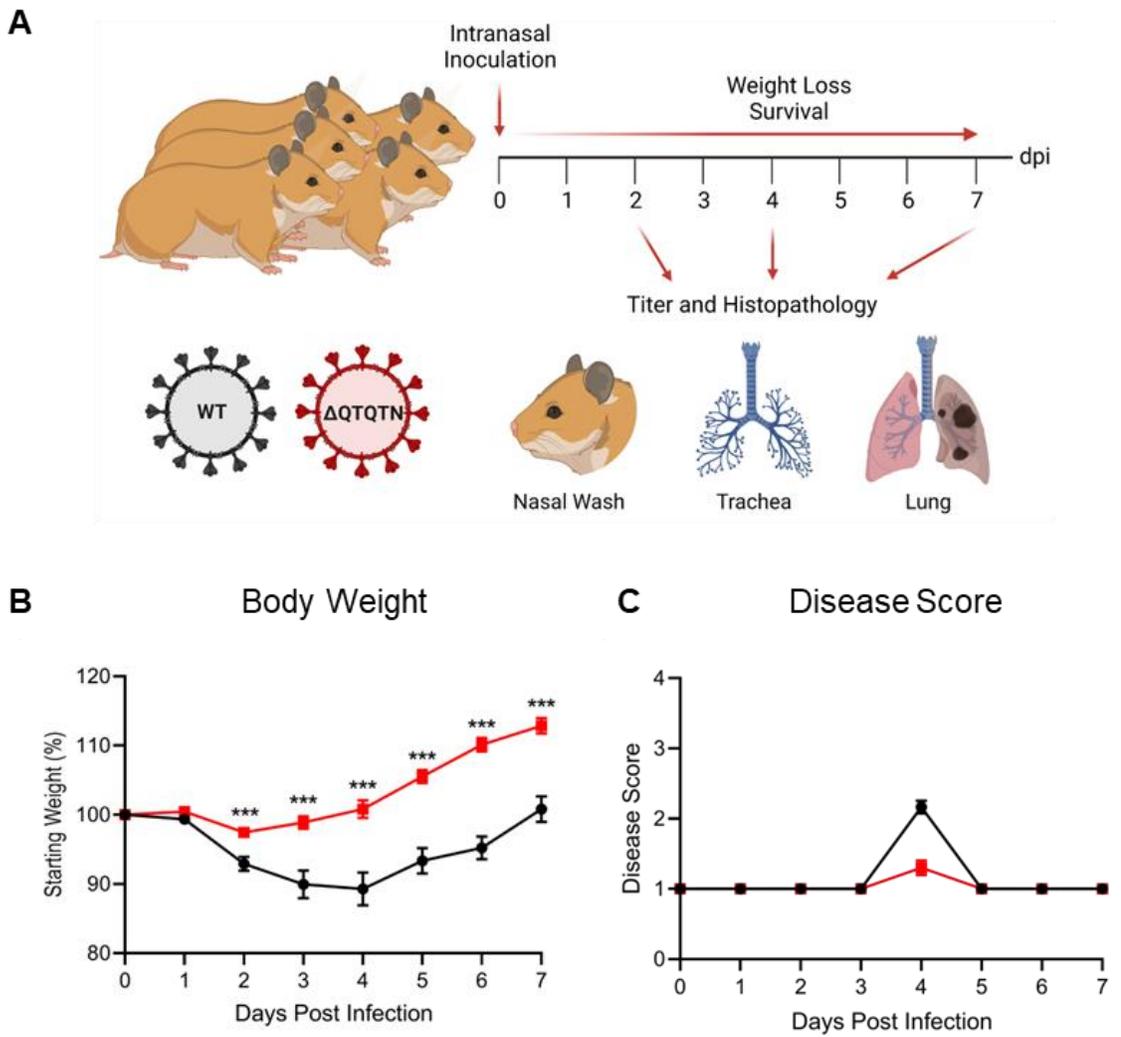
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.6. Vero E6 competition assay WT: $\Delta$ QTQTN**

Competition assay between WT and  $\Delta$ QTQTN SARS-CoV-2 at a ratio of 1 WT:1  $\Delta$ QTQTN (A), 100:0 (B), and 0:100 (C) showing RNA percentage from next generation sequencing.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

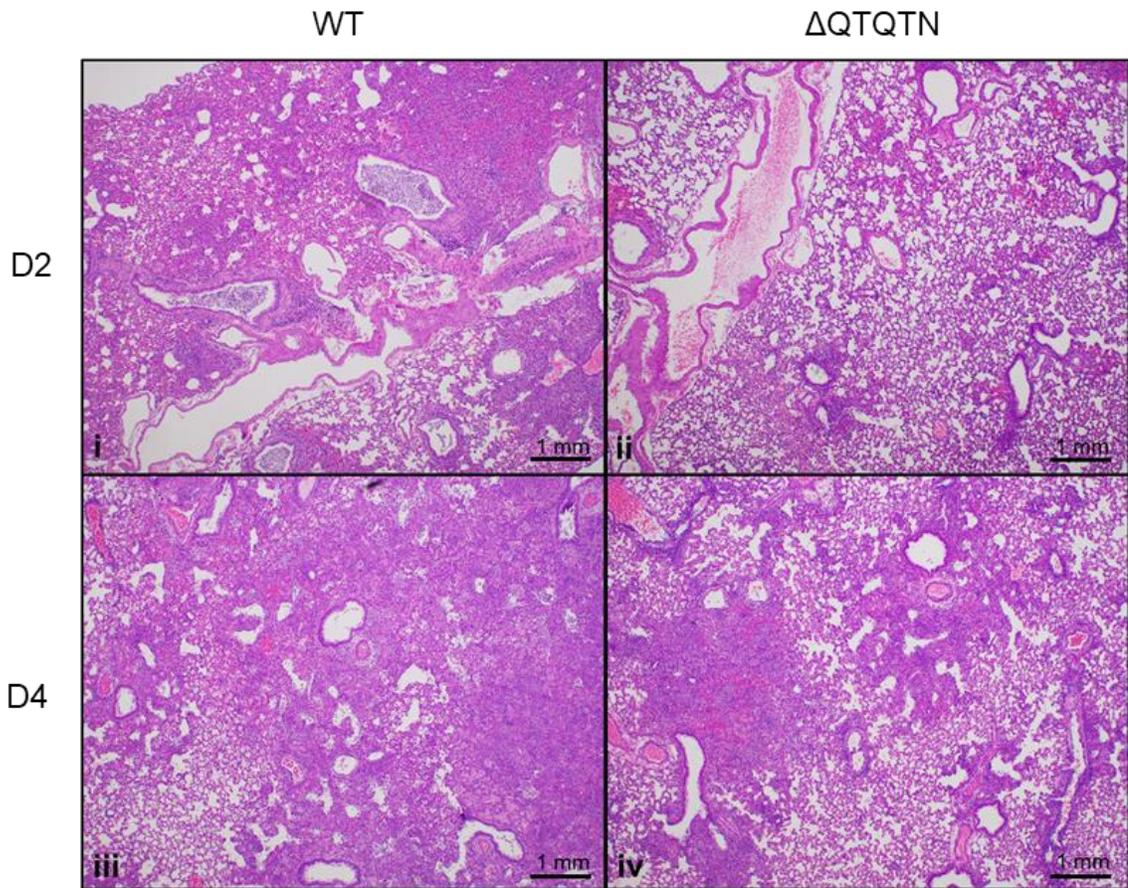


**Figure 2.7. *In vivo* characterization of SARS-CoV-2  $\Delta$ QTQTN in golden Syrian hamsters.**

(A) Schematic of golden Syrian hamster infection with WT (black) or  $\Delta$ QTQTN (red) SARS-CoV-2.

(B-C) Three- to four-week-old male hamsters were infected with 10<sup>5</sup> pfu of WT or  $\Delta$ QTQTN SARS-CoV-2 and monitored weight loss (B) and disease score (C) for seven days.

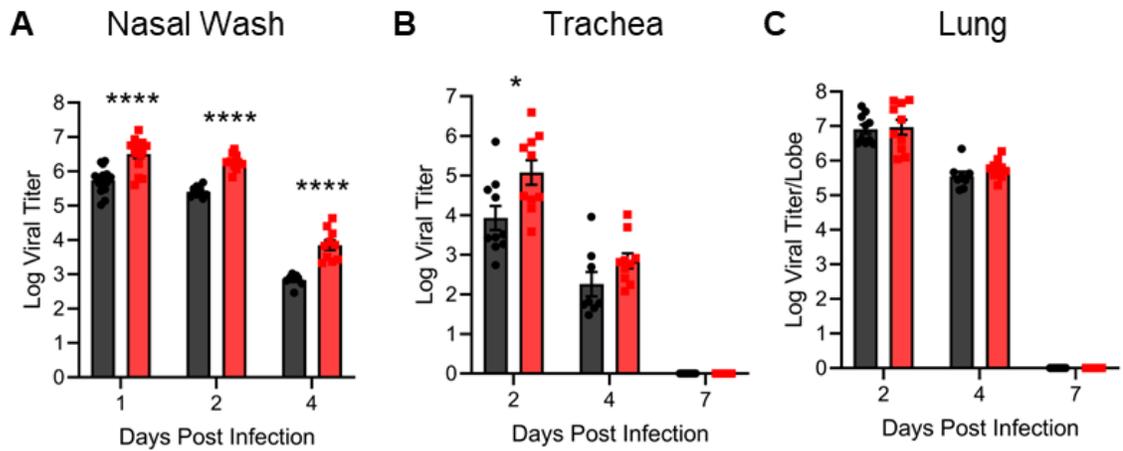
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.8. Histopathology of  $\Delta$ QTQTN infected hamsters.**

Histopathology of hamster lungs manifested more extensive lesions in animals infected with WT SARS-CoV-2 on day 2 (**i**) (4X) than in animals infected with  $\Delta$ QTQTN (**ii**) (4X). Lesions increased in volume on day 4 with greater proportions of the lungs affected in hamsters infected with WT (**iii**) (4X) than  $\Delta$ QTQTN (**iv**) (4X) on day 4.

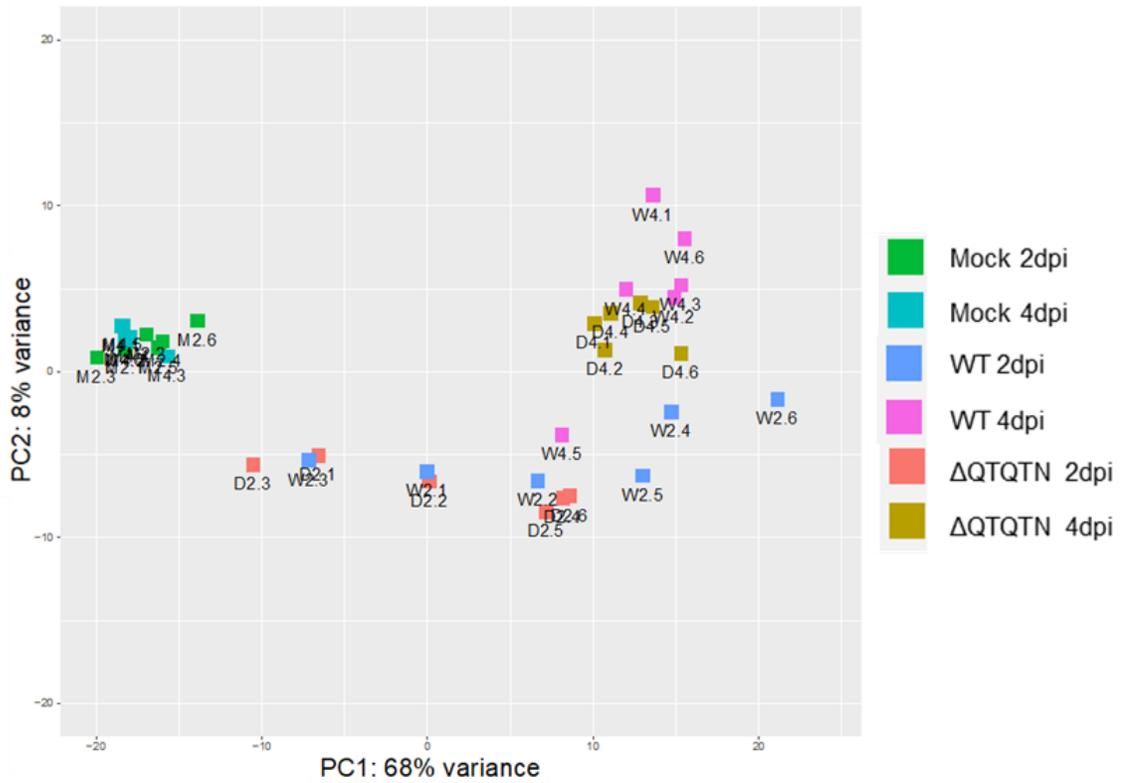
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.9. Viral replication of  $\Delta$ QTQTN infected hamsters.**

Viral titer was measured for nasal wash (A), trachea (B), and lung (C). Data are mean  $\pm$  s.e.m. Statistical analysis measured by two-tailed Student's t-test. \*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ; \*\*\*\*,  $p \leq 0.0001$ .

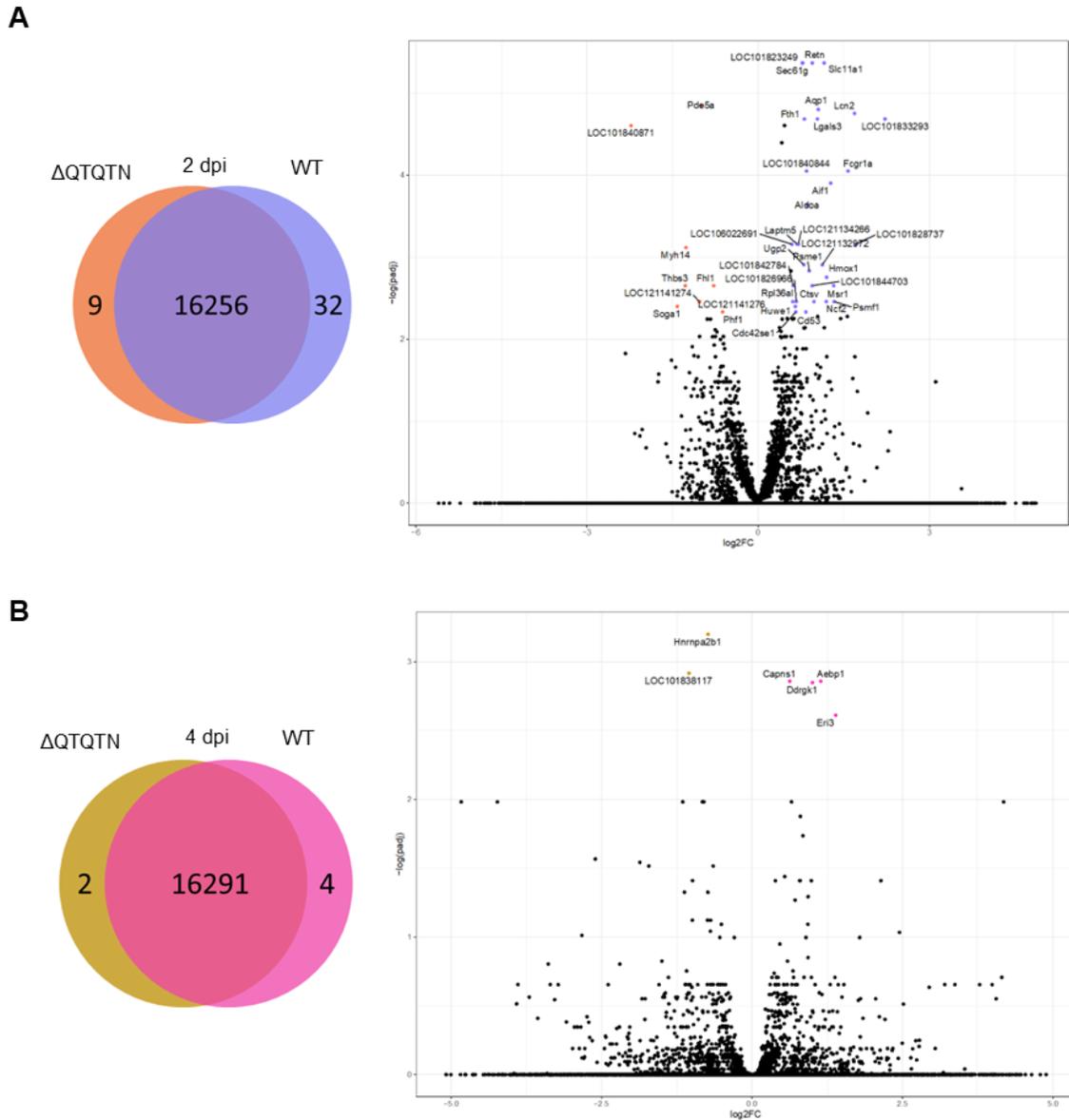
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.10. Clustering analysis of  $\Delta$ QTQTN SARS-CoV-2 infected hamster transcriptomics.**

Principal component analysis (PCA) plot of hamster lung samples.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

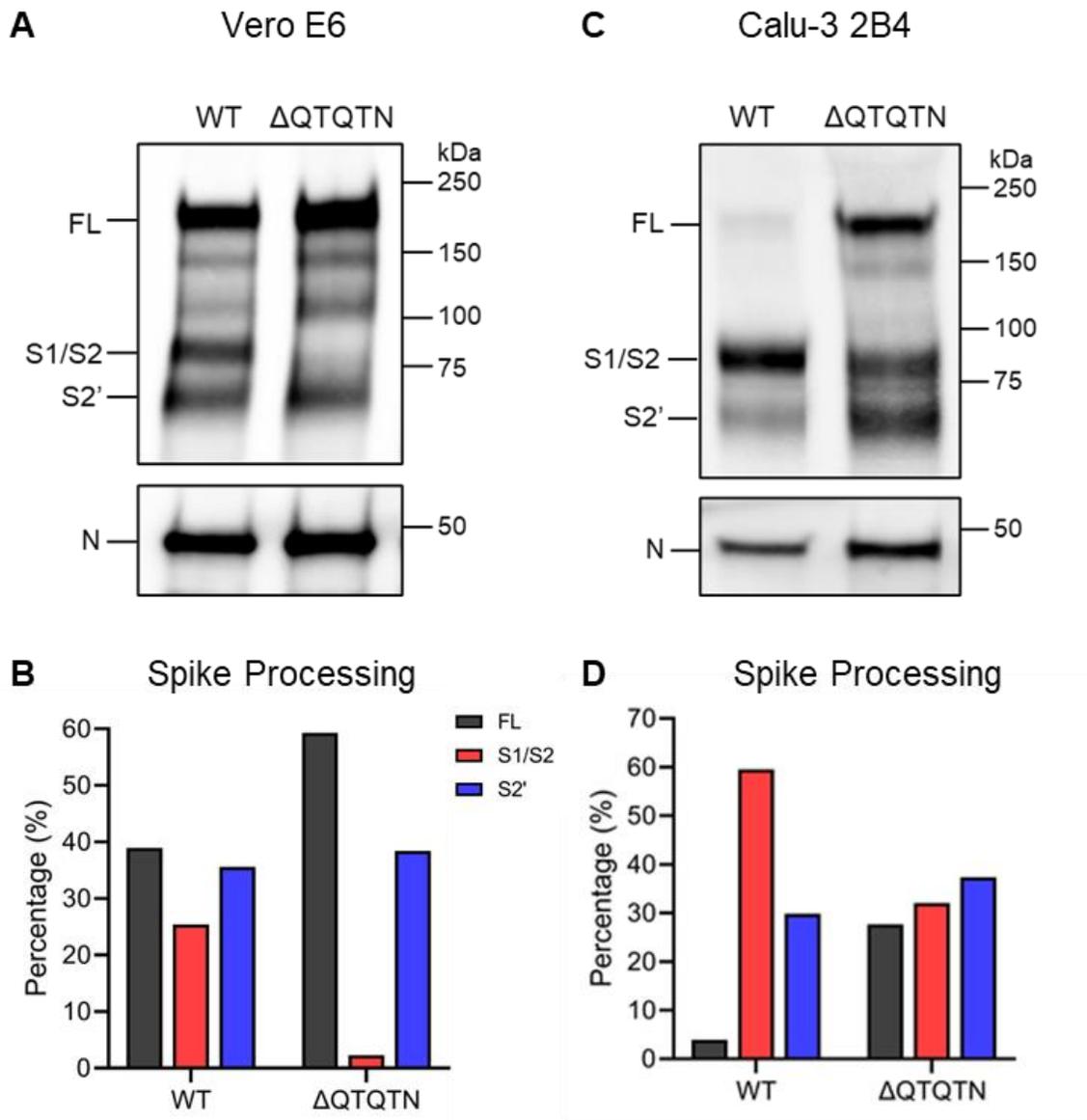


**Figure 2.11. Comparison of gene regulation in  $\Delta$ QTQTN SARS-CoV-2 infected hamster transcriptomics.**

(A) DESeq2 analysis of mapped genes between WT (purple) and  $\Delta$ QTQTN (orange) at 2 dpi (left) with upregulated genes indicated in volcano plot (right).

(B) DESeq2 analysis of mapped genes between WT (purple) and  $\Delta$ QTQTN (orange) at 4 dpi (left) with upregulated genes indicated in volcano plot (right).

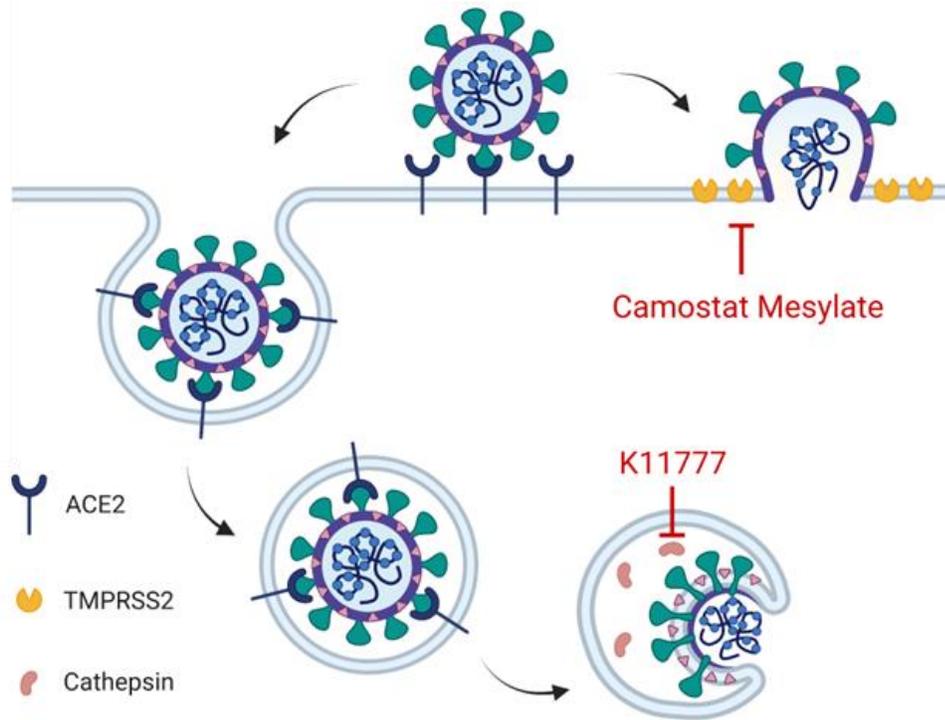
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.12. Spike processing of  $\Delta$ QTQTN.**

Purified WT and  $\Delta$ QTQTN SARS-CoV-2 virions from Vero E6 (**A-B**) and Calu-3 2B4 (**C-D**) probed with anti-S or anti-N antibody. Full-length (FL), S1/S2 cleavage product, and S2' cleavage product are indicated. Quantification of densitometry of FL (black), S1/S2 (red), and S2' (blue) normalized to N shown for Vero E6 (**B**) and Calu-3 2B4 (**D**). Results are representative of two experiments.

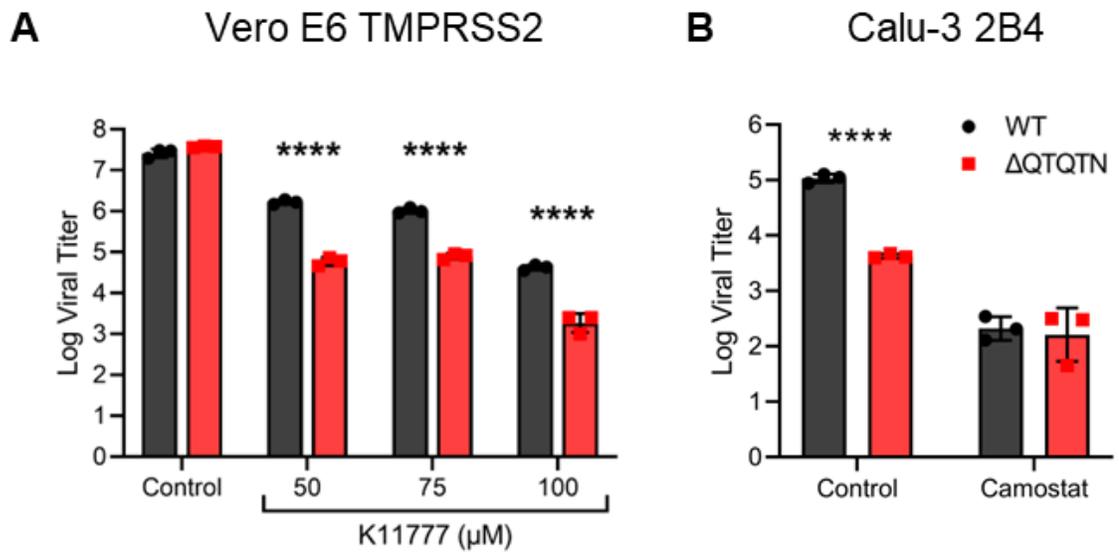
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.13. Inhibition of SARS-CoV-2 entry.**

Schematic of SARS-CoV-2 entry and use of host proteases. Inhibitors for TMRPSS2 (camostat mesylate) and cathepsin (K11777) indicated.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

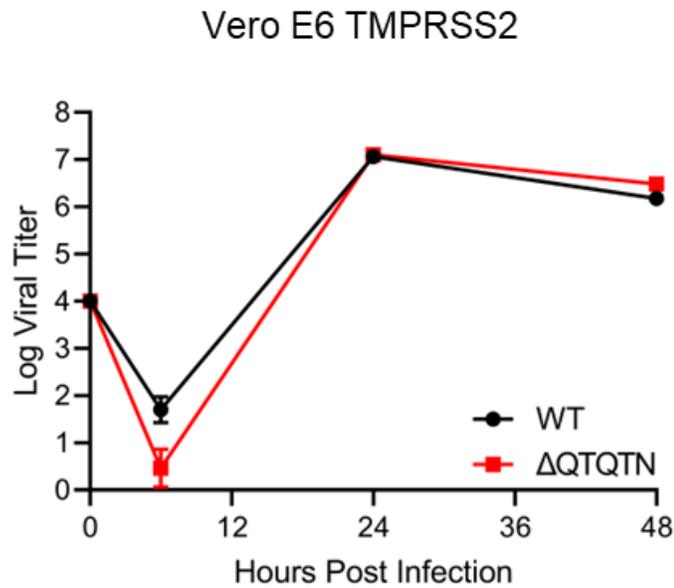


**Figure 2.14.  $\Delta$ QTQTN inefficiently uses TMPRSS2.**

(A) Viral titer at 24 hpi from TMPRSS2-expressing Vero E6 pretreated with varying doses of cathepsin inhibitor K11777 and infected with WT (black) or  $\Delta$ QTQTN SARS-CoV-2 (red) at an MOI of 0.01 (n=3).

(B) Viral titer at 24hpi from Calu-3 2B4 pretreated with 100  $\mu$ M of camostat mesylate and infected with WT or  $\Delta$ QTQTN SARS-CoV-2 at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test. \*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ; \*\*\*\*,  $p \leq 0.0001$ .

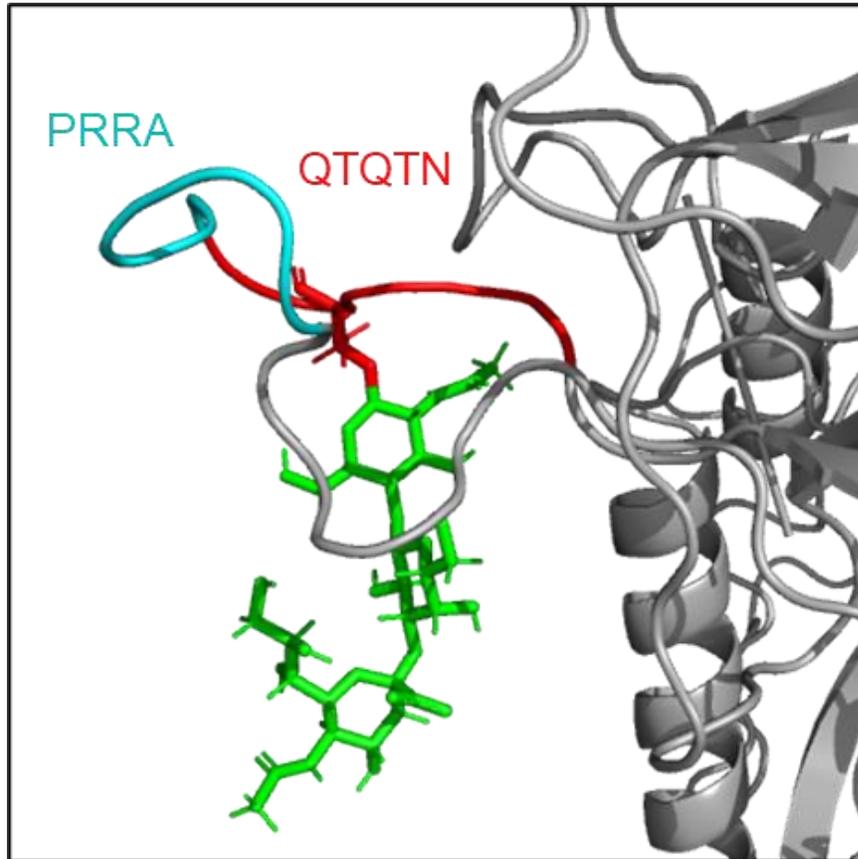
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.15.  $\Delta$ QTQTN SARS-CoV-2 replication in TMPRSS2-expressing Vero E6.**

Viral titer from TMPRSS2-expressing Vero E6 infected with WT or  $\Delta$ QTQTN SARS-CoV-2 at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test.

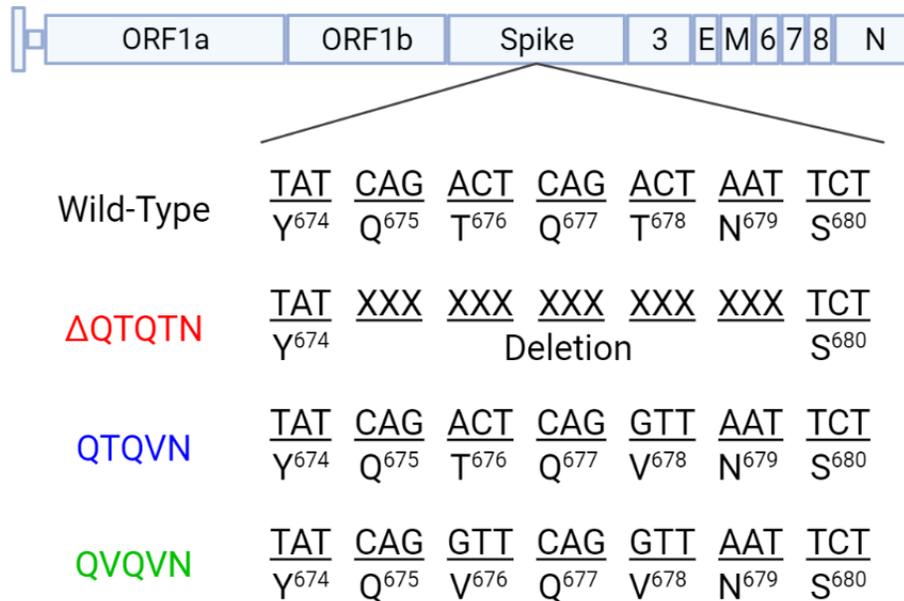
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.16. O-linked glycosylation of SARS-CoV-2 loop**

Structural modeling of O-linked glycosylation on threonine 678 (red) of QTQTN motif. PRRA (cyan) remains exposed.

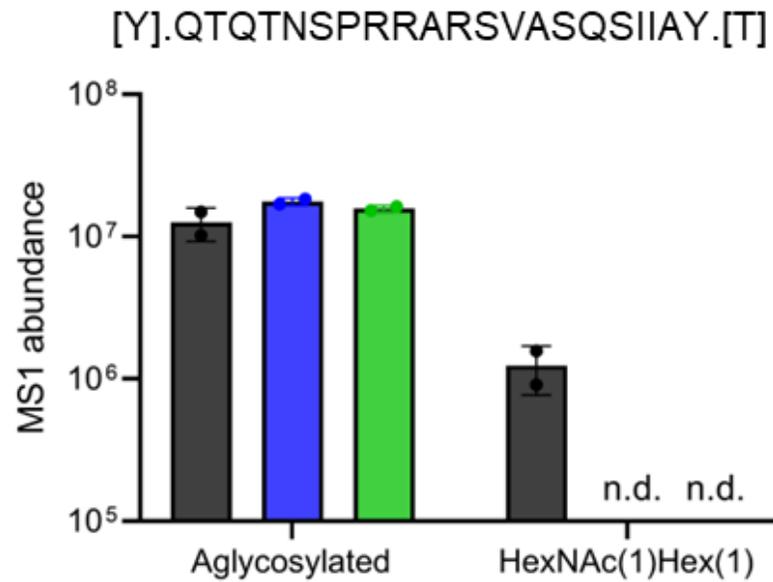
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.17. Cloning schematic of QTQTN glycosylation mutants.**

Schematic of SARS-CoV-2 genome with glycosylation mutations.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.18. Detection of glycosylation on QTQTN.**

MS1 label-free abundance of WT (black), QTQVN (blue), and QVQVN (green) SARS-CoV-2 spike chymotrypsin peptide Q675-Y975 from sucrose cushion-purified virions. n.d., not detected.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

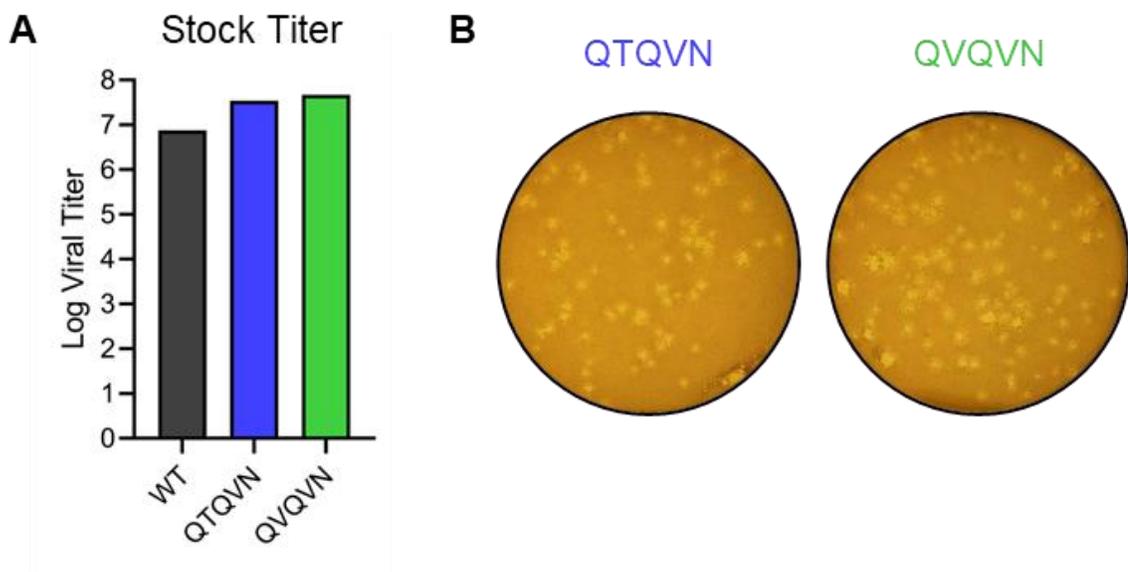


Virus	Coverage [%]	# Peptides	# Unique Peptides
SARS-CoV-2 WT	91	250	56
Annotated Sequence	Modifications	Abundance	
		ETHcD	HCD
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQNSPR.[R]	Carbamidomethyl [C3; C16; C25] Phospho [Y14] HexNAc(2)Hex(4)Fuc(1) [N11]	4.31E+05	1.76E+06
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQNSPR.[R]	Carbamidomethyl [C3; C16; C25] Deamidated [N12] Phospho [Y14] HexNAc(2)Hex(5)Fuc(1) [N11]	Not Detected	Below Quantitation Limit
Virus	Coverage [%]	# Peptides	# Unique Peptides
SARS-CoV-2 QTQVN	93	284	92
Annotated Sequence	Modifications	Abundance	
		ETHcD	HCD
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y14] 1xHexNAc(2)Hex(3)Fuc(1) [N11]	Not Detected	5.61E+05
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xDeamidated [N12] 1xPhospho [Y14] 1xHexNAc(2)Hex(3)Fuc(1) [N11]	Not Detected	5.61E+05
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y14] 1xHexNAc(2)Hex(5)Fuc(1) [N11]	3.51E+06	6.46E+06
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y14] 1xHexNAc(2)Hex(2)Fuc(1) [N11] 1xHexNAc(1)Hex(1)Fuc(1) [S13]	Not Detected	Below Quantitation Limit
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y] 1xHexNAc(2)Hex(1)Fuc(2) [S] 1xHexNAc(3)Hex(3) [S/T] 1xHexNAc(1)Hex(1)Fuc(1) [S13] 1xHexNAc(4)Hex(3)Fuc(1) [S27]	1.93E+06	1.24E+06
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQTQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [S13] 1xHexNAc(4)Hex(3)Fuc(1) [N11] 1xHexNAc(2)Hex(1)Fuc(1) [N11] 1xHexNAc(2)Hex(1)Fuc(2) [S27] 1xHexNAc(4)Hex(3)Fuc(2) [S27]	Below Quantitation Limit	Not Detected
Virus	Coverage [%]	# Peptides	# Unique Peptides
SARS-CoV-2 QVQVN	93	272	78
Annotated Sequence	Modifications	Abundance	
		ETHcD	HCD
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQVQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y14] 1xHexNAc(2)Hex(4)Fuc(1) [N11]	5.49E+05	7.28E+05
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQVQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xDeamidated [N12] 1xHexNAc(3)Hex(5) [N11]	1.21E+05	2.32E+05
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQVQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xPhospho [Y] 1xHexNAc(2)Hex(5)Fuc(1) [N11]	3.28E+05	1.59E+06
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQVQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xDeamidated [N12] 1xPhospho [Y14] 1xHexNAc(2)Hex(5)Fuc(1) [N11]	Not Detected	Below Quantitation Limit
[R].AGCLIGAEHVNNSEYCDIPIGAGICASYQVQVNSPR.[R]	3xCarbamidomethyl [C3; C16; C25] 1xHexNAc(4)Hex(7)NeuAc(1) [N11]	Not Detected	Below Quantitation Limit

**Table 2.1. Quantification of Nanoflow-LC-MS/MS analysis of QTQTN glycosylation mutants.**

Modifications and quantification of peptide spanning Spike 647-682 for WT (grey), QTQVN (blue) and QVQVN (green).

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

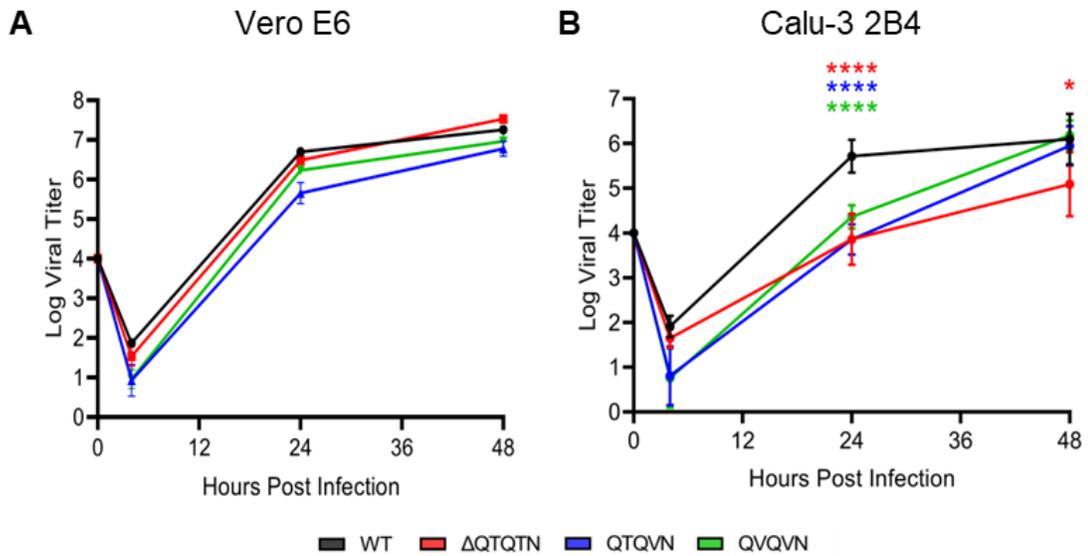


**Figure 2.20. Plaque morphology of QTQTN glycosylation mutants.**

(A) Virus stock titer of WT, QTQVN, and QVQVN SARS-CoV-2 in Vero E6.

(B) Plaque morphology of QTQVN and QVQVN in Vero E6.

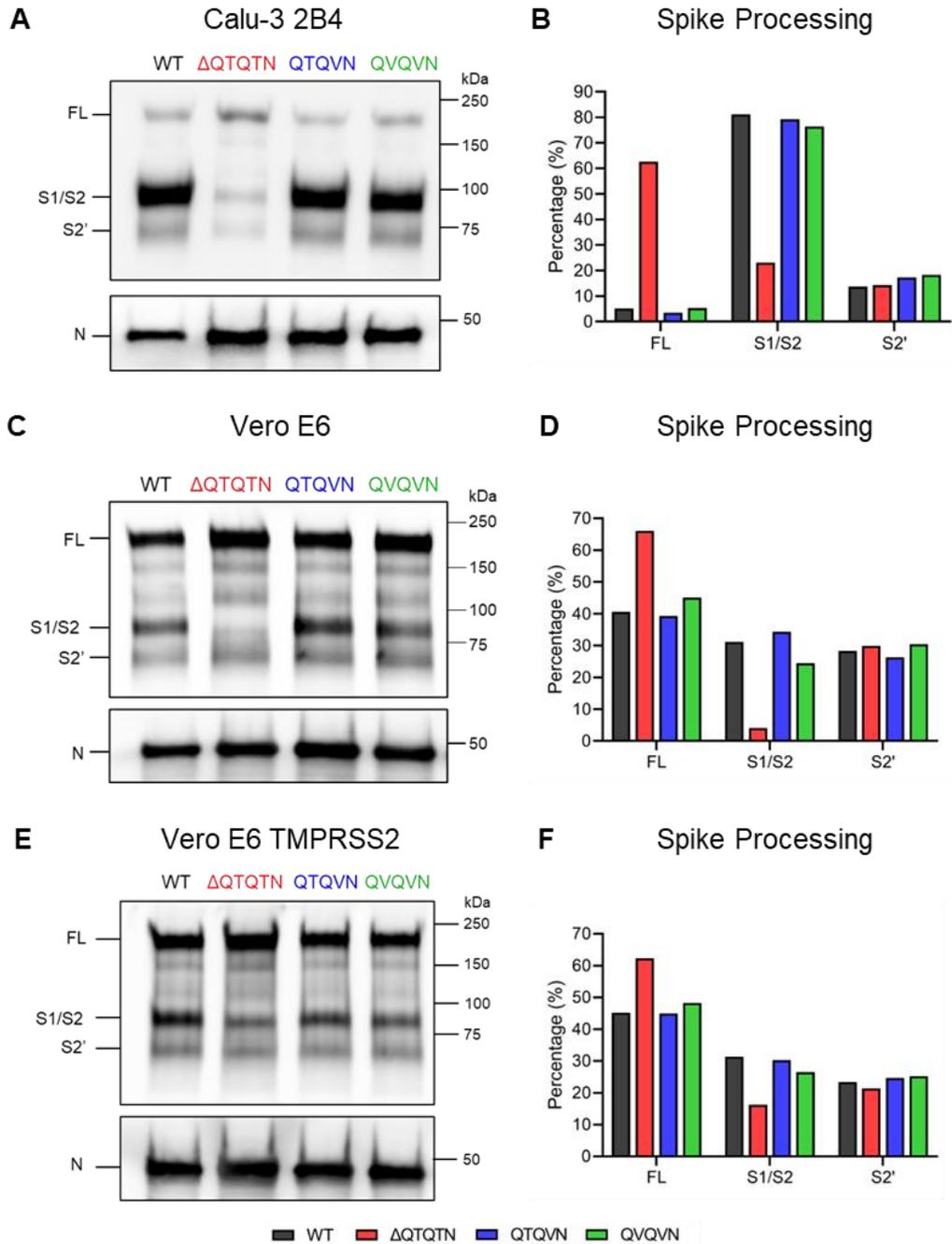
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.21. *In vitro* replication of QTQTN glycosylation mutants.**

(A) Viral titer from Vero E6 infected with WT (black),  $\Delta$ QTQTN (red), QTQVN (blue), or QVQVN (green) SARS-CoV-2 at an MOI of 0.01 (n=3). (B) Viral titer from Calu-3 2B4 infected with WT,  $\Delta$ QTQTN, QTQVN, or QVQVN SARS-CoV-2 at an MOI of 0.01 (n=6).

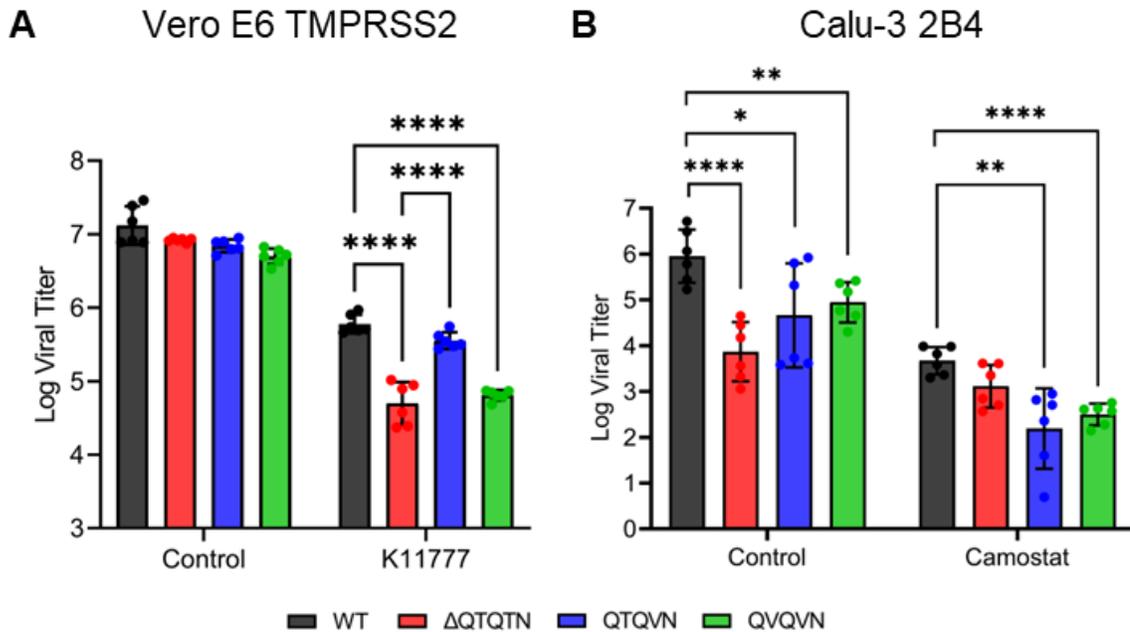
Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.



**Figure 2.22. Spike processing of QTQTN glycosylation mutants.**

Purified WT (black),  $\Delta$ QTQTN (red), QTQVN (blue), and QVQVN (green) SARS-CoV-2 virions from Calu-3 2B4 (A-B), Vero E6 (C-D) and TMRPSS2-expressing Vero E6 (E-F) probed with anti-S or anti-N antibody. Full-length (FL), S1/S2 cleavage product, and S2' cleavage product are indicated. Quantification of densitometry of FL, S1/S2, and S2' normalized to N shown for Calu-3 2B4 (B), Vero E6 (D), and Vero E6 TMRPSS2 (F). Results are representative of two experiments.

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

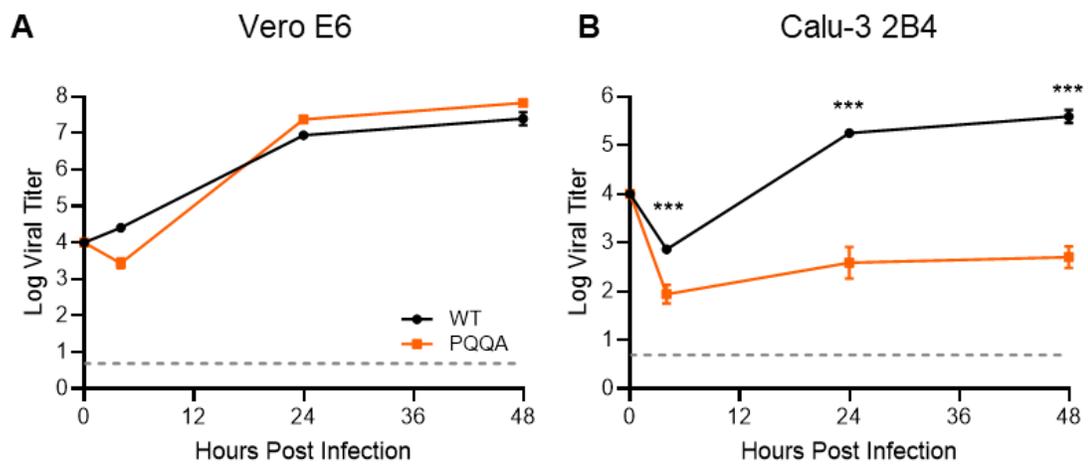


**Figure 2.23. Glycosylation of QTQTN motif influences protease usage.**

(A) Viral titer at 24 hpi from TMPRSS2-expressing Vero E6 pretreated with 50  $\mu$ M of K11777 and infected with WT,  $\Delta$ QTQTN, QTQVN, or QVQVN SARS-CoV-2 at an MOI of 0.01 (n=6).

(B) Viral titer at 24hpi from Calu-3 2B4 pretreated with 50  $\mu$ M of camostat mesylate and infected with WT,  $\Delta$ QTQTN, QTQVN, or QVQVN SARS-CoV-2 at an MOI of 0.01 (n=6). Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test. \*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ; \*\*\*\*,  $p \leq 0.0001$ .

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

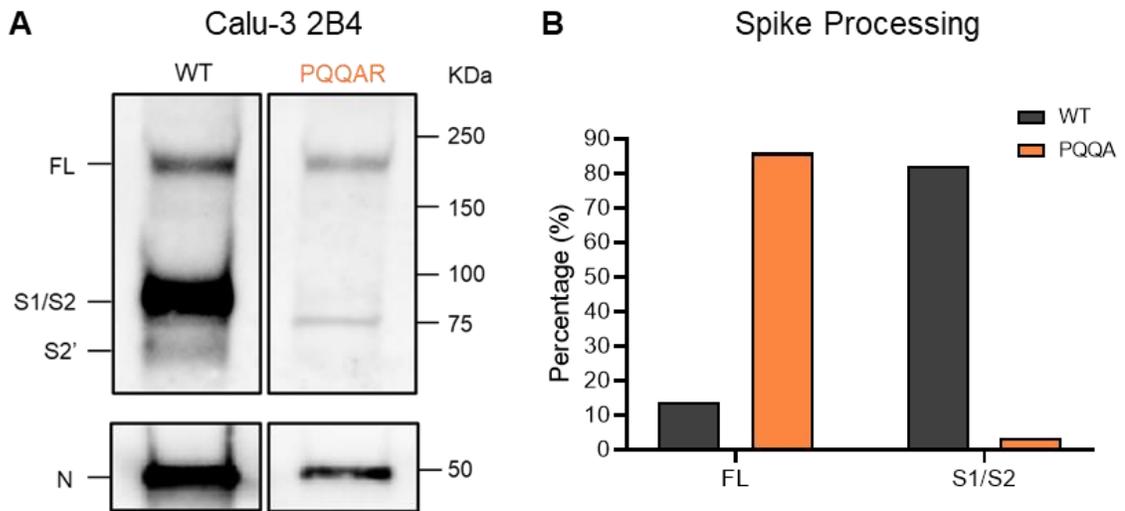


**Figure 2.24. *In vitro* replication of PQQA.**

**(A)** Viral titer from Vero E6 infected with WT (black) or PQQA (orange) SARS-CoV-2 at an MOI of 0.01 (n=3).

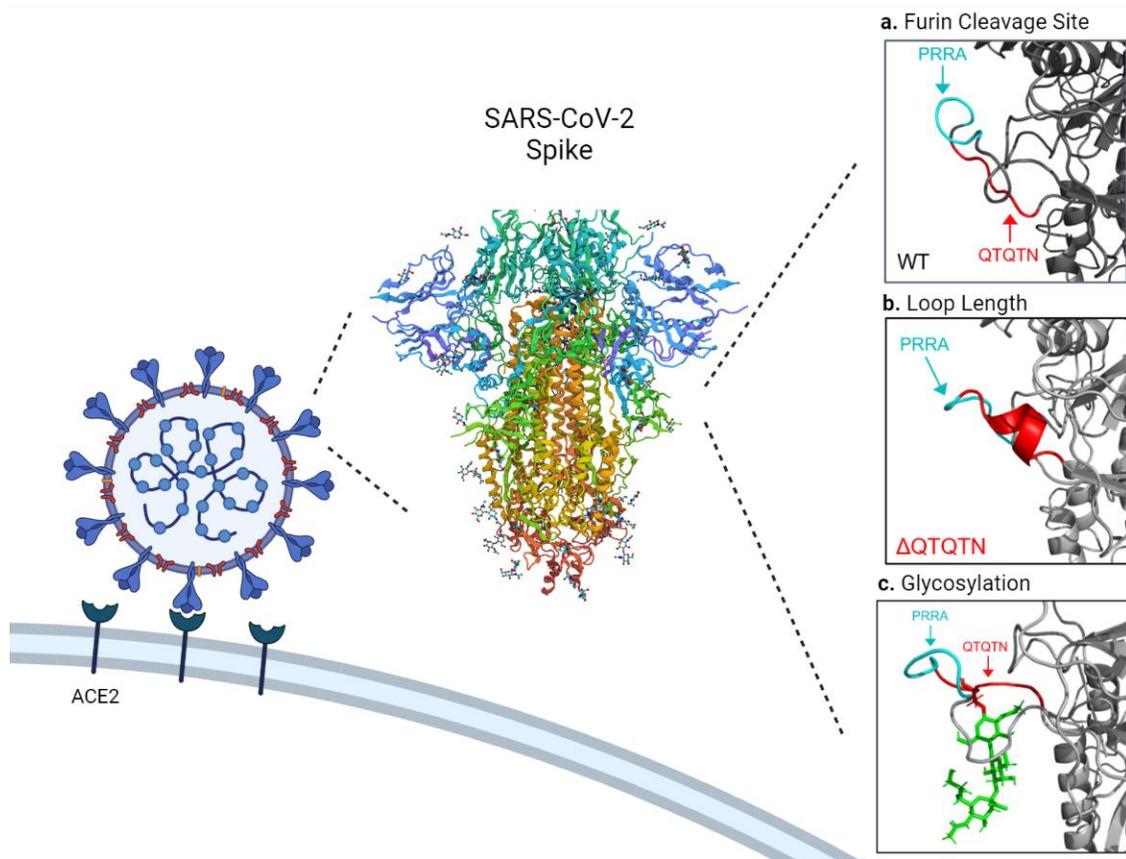
**(B)** Viral titer from Calu-3 2B4 infected with WT (black) or PQQA (orange) SARS-CoV-2 at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d.

Statistical analysis measured by two-tailed Student's t-test. \*,  $p \leq 0.05$ ; \*\*,  $p \leq 0.01$ ; \*\*\*,  $p \leq 0.001$ ; \*\*\*\*,  $p \leq 0.0001$ .



**Figure 2.25. Spike processing of PQQA.**

Purified WT and PQQA SARS-CoV-2 virions from Calu-3 2B4 (A-B) probed with anti-S or anti-N antibody. Full-length (FL), S1/S2 cleavage product, and S2' cleavage product are indicated. Quantification of densitometry of FL and S1/S2 cleavage product normalized to N shown (B). Results are representative of two experiments.



**Figure 2.26. Characteristics of SARS-CoV-2 S1/S2 cleavage site for efficient infection.**

The SARS-CoV-2 S1 cleavage site contains multiple components required for efficient infection and virulence: the furin cleavage site (FCS), PRRA, is important for spike processing (A); the loop length/composition manages FCS accessibility and protease interaction (B); and glycosylation is involved in protease interaction (C). SARS-CoV-2 spike protein (Walls et al., 2020).

Reproduced from Vu et al. (2022) *PNAS* with permission<sup>95</sup>.

### Chapter 3: Interplay between Omicron CTS1 mutations

The following manuscript is being prepared for submission.

#### INTRODUCTION

Since its introduction, SARS-CoV-2 has continuously evolved giving rise to multiple Variants of Concerns (VOCs) with diverse mutations in the spike protein <sup>96</sup>. Presented as a trimer on virions, spike is comprised of the S1 and S2 subunits, responsible for receptor binding and membrane fusion, respectively <sup>1,65</sup>. The S1 subunit contains the N-terminal domain (NTD), receptor binding domain (RBD), and the C-terminal of the S1 subunit (CTS1) which harbors a furin cleavage site in SARS-CoV-2. Following receptor binding, the spike is cleaved at the S1/S2 site by host proteases to expose the fusion machinery for entry. With the diverse mutations in spike, most Omicron studies have focused on the spike RBD and the impact on vaccine efficacy. However, mutations surrounding the FCS and S1/S2 cleavage site have been demonstrated to drive SARS-CoV-2 pathogenesis <sup>53,68,70,95,97,98</sup> and have been largely unstudied.

With this in mind, we set out to evaluate the role of Omicron CTS1 mutations on infection and pathogenesis. Omicron maintains three mutations adjacent to the FCS and S1/S2 cleavage site: H655Y, N679K, and P681H. Both H655Y and P681H have both previously been observed individually in the Gamma and Alpha variants <sup>99,100</sup>; in contrast, N679K is unique to and maintained by all Omicron subvariants <sup>101</sup>. To evaluate the role of these mutations, we used reverse genetics to generate SARS-CoV-2 mutants with all three CTS1 mutations (YKH) or N679K alone in the original WA1 backbone. The YKH mutant enhanced viral replication and spike processing, similar to prior reports with H655Y and P681H individually <sup>99,100</sup>. In contrast, the N679K mutant attenuated SARS-CoV-2 *in vitro*

and *in vivo*. Mechanistically, the N679K mutation decreased spike protein on virions, in infected cells, and from exogenous expression. Collectively, our studies suggest that N679K functions as a loss of function mutation through decreased spike protein resulting in attenuated Omicron replication and pathogenesis. Given the importance of spike protein for immunity, our finding may have major implications for vaccine efficacy and breakthrough infections.

## RESULTS

### **H655Y, N679K, and P681H together increase viral replication and spike processing.**

While the majority of the 30+ spike mutations Omicron acquired are localized to the RBD, three are harbored in the CTS1 adjacent to the furin cleavage site – H655Y, N679K, and P681H (**Figure 2.27**). Both H655Y and P681H have individually been observed in Gamma and Alpha variants and are associated with increased spike processing. In contrast, N679K is a unique mutation to Omicron and is maintained in all subsequent Omicron subvariants despite being a single nucleotide change (T/C to A/G) in the wobble position<sup>101</sup>. Importantly, N679K is adjacent to an important O-linked glycosylation site at T678<sup>79,80</sup>; our group has previously shown this glycosylation is important for SARS-CoV-2 infection and protease usage<sup>95</sup>.

Several motifs within the CTS1 spike domain, including the furin cleavage site and the upstream QTQTN motif, are key to spike cleavage and host protease interactions which drive SARS-CoV-2 infection and pathogenesis. All three Omicron mutations in the CTS1, H655Y, N679K, and P681H, are adjacent or within these motifs and may impact their function (**Figures 2.27 and 2.28**). To evaluate this question, we generated a mutant SARS-CoV-2 harboring H655Y, N679K, and P681H in the original WA1 backbone (YKH)

(**Figure 2.29**)<sup>77,102</sup>. Plaques produced by the YKH mutant are smaller in size compared to wild type WA1 (WT) (**Figure 2.30A**). However, the YKH mutant did not attenuate stock titers nor replication kinetics in Vero E6 cells as compared to WT SARS-CoV-2 (**Figures 2.30B and 2.31A**). Notably, while replication was slightly reduced at 24 hpi, endpoint titers for YKH were augmented at 48 hpi in Calu-3 2B4 cells compared to WT. The results suggests that the combination of the three mutations alters infection dynamics which may offer some advantages to the Omicron variant in human respiratory cells (**Figure 2.31B**). As H655Y and P681H have individually been shown to increase spike processing, we next evaluated spike processing on purified virions from YKH and WT infection. Similar to Delta and Omicron, YKH spike was more processed than to WT (**Figures 2.32A-B**). At 24 hpi, the S1/S2 cleavage to full length spike ratio was ~2.4:1 for the YKH spike (55% S1/S2 product, 23% full-length); in contrast, WT had a roughly equivalent amounts of S1/S2 product and full length. Overall, the combination of H655Y, N679K, and P681H in the YKH mutant resulted in increased viral endpoint yields in human respiratory cells and contributes to Omicron's enhanced spike processing.

#### **N679K mutation attenuates SARS-CoV-2 infection.**

The increase in spike processing found in the YKH mutant is consistent with prior work examining H655Y and P681H mutations individually; however, the contribution of N679K had yet to be evaluated. Based on its location adjacent to a key O-linked glycosylation site, we expected N679K may impact SARS-CoV-2 infection (**Figure 2.33**). To evaluate change, we generated a SARS-CoV-2 mutant in the original WA1 backbone (N679K) (**Figure 2.34**). Our initial characterization found that the N679K plaque sizes were distinctly at days 2 and 3 post infection (**Figure 2.35A-B**) and stock titers were

slightly lower than WT (**Figure 2.35C**). These changes in plaque size and stock titers are consistent with observation of most of the omicron strains <sup>103-106</sup>. Notably, unlike the minimal differences seen in YKH growth kinetics, the N679K human had attenuated replication in both Vero E6 and Calu-3 2B4 at 24 hpi (**Figures 2.36A-B**). While recovering at 48 hpi, the results suggest that N679K is a loss of function mutation in terms of replication in both cell lines.

We next evaluated N679K *in vivo* infecting 3- to 4-week-old golden Syrian hamsters and monitoring weight loss and disease over 7 days (**Figure 2.37A**). Hamsters infected with N679K displayed significantly attenuated body weight loss compared to those infected with WT (**Figure 2.37B**). Despite the stark attenuation seen in weight loss, N679K viral titers were equivalent to WT at 2dpi and 4dpi in the lungs (**Figure 2.37C**). Similarly, N679K viral titer was comparable to WT at 2dpi in nasal washes; however, the mutant virus was attenuated in replication at 4dpi (**Figure 2.37D**). Taken together, N679K has a distinct loss of function phenotype *in vitro* and *in vivo*; given this result, N679K may also function to mitigate the impact H655Y and P681H mutations.

### **N679K mutation results in decreased spike protein expression.**

We next set out to determine the mechanism driving the loss of function observed with the N679K mutant. Given its location adjacent to the furin cleavage site, we first evaluated N679K effects on spike processing. Virions were purified from WT, N679K or the Omicron variant BA.1 (Omicron) and probed for spike processing. Nearly identical to YKH, the N679K mutant had increased spike processing with a ~2.5:1 ratio of S1/S2 cleavage to full-length site compared to 1:1 for WT; at 24 hpi (**Figures 2.38A-B**). However, we noted distinct differences in total spike protein with N679K and Omicron

compared to WT despite similar levels of nucleocapsid protein. Densitometry analysis revealed that the total spike to nucleocapsid (S/N) ratio incorporated in N679K and Omicron virions was reduced 21% and 36%, respectively, as compared to WT (**Figure 2.38C**). Overall, our results indicate that the N679K mutant and Omicron variant incorporate less spike protein onto their virions.

Next, we sought to determine if changes in the virion spike were due to changes to total protein expression in the cell or incorporation into the particle. To examine spike protein expression, we measured total spike relative to nucleocapsid in infected Vero E6 cell lysate 24 hpi (**Figures 2.39A-B**). N679K resulted in a total S/N ratio 66% less than WT, displaying an even further decrease in spike protein compared to purified virions. Additionally, a similar decreased S/N ratio was observed in Omicron, indicating that the phenotype is maintained in the context of all Omicron mutations (**Figures 2.39A-B**). Together, the results indicate that the N679K mutation reduces the Omicron spike protein levels compared to WT.

Having established reduced spike protein in the context of N679K, we next set out to determine if this reduction only occurs in the context of virus infection or is inherent to the protein. Therefore, we introduced the mutation into the Spike HexaPro plasmid to exogenously express spike and separate from other aspects of viral infection<sup>107</sup>. Vero E6 cells were transfected with the WT or N679K mutant spike HexaPro and harvested at 24 and 48 hours post transfection (hpt). Similar to what was observed in viral infection, N679K spike was reduced 43% at 24 hpt and 46% at 48 hpt (**Figures 2.40A-B**). Overall, the results suggest that the loss of Omicron spike is governed by N679K in a manner independent of viral infection.

## DISCUSSION

Most Omicron studies have focused on determining the impact the RBD mutations have on immune escape, largely overlooking mutations in other spike domains like the CTS1. Harboring the FCS and S1/S2 cleavage site, the CTS1 has been demonstrated as a hotspot for attenuating and augmenting mutations<sup>53,68,70,95,97,98</sup>. Focusing on Omicron's three CTS1 mutations – H655Y, N679K, and P681H, we generated infectious clones with all three (YKH) or N679K alone in the SARS-CoV-2 WA1 background. The combination of YKH produced a modest increase in endpoint titer in human respiratory cells and augmented spike processing, consistent with prior studies on H655Y and P681H individually<sup>99,100</sup>. However, the N679K mutant was attenuated viral replication *in vitro* and disease *in vivo*. Mechanistic studies determined that both N679K and Omicron have reduced spike incorporated into their virions, less spike in infected cell lysates, and inferior production using exogenous expression systems compared to WT. Our results argue reduced spike protein in the context the N679K mutation attenuates Omicron strains and may have implication for SARS-CoV-2 immunity.

N679K is likely attenuated due its decrease in spike protein. Starting with ~30% less spike in its virions, one possibility was a change in spike incorporation into the virions. However, an even greater decrease (66%) in Vero infected cell lysates indicated that spike protein levels were an issue. To confirm the reduction in spike was not a product of virus infection or host interactions, we used exogenous expression to demonstrate that the spike protein itself was less stable than control. N679K, a single nucleotide and amino acid change, is unlikely to alter spike RNA expression. Instead, protein stability may have been altered, although the S1/S2 loop is an ordered structure. One possible mechanism is the

asparagine to lysine change introduces a ubiquitination site that could lead to degradation. Another possible mechanism leading to decreased spike protein is that the N679K mutation may destabilize the protein. Similarly, N679K adds another basic amino acid to the stretch including the FCS; the additional positively charged lysine extends the polybasic cleavage motif and may facilitate activation by additional host proteases<sup>108</sup>. Overall, while the exact mechanism is unclear, the N679K mutation results in less stable spike that impacts infection and pathogenesis of SARS-CoV-2.

Surprisingly, N679K is uniformly maintained in 100% of the Omicron sequences in GSAID, despite being a single nucleotide change in the wobble position<sup>101</sup>. Though attenuated *in vitro*, N679K does replicate to similar titers as WT in lungs and at day 2 in nasal washes. These results suggest no deficits in transmission and may explain how N679K is maintained despite clear attenuation of SARS-CoV-2 infection. Notably, addition of H655Y and P681H in the YKH mutant rescues replication in Calu-3 cells and suggest that other Omicron mutations may compensate for N679K. However, it is unclear if reverting N679K in the Omicron strains would result in a gain-of-function in terms of *in vitro* replication or *in vivo* pathogenesis. While N679K in SARS-CoV-2 WA1 produces a clear loss of function, the constellation of mutations and epistatic interactions may mitigate the deficit in Omicron. Importantly, the complete conservation of N679K in Omicron also implies some advantage is conferred; this may be in transmission, replication in certain cells/tissues, or evasion of host immune responses<sup>109-114</sup>. While more studies are required to decipher why it is maintained, N679K is clearly a mutation that impacts overall spike protein expression.

In addition to impacting primary infection, the reduction in spike protein may have important implications for SARS-CoV-2 immunity. Compared to WT, the N679K mutation produce less spike upon infection and can potentially skew the antibody ratio targeting spike and nucleocapsid. Prior work with SARS-CoV had shown that an altered spike/nucleocapsid antibody ratio contributed to vaccine failure in aged mice <sup>6</sup>. Therefore, infection with Omicron could increase N targeting antibodies at the expense of spike antibodies. The result would be less protective neutralizing antibody which may facilitate more breakthrough infections. Furthermore, SARS-CoV-2 vaccines based on the Omicron spike may produce less spike protein due to N679K mutation. In the context of the mRNA bivalent vaccines, the 1:1 ratio of WT to Omicron spike may bias responses towards WT spike instead of equally between both spikes. In addition, the total amount of spike produced may be less than previous vaccines formulations, thus diminishing the antibody response. These factors potentially contribute to the less than expected induced immunity against Omicron despite the new vaccine formulations. Moving forward, reverting K679 back to N679 may improve spike protein yields and subsequently improve vaccine response to the Omicron variants.

Together, the results demonstrate that Omicron N679K is a loss of function mutation consistently maintained in subvariants. Mechanistically, the N679K mutation attenuates the virus *in vitro* and *in vivo* by increasing spike degradation. While attenuating in isolation, other Omicron mutations like H655Y and P681H may mitigate the N679K loss of function by amplifying spike processing and infection. However, the decreased spike expression by N679K may have implications on immunity induced by infection and vaccines through skewing the antibody responses. Overall, the interplay between the

Omicron CTS1 mutations has a significant impact on SARS-CoV-2 infection and pathogenesis and require further study for clarification.

## **METHODS**

### **Cell Culture**

Vero E6 cells were grown in high glucose DMEM (Gibco #11965092) with 10% fetal bovine serum and 1x antibiotic-antimycotic. TMPRSS2-expressing Vero E6 cells were grown in low glucose DMEM (Gibco #11885084) with sodium pyruvate, 10% FBS, and 1 mg/mL Geneticin<sup>TM</sup> (Invitrogen #10131027). Calu-3 2B4 cells were grown in high glucose DMEM (Gibco #11965092) with 10% defined fetal bovine serum, 1 mM sodium pyruvate, and 1x antibiotic-antimycotic.

### **Viruses**

The SARS-CoV-2 infectious clones were based on the USA-WA1/2020 sequence provided by the World Reference Center of Emerging Viruses and Arboviruses and the USA Centers for Disease Control and Prevention <sup>84</sup>. Mutant viruses (YKH and N679K) were generated with restriction enzyme-based cloning using gBlocks encoding the mutations (Integrated DNA Technologies) and our reverse genetics system as previously described <sup>76,77</sup>. Virus stock was generated in TMPRSS2-expressing Vero E6 cells to prevent mutations from occurring at the FCS. Viral RNA was extracted from virus stock and cDNA was generated to verify mutations by Sanger sequencing.

Delta isolate (B.1.617.2) was obtained from the World Reference Center of Emerging Viruses and Arboviruses. Infectious clone of Omicron (BA.1) was obtained from Dr. Pei Yong Shi and Dr. Xuping Xie.

### ***In vitro* Infection**

Vira infections in Vero E6 and Calu-3 2B4 were carried out as previously described<sup>95</sup>. Briefly, growth media was removed, and cells were infected with WT or mutant SARS-CoV-2 at an MOI of 0.01 for 45 min at 37°C with 5% CO<sub>2</sub>. After absorption, cells were washed three times with PBS and fresh complete growth media was added. Three or more biological replicates were collected at each time point and each experiment was performed at least twice. Samples were titrated with plaque assay or focus forming assays.

### **Plaque Assay**

Vero E6 cells were seeded in 6-well plates and grown to 80-100% confluency in complete growth media. Ten-fold serial dilutions in PBS were performed on virus samples. Growth media was removed from cells and 200 µl of inoculum was added to monolayers. Cells were incubated for 45 min at 37°C with 5% CO<sub>2</sub>. After absorption, 0.8% agarose overlay was added, and cells were incubated at 37°C with 5% CO<sub>2</sub> for 2 days. Plaques were visualized with neutral red stain. Average plaque size was determined using ImageJ.

### **Focus Forming Assay**

Focus forming assays (FFAs) were performed as previously described with some adaptations<sup>115</sup>. Briefly, Vero E6 cells were seeded in 96-well plates to be 100% confluent. Samples were 10-fold serially diluted in serum-free media and 20 µl was to infect cells. Cells were incubated for 45 min at 37°C with 5% CO<sub>2</sub> before 100 µl of 0.85% methylcellulose overlay was added. Cells were incubated for 24 h 45 min at 37°C with 5% CO<sub>2</sub>. After incubation, overlay was removed, and cells were washed three times with PBS before fixed and virus inactivated by 10% formalin for 30 min at room temperature. Cells

were then permeabilized and blocked with 0.1% saponin/0.1% BSA in PBS before incubated with  $\alpha$ -SARS-CoV-2 Nucleocapsid primary antibody (Cell Signaling Technology) at 1:1000 in permeabilization/blocking buffer overnight at 4°C. Cells are then washed three times with PBS before incubated with Alexa Fluor™ 555-conjugated  $\alpha$ -mouse secondary antibody (Invitrogen #A28180) at 1:2000 in permeabilization/blocking buffer for 1 h at room temperature. Cells were washed three times with PBS. Fluorescent foci images were captured using a Cytation 7 cell imaging multi-mode reader (BioTek), and foci were counted manually.

### **Hamster Infection**

Three- to four-week-old male golden Syrian hamsters (HsdHan:AURA strain) were purchased from Envigo. All studies were conducted under a protocol approved by the UTMB Institutional Animal Care and Use Committee and complied with USDA guidelines in a laboratory accredited by the Association for Assessment and Accreditation of Laboratory Animal Care. Procedures involving infectious SARS-CoV-2 were performed in the Galveston National Laboratory ABSL3 facility. Hamsters were intranasally infected with  $10^5$  pfu of WT or N679K SARS-CoV-2 in 100  $\mu$ l. Infected hamsters were weighed and monitored for illness over 7 days. Hamsters were anesthetized with isoflurane and nasal washes were collected with 400  $\mu$ l of PBS on endpoint days (2, 4, and 7 dpi). Hamsters were euthanized by CO<sub>2</sub> for organ collection. Nasal wash and lung were collected to measure viral titer and RNA. Left lungs were collected for histopathology.

### **Virion Purification**

Vero E6 cells were grown in T175 flasks to be 100% confluent at time of infection.

Cells were infected with 50  $\mu$ l of virus stock in PBS for 45 min at 37°C with 5% CO<sub>2</sub>, and growth media with 5% FBS was added after absorption. Supernatant was harvested at 24 hpi and clarified by low-speed centrifugation. Virions were purified from supernatant by ultracentrifugation through a 20% sucrose cushion at 26,000 rpm for 3 hrs using a Beckman SW28 rotor. Pellets were resuspended with 2x Laemmli buffer to obtain protein samples for Western blot.

### **Western Blot**

Protein levels were determined by SDS-PAGE followed by western blot analysis as previously described<sup>95</sup>. In brief, sucrose-purified SARS-CoV-2 virions were inactivated by resuspending in 2x Laemmli buffer and boiling. SDS-PAGE gels were run with equal volumes of samples on Mini-PROTEAN TGX gels (Bio-Rad #4561094) followed by transfer onto PVDF membrane. Membranes were incubated with  $\alpha$ -SARS-CoV S primary antibody (Novus Biologicals #NB100-56578) at 1:1000 dilution in 5% BSA in TBST to measure spike protein processing and expression. For loading control,  $\alpha$ -SARS Nucleocapsid primary antibody (Novus Biologicals #NB100-56576) at 1:1000 in 5% BSA in TBST was used for viral loading control and  $\alpha$ -GAPDH primary antibody (Invitrogen #AM4300) at 1:1000 in 5% BSA in TBST for cellular loading control. Primary antibody incubation was followed by HRP-conjugated  $\alpha$ -rabbit secondary antibody (Cell Signaling Technology #7074) or HRP-conjugated  $\alpha$ -mouse secondary antibody (Cell Signaling Technology #7076) at 1:3000 in 5% milk in TBST. Chemiluminescence signal was developed using Clarity Western ECL substrate (Bio-Rad #1705060) or Clarity Max Western ECL substrate (Bio-Rad #1705062) and imaged with a ChemiDoc MP System (Bio-Rad). Densitometry analysis was performed using ImageLab 6.0.1 (Bio-Rad).

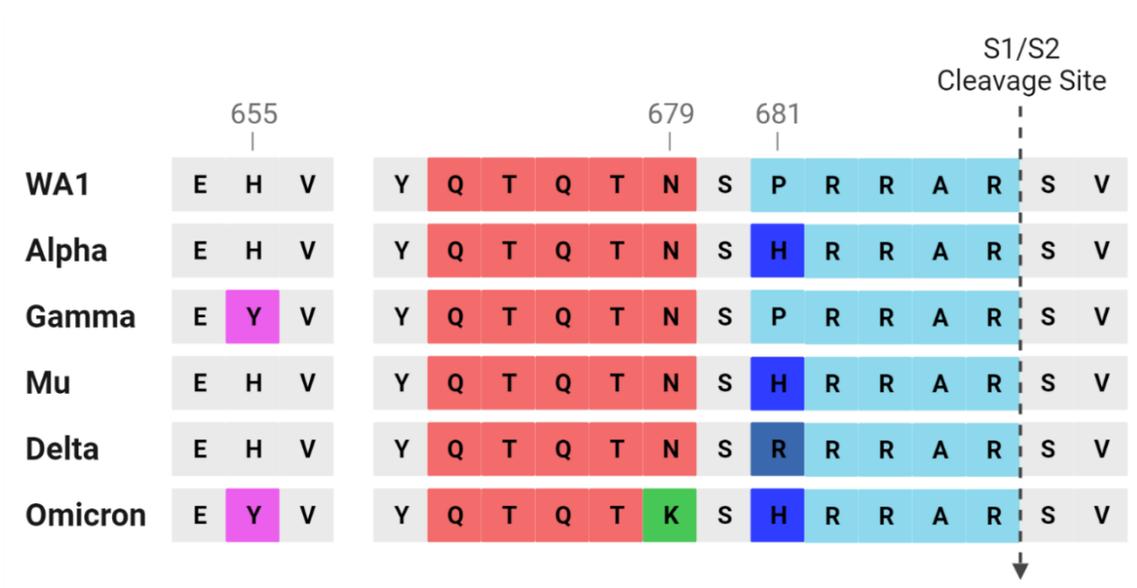
## **Spike HexaPro Cloning and Transfection**

SARS-CoV-2 S HexaPro was a gift from Jason McLellan (Addgene plasmid #154754)<sup>107</sup>. The N679K mutation was cloned into spike HexaPro using a gBlock encoding the mutation (Integrated DNA Technologies) and restriction enzyme-based cloning. Sequences were verified by Sanger sequencing.

Vero E6 cells were grown in 24-well plates to be 100% confluent at time of transfection. Cells were transfected with spike HexaPro WT or N679K plasmid and Lipofectamine 2000 following manufacturer's instructions (Invitrogen). Briefly, 100 ng of spike HexaPro plasmid and 1.5  $\mu$ l of Lipofectamine 2000 were separately diluted in 50  $\mu$ l Opti-MEM (Gibco #31985070) before mixing together. After 20 min of room temperature incubation, 100  $\mu$ l of the transfection mixture was added to cells, and cells were incubated at 37°C with 5% CO<sub>2</sub>. Cell lysate was harvested with 2x Laemmli buffer at 24 and 48 hours post transfection to be analyzed by Western blot.

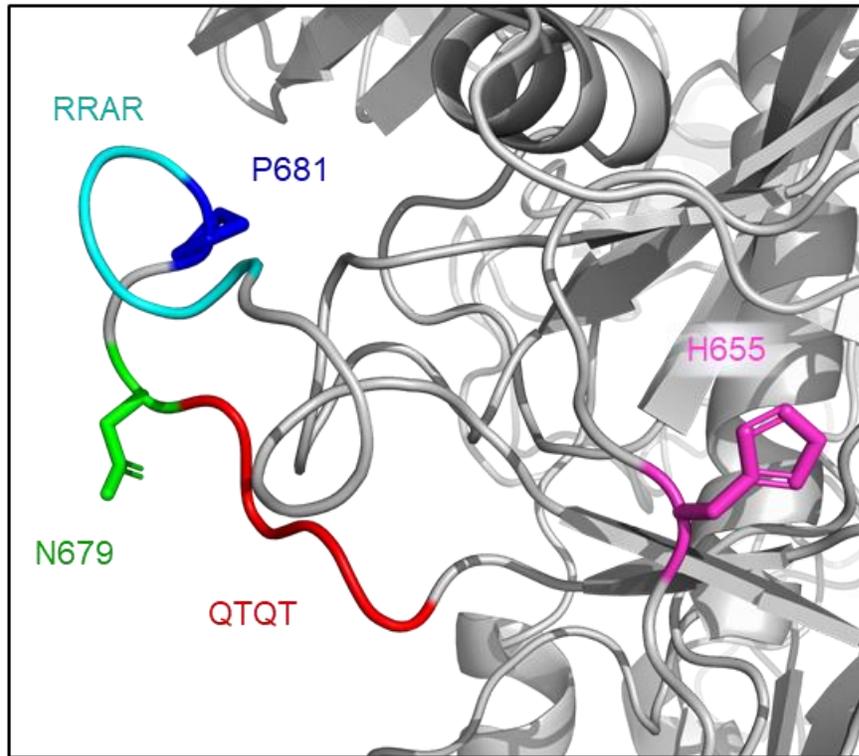
## **Structural Modeling**

Structural models previously generated were used as a base to visualize residues mutated in Omicron<sup>95</sup>. Briefly, structural models were generated using SWISS-Model to generate homology models for WT and glycosylated SARS-CoV-2 spike protein on the basis of the SARS-CoV-1 trimer structure (Protein Data Bank code 6ACD). Homology models were visualized and manipulated in PyMOL (version 2.5.4) to visualize Omicron mutations.



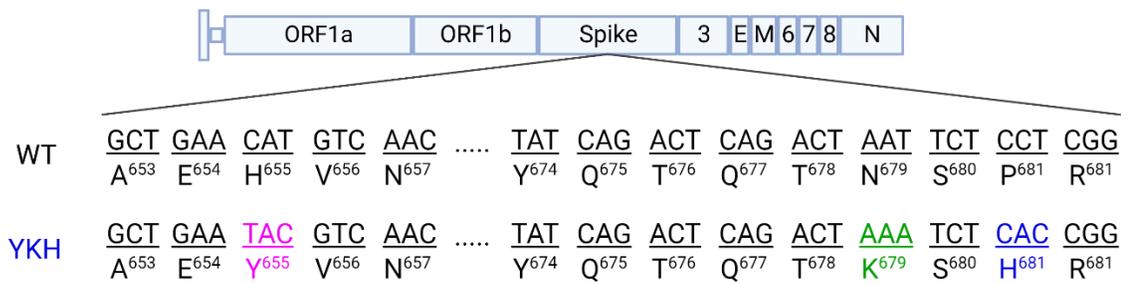
**Figure 2.27. Comparison of CTS1 across SARS-CoV-2 variants.**

Comparison of CTS1 region near the S1/S2 cleavage site between SARS-CoV-2 variants.



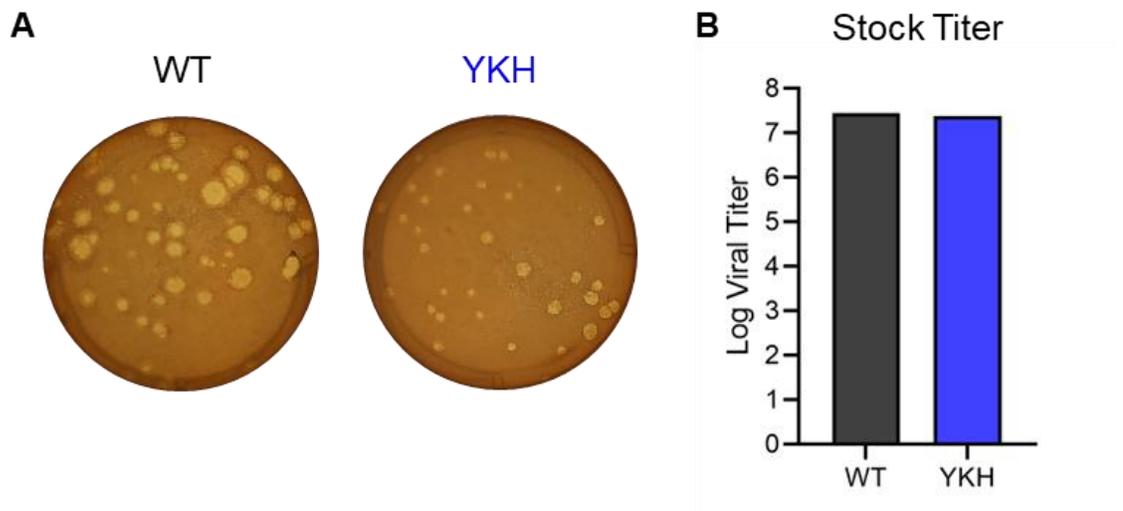
**Figure 2.28. Structural modeling of spike CTS1 residues mutated in Omicron.**

Structure of loop containing the S1/S2 cleavage site on SARS-CoV-2 spike protein. The residues that are mutated in Omicron are shown – H655 (magenta), N679 (green), and P681 (blue). The furin cleavage site RRAR (cyan) and QTQT motif (red) are also shown.



**Figure 2.29. Cloning schematic of SARS-CoV-2 YKH.**

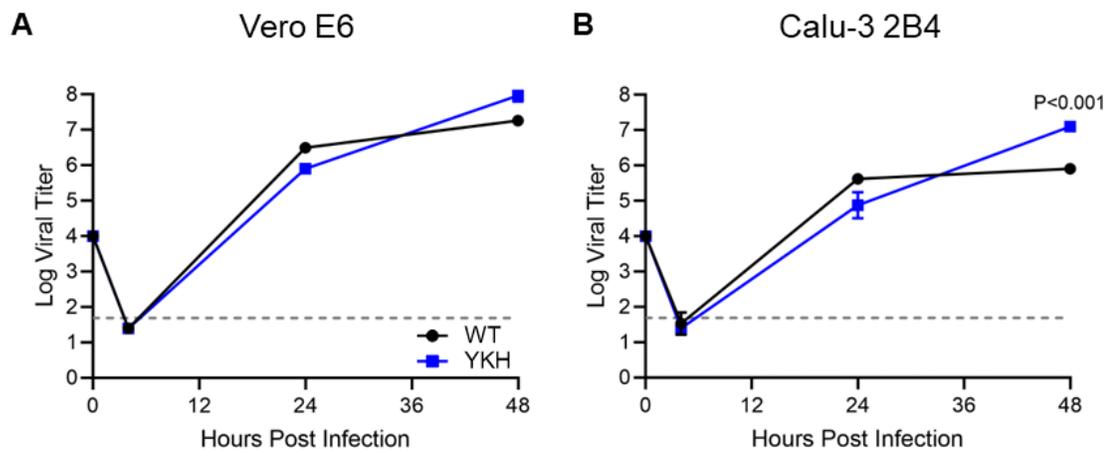
Schematic of WT and YKH SARS-CoV-2 mutant genomes.



**Figure 2.30. Plaque morphology of YKH.**

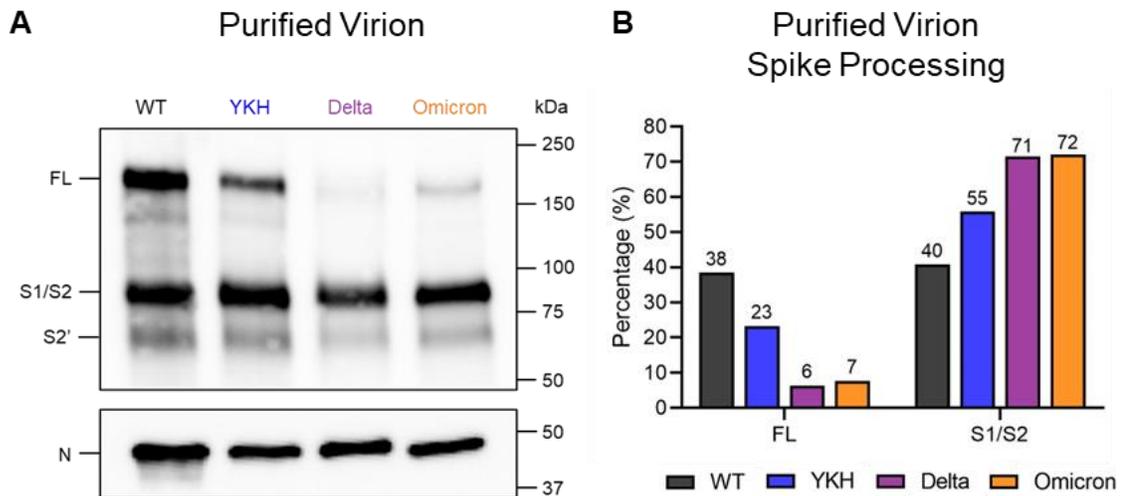
(A) WT and YKH SARS-CoV-2 plaques on Vero E6 cells at 2 dpi.

(B) Viral titer from WT and YKH virus stock with the highest yield generated from TMPRSS2-expressing Vero E6 cells.



**Figure 2.31. *In vitro* replication of YKH.**

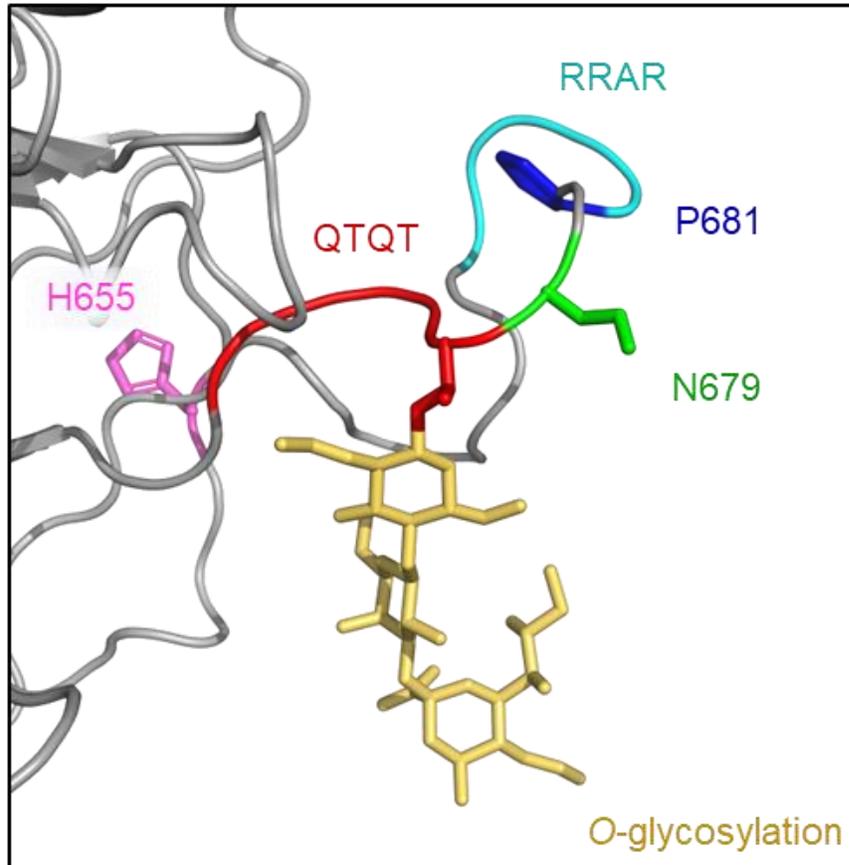
Growth kinetics of WT and YKH in Vero E6 (**A**) and Calu-3 2B4 (**B**) cells. Cells were infected at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test.



**Figure 2.32. Spike processing of YKH purified virions.**

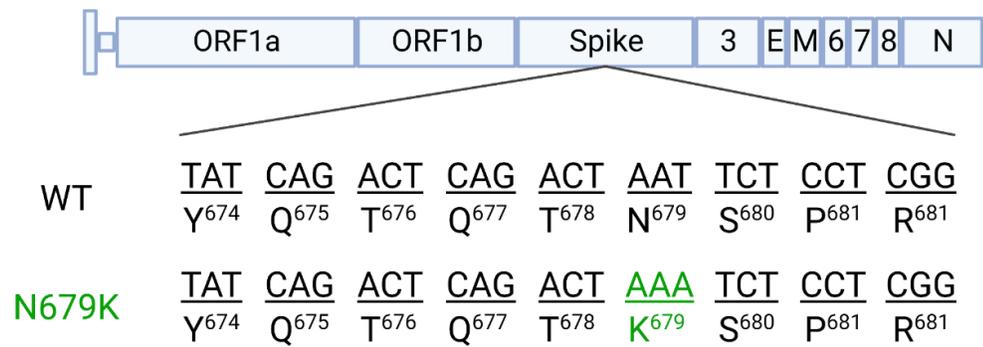
(A) Purified WT, YKH, Delta isolate (B.1.617.2), and Omicron (BA.1) virions from Vero E6 supernatant were probed with  $\alpha$ -Spike and  $\alpha$ -Nucleocapsid (N) antibodies in Western blots. Full-length spike (FL), S1/S2 cleavage product, and S2' cleavage product are indicated.

(B) Densitometry of FL and S1/S2 cleavage product was performed, and quantification of FL and S1/S2 cleavage product percentage of total spike is shown. Quantification was normalized to N for viral protein loading control. WT (black), YKH (blue), Delta isolate (purple), Omicron (orange). Results are representative of two experiments.



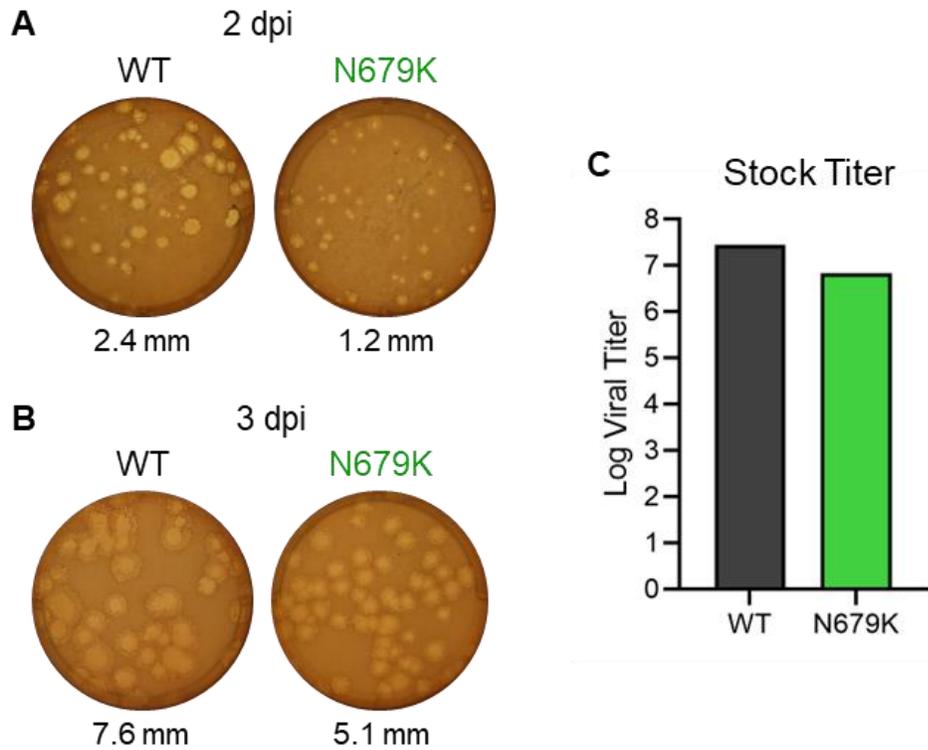
**Figure 2.33. Structural modeling of spike CTS1 residues mutated in Omicron adjacent to O-linked glycosylation.**

Structural modeling of O-linked glycosylation of threonine 678 (yellow) of QTQTN motif (red) and the residues mutated in Omicron – H655 (magenta), N679 (green), and P681 (blue) – with N679 adjacent to the glycosylation. The furin cleavage site RRAR is also shown (cyan).



**Figure 2.34. Cloning schematic of SARS-CoV-2 N679K.**

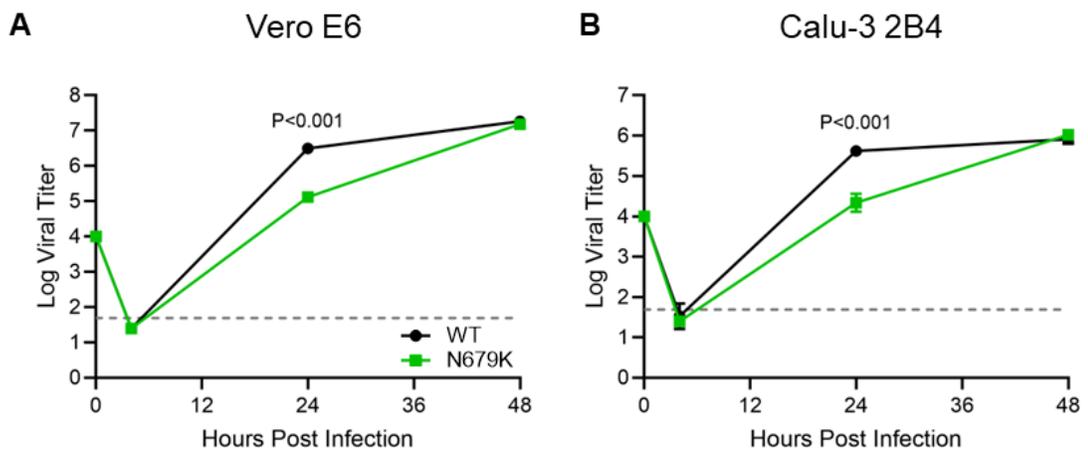
Schematic of WT and N679K SARS-CoV-2 mutant genomes.



**Figure. 2.35. Plaque morphology of N679K.**

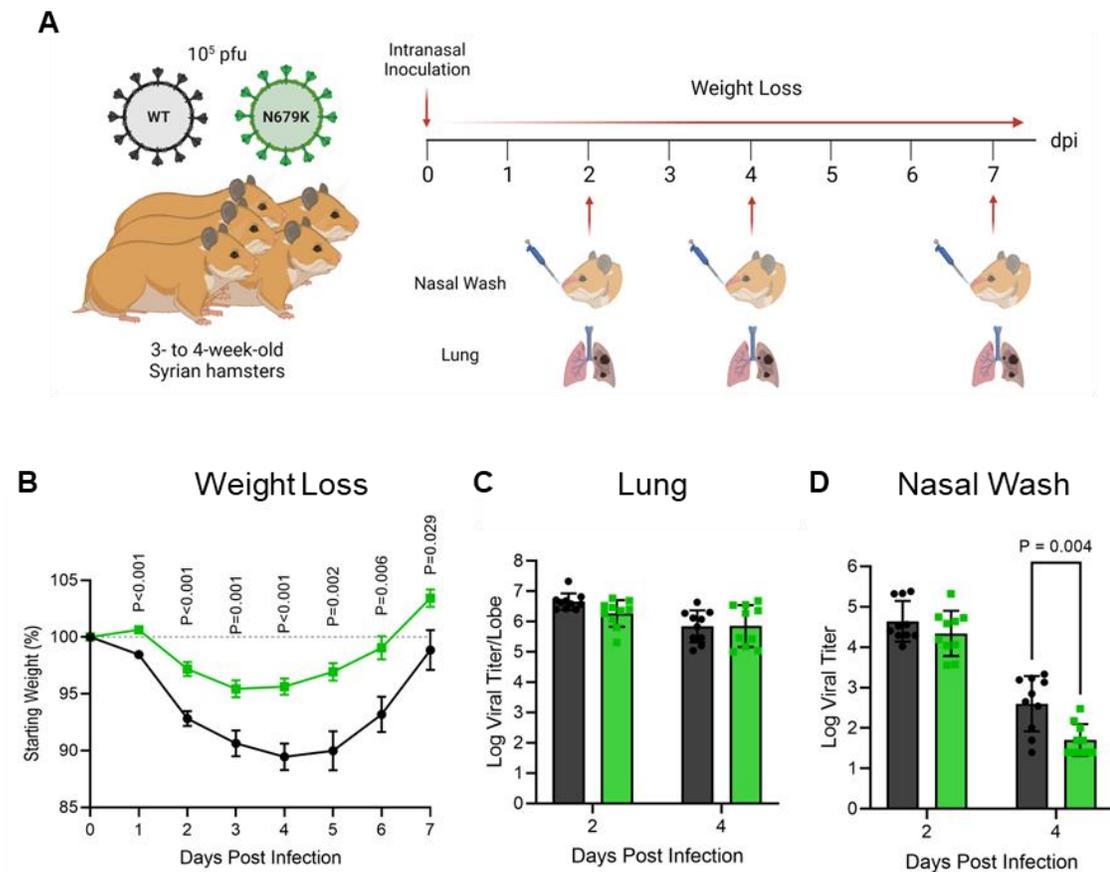
(A-B) WT and N679K SARS-CoV-2 plaques on Vero E6 cells at 2 dpi (A) and 3 dpi (B). Average plaque size noted below.

(C) Viral titer from WT and N679K virus stock with the highest yield generated from TMPRSS2-expressing Vero E6 cells.



**Figure 2.36. *In vitro* replication of N679K.**

Growth kinetics of WT and N679K in Vero E6 (**A**) and Calu-3 2B4 (**B**) cells. Cells were infected at an MOI of 0.01 (n=3). Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test.

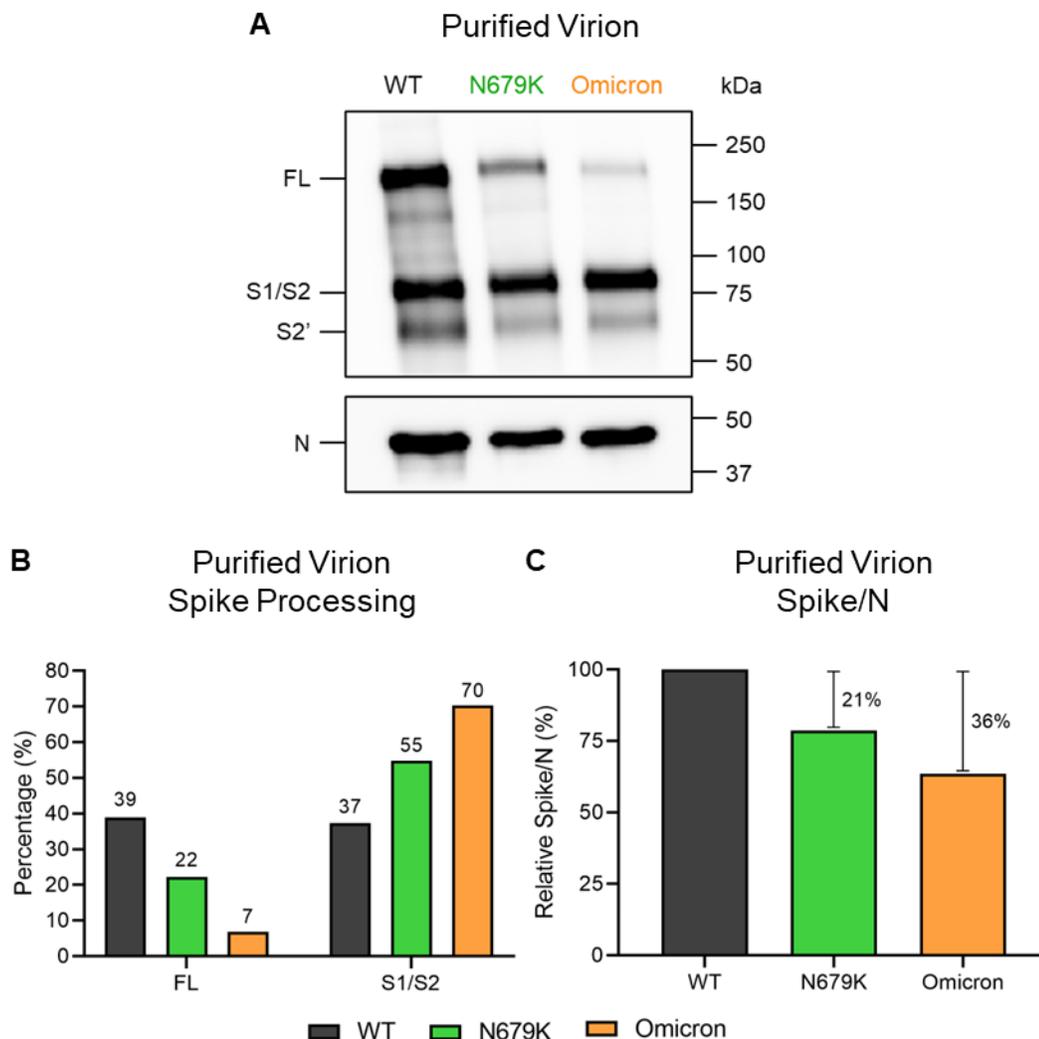


**Figure 2.37. *In vivo* characterization of SARS-CoV-2 N679K in golden Syrian hamsters.**

**(A)** Schematic of experiment design for golden Syrian hamster infection with WT (black) or N679K (green) SARS-CoV-2. Three- to four-week-old golden Syrian hamsters were infected with 10<sup>5</sup> pfu and monitored for weight loss over 7 days. At 2, 4, and 7 dpi, nasal wash and lung were collected for viral titer, and lung was collected for histopathology.

**(B)** Weight loss of hamsters infected with WT (black) or N679K (green) SARS-CoV-2 over 7 days. Data are mean  $\pm$  s.e.m. Statistical analysis measured by two-tailed Student's t-test.

**(C-D)** Viral titer of lung **(C)** and nasal wash **(D)** collected at 2 and 4 dpi from hamsters infected with WT (black) or N679K (green) SARS-CoV-2. Data are mean  $\pm$  s.d. Statistical analysis measured by two-tailed Student's t-test.

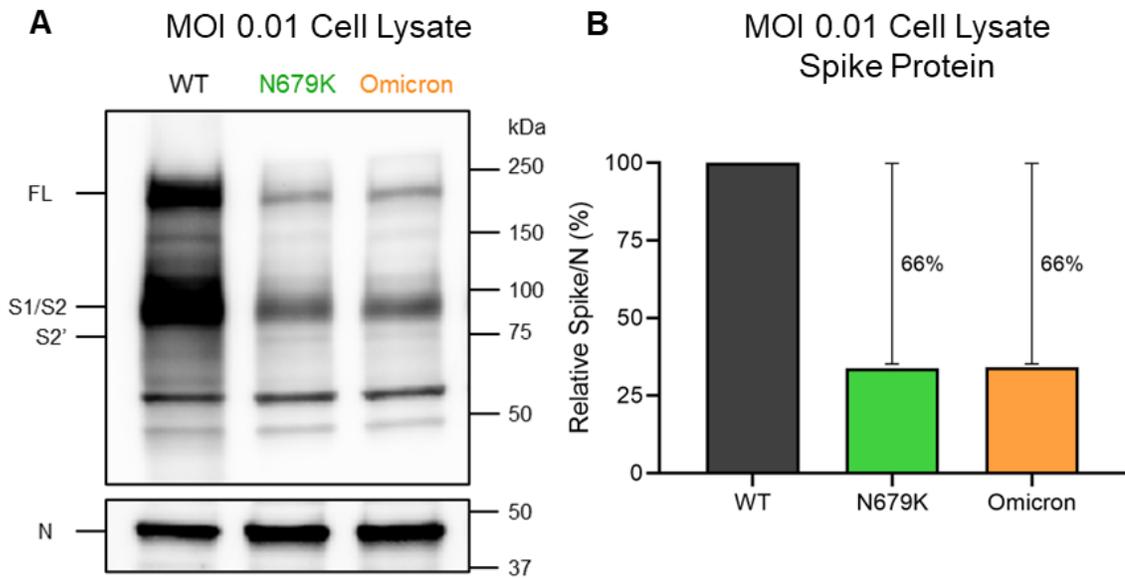


**Figure 2.38. Spike processing and expression in N679K purified virions.**

(A) Purified WT, N679K, and Omicron (BA.1) virions from Vero E6 supernatant were probed with  $\alpha$ -Spike and  $\alpha$ -Nucleocapsid (N) antibodies in Western blots. Full-length spike (FL), S1/S2 cleavage product, and S2' cleavage product are indicated.

(B) Densitometry of spike processing from purified virion Western blot in (A) was performed, and quantification of FL and S1/S2 cleavage product percentage of total spike is shown. Quantification was normalized to N as viral protein loading control. WT (black), N679K (green), Omicron (orange). Results are representative of two experiments.

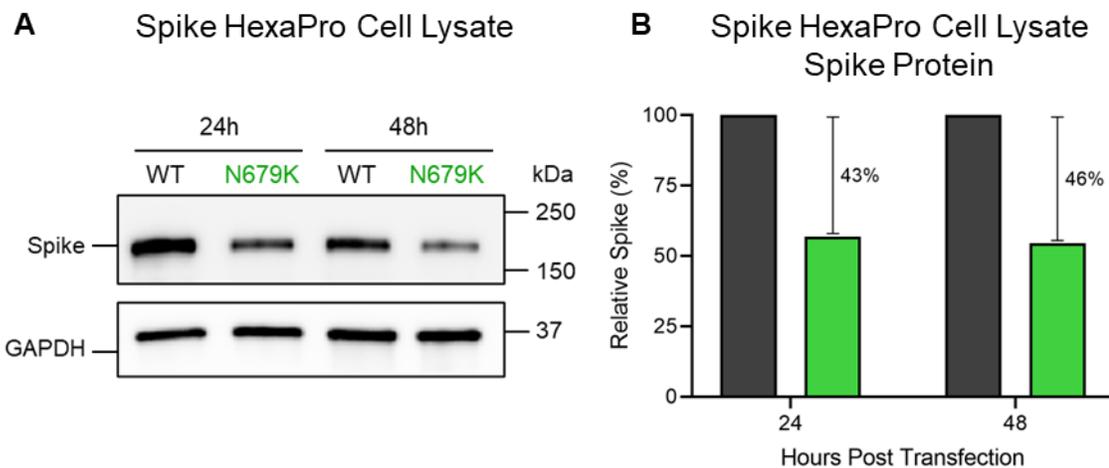
(C) Densitometry of spike expression from purified virion Western blot in (A) was performed, and quantification of total spike protein to nucleocapsid ratio is shown. Spike/N ratio is relative to WT. WT (black), N679K (green), Omicron (orange). Results are representative of two experiments.



**Figure 2.39. Intracellular spike expression of N679K.**

(A) Vero E6 cells were infected with WT, N679K, or Omicron at an MOI of 0.01. Cell lysate was collected at 24 hpi and probed with  $\alpha$ -Spike and  $\alpha$ -Nucleocapsid (N) antibodies in Western blots. Full-length spike (FL), S1/S2 cleavage product, and S2' cleavage product are indicated.

(B) Densitometry of spike expression from infected cell lysate Western blot in (A) was performed, and quantification of total spike protein to nucleocapsid ratio is shown. Spike/N ratio is relative to WT. WT (black), N679K (green), Omicron (orange). Results are representative of three biological replicates.



**Figure 2.40. Expression of exogenous N679K Spike HexaPro.**

(A) Vero E6 cells were transfected with WT and N679K Spike HexaPro and cell lysate was collected at 8, 24, and 48 hpt. Lysates were probed with  $\alpha$ -Spike and  $\alpha$ -GAPDH antibodies in Western blots.

(B) Densitometry of spike expression from transfected cell lysate Western blot in (B) was performed, and quantification of relative total spike protein is shown. Spike protein levels were normalized to GAPDH and are relative to WT. WT (black), N679K (green). Results are representative of three biological replicates.

## **Chapter 4: Discussion and Future Directions**

### **IMPACT OF THE CTS1**

While incompatibilities between the spike RBD and host receptor have been considered the primary barrier to zoonotic emergence, the RBD alone does not accurately predict epidemic strains<sup>6-9</sup>. Spike chimera with SARS-CoV and bat SHC014-CoV revealed that the CTS1 was the driver behind SARS-CoV pathogenesis (data not shown). Although highly conserved in comparison to the NTD and RBD, the differences in the CTS1 can impact its functions in spike processing and host protease usage. CoVs have displayed variability in these processes with different spike cleavage efficiency and requiring different host proteases. Additionally, the barrier to spike cleavage is thought to be higher for bat CoVs due to the protease-rich environment in the gut where bat CoVs are thought to reside. Yet, mutations that enhance spike processing in the less protease-rich respiratory tract are selected for in CoVs that have emerged in the human host. Therefore, the CTS1 contains elements not yet determined that contributes to pathogenesis. Throughout this dissertation, we have demonstrated that the CTS1 contains multiple attributes that alter spike processing, protease usage, and spike stability, highlighting complexities to a domain previously underappreciated.

### **SPIKE CLEAVAGE EFFICIENCY**

The emergence and increased spike cleavability of SARS-CoV-2 compared to previous zoonotic CoVs was credited to the presence of a furin cleavage site, an unusual feature for sarbecoviruses<sup>52,53,68,97</sup>. The FCS allowed the use of furin to cleave the spike intracellularly prior to virion release. The sequence of the FCS is vital for SARS-CoV-2

pathogenesis as observed when the sequence was mutated in PQQA disrupting the polybasic sequence. In addition, this stretch of charged amino acids has been further extended with mutations like P681R in Delta, P681H in Alpha and Omicron, and N679K in Omicron. This introduces the question of if there is an optimal length or sequence to the polybasic sequence that would efficiently promote spike cleavage and not be unstable. Nevertheless, the FCS/polybasic sequence and its iterations remain an integral part to SARS-CoV2 virulence.

We show that another motif QTQTN, also unusual in sarbecoviruses, plays a role in spike processing by introducing additional elements that impact spike cleavage efficiency. Structural modeling predicted that deletion of the QTQTN motif produces a shorter and more rigid loop that the S1/S2 cleavage site is present on. This shortening of the loop may hinder protease interaction as the cleavage site would be closer to the spike core and more inaccessible. The flexibility offered with a longer loop may also allow interaction with an increased range of host proteases that were previously unable to bind at the cleft. A longer loop may be a clear indicator of increased virulence as a strain of hCoV-OC43 with a 4 amino acid insertion adjacent to its S1/S2 cleavage site resulted in a fatal case<sup>83</sup>. Placement of the insertion to extend the loop may not be particularly important as while QTQTN is upstream of the S1/S2 cleavage site, the OC43 insertion is downstream. Similar to the optimal extension of the FCS, it would be interesting to assess if there is an optimal length to the loop before potentially disrupting the more ordered surrounding structure. Also, the addition of two more basic amino acids (N679K and P681H) further extends the FCS/polybasic cleavage sequence. It would be curious to examine if the mutations alter any electrostatic interactions with the surrounding spike, influencing the

position and extension of the loop from the spike body and subsequent interaction with host proteins. Overall, increasing the loop length allows the cleavage site to be more accessible to host proteases and surveillance of such would help in predicting the next epidemic CoV strain.

### **PROTEASE USAGE**

Spike processing and protease usage are typically thought to be intrinsically linked as two concurrent processes. However, our studies have demonstrated the opposite to be true; spike processing and protease usage are not dependent. Mutating the glycosylation sites on the QTQTN motif did not affect spike processing but did alter protease usage. The glycosylation mutants were more affected by the lack of TMPRSS2 compared to  $\Delta$ QTQTN, demonstrating overall inefficient protease usage. It is worth to note that inefficient protease usage does not equate to the inability to use proteases for spike cleavage as the spike is eventually processed. The reduced interaction with one protease may allow for the spike to be processed by a different protease, shifting the entry route used by the virus and infection dynamics. With inefficient TMPRSS2 usage as an example, the viral entry may be delayed and more virus would enter through endosomes, potentially delaying the infection course. Additionally, the mechanism of spread may also be altered, causing the virus to switch between virion release into the supernatant and cell-to-cell spread. Altogether, we have shown that protease usage is distinct and independent of spike processing, subverting what is traditionally thought to occur.

Both  $\Delta$ QTQTN and the glycosylation mutants exhibited differential protease usage with TMPRSS2 usage impacted more so than cathepsin. However, protease interaction measured indirectly by inhibiting either protease using compounds. A biochemical binding

assay such as surface plasmon resonance can more directly assess the effects of loop length and glycosylation on protease interaction. While it seems the inefficient TMPRSS2 usage was due to the lack of glycosylation, it would not be improbable for a shortened loop to alter protease interaction as it could physically shorten the connection to the protease.

In addition to evaluating the effects spike mutations have on protease interaction, it would be interesting to determine whether SARS-CoV-2 preferentially utilizes TMPRSS2 rather than cathepsin due to differences in spike-protease interaction. To that end, Omicron has been observed to prefer the endosomal route and cathepsin instead of the cell surface route and TMPRSS2, representing an additional shift from previous variants<sup>103,104</sup>. It would be interesting to assess whether the Omicron spike has differential interactions to both proteases or if the YKH/N679K mutations are enough to alter protease usage. The ubiquitination on K679 may affect the glycosylation on the adjacent T678, possibly preventing it to become glycosylated and subsequently affecting spike-protease interaction. However, the QTQVN glycosylation mutant demonstrated that a compensatory glycosylation could occur upstream of the motif. The possibility of multiple post translational modifications existing on and surrounding the loop brings forth more complexity to the CTS1 with balancing their interactions.

To be involved in protease interaction, the glycosylation on the QTQTN motif must be interacting with another carbohydrate on the protease. TMPRSS2 contains two glycosylation sites – N213 in the scavenger receptor cysteine-rich (SRCR) domain and N249 in between the SRCR domain and activation domain<sup>116</sup>. It would be interesting if the same inefficient spike-protease interaction as observed with the QTQTN glycosylation mutants is recapitulated in the converse experiment of disrupting the TMPRSS2

glycosylation sites. Follow-up experiments would include infection with WT,  $\Delta$ QTQTN and the QTQTN glycosylation mutants in WT and mutant TMPRSS2-expressing Vero cell line as well as measuring binding affinity between the different spikes and TMPRSS2. Further analysis of possible glycosylation sites on cathepsin would also be interesting if it contributes to protease preference.

### **SPIKE STABILITY**

While we demonstrated that N679K resulted in decreased spike protein expression, the exact mechanism has yet to be determined. With the substitution to a lysine, the most probable mechanism is the introduction of a new ubiquitination site that allows for an increased degradation of the spike protein. Ubiquitination of K679 can be confirmed through mass spectrometry as well as any effect on glycosylation of the motif. Although it may seem unlikely for a single ubiquitination site to have such a profound effect on spike stability, N679K is situated on an exterior loop readily accessible to host proteins. To examine whether N679K indeed promotes spike degradation, a pulse-chase assay comparing WT and N679K Spike HexaPro would be able to measure the rate of degradation for the small amount of labelled spike protein. Albeit unlikely, the decreased spike protein may be a result of decreased spike RNA expression; therefore, comparing the subgenomic RNA levels between WT and N679K is necessary to exclude the possibility.

The implications of reduced spike protein due to an Omicron mutation raise concerns for vaccine efficacy. As Omicron emerged with 30+ spike mutations, it is understandable for all Omicron mutations to be incorporated in the new vaccine design. Most of the Omicron vaccines are mRNA vaccines whose spike protein production elicits immune response. N679K may result in less spike protein to be produced, thus eliciting a

weaker immune response than intended. Furthermore, host ubiquitination-proteasome system (UPS) differs in across cell types, resulting in different spike protein degradation. Additionally, as N679K in the context of all Omicron displayed a similar reduction in spike protein, it would be interesting to generate an Omicron spike with the K679N mutation and assess if spike protein levels would increase, taking care to work in an exogenous expression system to avoid gain-of-function issues.

## **TRANSMISSION**

It is interesting to note that although both  $\Delta$ QTQTN and N679K were attenuated *in vitro* and *in vivo*, their viral replication in the upper airways were altered compared to WT. Whereas  $\Delta$ QTQTN replication in nasal washes was increased at 1, 2 and 4 dpi, N679K was comparable to WT at 2 dpi before dropping off at 4 dpi. As we did not measure viral titers in nasal washes at 1 dpi for N679K, it is possible that we may have missed the window for peak titer and should address this in follow-up experiments. Additionally, with the alterations to upper airway replication, it would of interest to evaluate if transmission dynamics change in terms of virions per droplet or establishing infection in the recipient. For  $\Delta$ QTQTN, transmission may possibly increase as  $\Delta$ QTQTN displayed increased replication. For N679K, despite being a loss of function mutation, transmission may not be greatly attenuated as N679K replicated similarly to WT at 2 dpi within the window for transmission. However, if transmission is lower, the other Omicron mutations may make up the deficit caused by N679K. While it would be interesting to examine the effects on transmission and *in vivo* aspects of other Omicron CTS1 mutation combinations, considerations on possible combinations must take into account gain-of-function potentials.

When examining the relationship between transmissibility and pathogenicity in coronaviruses, a virus typically skews to be more transmissible than pathogenic, like the common cold CoVs, or more pathogenic than transmissible, like SARS-CoV and MERS-CoV. Unlike the previous epidemic SARS-CoV and MERS-CoV, SARS-CoV-2 seems to have established a balance between being capable of spreading and infecting more people without being overly pathogenic. It could be argued that with this balance, SARS-CoV-2 is a more fit virus when compared to SARS-CoV and MERS-CoV in terms of allowing the virus to continue spreading. Through our various studies, we have demonstrated that mutations in the CTS1 can alter the balance between the two, being beneficial in one aspect while disadvantageous in the other. We have shown that deleting the QTQTN motif decreases the pathogenicity of the virus. However,  $\Delta$ QTQTN viral titer was increased in the upper airways, which potentially can increase the transmissibility of the virus. Additionally, subsequent SARS-CoV-2 variants seem to have fine-tuned this balance even more with Omicron shifting its ratio. While Omicron contains mutations that increase spike processing and could therefore increase pathogenicity, the presence of N679K mitigates the virulent effects. The decreased spike protein expression due to N679K aids in tempering the pathogenicity. The reduction could also contribute to immune evasion by having less spike protein be available for immune stimulation and antibody production. The result would therefore allow the virus to better infect and reinfect a host. Overall, the CTS1 can give indications to the transmissibility and pathogenicity of the virus with mutations in the CTS1 potentially skewing the balance.

### **CTS1 IN EMERGING ZOOONOTIC CORONAVIRUSES**

The surveillance of emergent zoonotic coronaviruses has been focused primarily

on receptor compatibility. However, predictions of receptor compatibility may not always be sufficient to predict epidemic strains 9,23. Successful coronavirus infection requires both receptor binding (ruled by the RBD) and spike processing, where the CTS1 is critically involved. The efficiency of spike cleavage at the CTS1 has been previously thought to only be driven by for different proteases. This has led to proposed surveillance focusing on defining new or additional cleavage sites as a metric for emergence. The work done in this dissertation expands upon the complexity of the CTS1 and the aspects that contribute to spike processing and protease usage. The  $\Delta$ QTQTN studies introduced both loop length and glycosylation near the cleavage site as new factors that can alter spike processing and protease interaction. These aspects can serve as additional factors to survey for to more accurately predict epidemic strains. Additionally, the  $\Delta$ QTQTN studies distinguished spike processing and protease usage as independent mechanisms that each have their own role in infection. The Omicron studies reinforce the significance of post translational modifications in the CTS1 region and demonstrate that consequential outcomes can result from a single nucleotide and amino acid change that greatly alter infection dynamics. The Omicron studies also demonstrate the importance of studying mutation constellation. Interplay between the mutations may be functionally distinct compared to when mutations are studied individually. In conclusion, the CTS1 of coronaviruses has been an underappreciated region of the spike protein in terms of its importance to coronavirus emergence and pathogenesis. This body of work introduced new complexities to the CTS1 that may allow for the better surveillance and preparation for the next epidemic CoV strain.

## References

- 1 Belouzard, S., Millet, J. K., Licitra, B. N. & Whittaker, G. R. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **4**, 1011-1033 (2012).
- 2 Magazine, N. *et al.* Mutations and Evolution of the SARS-CoV-2 Spike Protein. *Viruses* **14**, 640 (2022).
- 3 Planas, D. *et al.* Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276-280 (2021). <https://doi.org/10.1038/s41586-021-03777-9>
- 4 Xia, S. *et al.* Structure-based evidence for the enhanced transmissibility of the dominant SARS-CoV-2 B.1.1.7 variant (Alpha). *Cell Discovery* **7**, 109 (2021). <https://doi.org/10.1038/s41421-021-00349-z>
- 5 Gobeil, S. M.-C. *et al.* Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. *Science* **373**, eabi6226 (2021). <https://doi.org/doi:10.1126/science.abi6226>
- 6 Deming, D. *et al.* Vaccine Efficacy in Senescent Mice Challenged with Recombinant SARS-CoV Bearing Epidemic and Zoonotic Spike Variants. *PLOS Medicine* **3**, e525 (2006). <https://doi.org/10.1371/journal.pmed.0030525>
- 7 Sheahan, T., Rockx, B., Donaldson, E., Corti, D. & Baric, R. Pathways of Cross-Species Transmission of Synthetically Reconstructed Zoonotic Severe Acute Respiratory Syndrome Coronavirus. *Journal of Virology* **82**, 8721-8732 (2008). <https://doi.org/doi:10.1128/JVI.00818-08>
- 8 Sheahan, T. *et al.* Mechanisms of Zoonotic Severe Acute Respiratory Syndrome Coronavirus Host Range Expansion in Human Airway Epithelium. *Journal of Virology* **82**, 2274-2285 (2008). <https://doi.org/doi:10.1128/JVI.02041-07>
- 9 Menachery, V. D. *et al.* A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nature Medicine* **21**, 1508-1513 (2015). <https://doi.org/10.1038/nm.3985>
- 10 Roberts, A. *et al.* A Mouse-Adapted SARS-Coronavirus Causes Disease and Mortality in BALB/c Mice. *PLOS Pathogens* **3**, e5 (2007). <https://doi.org/10.1371/journal.ppat.0030005>
- 11 Leist, S. R. *et al.* A Mouse-Adapted SARS-CoV-2 Induces Acute Lung Injury and Mortality in Standard Laboratory Mice. *Cell* **183**, 1070-1085.e1012 (2020). <https://doi.org/10.1016/j.cell.2020.09.050>
- 12 Lyke, K. E. *et al.* Rapid decline in vaccine-boosted neutralizing antibodies against SARS-CoV-2 Omicron variant. *Cell Reports Medicine* **3** (2022). <https://doi.org/10.1016/j.xcrm.2022.100679>
- 13 Planas, D. *et al.* Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671-675 (2022). <https://doi.org/10.1038/s41586-021-04389-z>
- 14 Saso, W. *et al.* Significant role of host sialylated glycans in the infection and spread of severe acute respiratory syndrome coronavirus 2. *PLOS Pathogens* **18**, e1010590 (2022). <https://doi.org/10.1371/journal.ppat.1010590>

- 15 Petitjean, S. J. L. *et al.* Multivalent 9-O-Acetylated-sialic acid glycoclusters as potent inhibitors for SARS-CoV-2 infection. *Nature Communications* **13**, 2564 (2022). <https://doi.org:10.1038/s41467-022-30313-8>
- 16 Tomris, I. *et al.* The SARS-CoV-2 spike N-terminal domain engages 9-  
<em>O</em>-acetylated  $\alpha$ 2-8-linked sialic acids. *bioRxiv*,  
2022.2009.2014.507904 (2022). <https://doi.org:10.1101/2022.09.14.507904>
- 17 Buchanan, C. J. *et al.* Pathogen-sugar interactions revealed by universal saturation transfer analysis. *Science* **377**, eabm3125 (2022).  
<https://doi.org:doi:10.1126/science.abm3125>
- 18 Zehr, J. D. *et al.* Recent Zoonotic Spillover and Tropism Shift of a Canine Coronavirus Is Associated with Relaxed Selection and Putative Loss of Function in NTD Subdomain of Spike Protein. *Viruses* **14**, 853 (2022).
- 19 Graham, C. *et al.* Neutralization potency of monoclonal antibodies recognizing dominant and subdominant epitopes on SARS-CoV-2 Spike is impacted by the B.1.1.7 variant. *Immunity* **54**, 1276-1289.e1276 (2021).  
<https://doi.org:https://doi.org/10.1016/j.immuni.2021.03.023>
- 20 Suryadevara, N. *et al.* Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. *Cell* **184**, 2316-2331.e2315 (2021).  
<https://doi.org:https://doi.org/10.1016/j.cell.2021.03.029>
- 21 McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332-2347.e2316 (2021).  
<https://doi.org:10.1016/j.cell.2021.03.028>
- 22 Meng, B. *et al.* Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Reports* **35** (2021).  
<https://doi.org:10.1016/j.celrep.2021.109292>
- 23 Menachery, V. D. *et al.* SARS-like WIV1-CoV poised for human emergence. *Proceedings of the National Academy of Sciences* **113**, 3048-3053 (2016).  
<https://doi.org:10.1073/pnas.1517719113>
- 24 Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533-534 (2020).  
[https://doi.org:10.1016/s1473-3099\(20\)30120-1](https://doi.org:10.1016/s1473-3099(20)30120-1)
- 25 Gralinski, L. E. & Menachery, V. D. Return of the Coronavirus: 2019-nCoV. *Viruses* **12** (2020). <https://doi.org:10.3390/v12020135>
- 26 Su, S. *et al.* Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends in Microbiology* **24**, 490-502 (2016).  
<https://doi.org:10.1016/j.tim.2016.03.003>
- 27 Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *The Lancet* **391**, 1285-1300 (2018).  
[https://doi.org:10.1016/s0140-6736\(17\)33293-2](https://doi.org:10.1016/s0140-6736(17)33293-2)
- 28 World Health Organization. *Cumulative number of confirmed human cases of avian influenza A(H5N1) reported to WHO*,  
<[https://www.who.int/influenza/human\\_animal\\_interface/2020\\_MAY\\_tableH5N1.pdf?ua=1](https://www.who.int/influenza/human_animal_interface/2020_MAY_tableH5N1.pdf?ua=1)> (2020).

- 29 Food and Agriculture Organization of the United Nations. *H7N9 situation update*, <[http://www.fao.org/ag/againfo/programmes/en/empres/H7N9/situation\\_update.html](http://www.fao.org/ag/againfo/programmes/en/empres/H7N9/situation_update.html)> (
- 30 Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the Mother of All Pandemics. *Emerging Infectious Diseases* **12**, 15-22 (2006). <https://doi.org/10.3201/eid1201.050979>
- 31 Van Kerkhove, M. D., Hirve, S., Koukounari, A. & Mounts, A. W. Estimating age-specific cumulative incidence for the 2009 influenza pandemic: a meta-analysis of A(H1N1)pdm09 serological studies from 19 countries. *Influenza and Other Respiratory Viruses* **7**, 872-886 (2013). <https://doi.org/10.1111/irv.12074>
- 32 Acevedo, M. A., Dillemoth, F. P., Flick, A. J., Faldyn, M. J. & Elder, B. D. Virulence-driven trade-offs in disease transmission: A meta-analysis. *Evolution* **73**, 636-647 (2019). <https://doi.org/10.1111/evo.13692>
- 33 Hector, T. E. & Booksmythe, I. Digest: Little evidence exists for a virulence-transmission trade-off. *Evolution* **73**, 858-859 (2019). <https://doi.org/10.1111/evo.13724>
- 34 Lepiller, Q. *et al.* High incidence but low burden of coronaviruses and preferential associations between respiratory viruses. *Journal of clinical microbiology* **51**, 3039-3046 (2013). <https://doi.org/10.1128/JCM.01078-13>
- 35 Cheng, P. K. *et al.* Viral shedding patterns of coronavirus in patients with probable severe acute respiratory syndrome. *The Lancet* **363**, 1699-1700 (2004). [https://doi.org/10.1016/s0140-6736\(04\)16255-7](https://doi.org/10.1016/s0140-6736(04)16255-7)
- 36 Corman, V. M. *et al.* Viral Shedding and Antibody Response in 37 Patients With Middle East Respiratory Syndrome Coronavirus Infection. *Clin Infect Dis* **62**, 477-483 (2016). <https://doi.org/10.1093/cid/civ951>
- 37 Liu, R. *et al.* Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clin Chim Acta* **505**, 172-175 (2020). <https://doi.org/10.1016/j.cca.2020.03.009>
- 38 De Graaf, M. & Fouchier, R. A. M. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. *The EMBO Journal* **33**, 823-841 (2014). <https://doi.org/10.1002/emboj.201387442>
- 39 Shinya, K. *et al.* Influenza virus receptors in the human airway. *Nature* **440**, 435-436 (2006). <https://doi.org/10.1038/440435a>
- 40 Xiong, X. *et al.* Receptor binding by an H7N9 influenza virus from humans. *Nature* **499**, 496-499 (2013). <https://doi.org/10.1038/nature12372>
- 41 Zeng, H. *et al.* Tropism and Infectivity of a Seasonal A(H1N1) and a Highly Pathogenic Avian A(H5N1) Influenza Virus in Primary Differentiated Ferret Nasal Epithelial Cell Cultures. *Journal of Virology* **93**, e00080-00019 (2019). <https://doi.org/doi:10.1128/JVI.00080-19>
- 42 Richard, M. *et al.* Influenza A viruses are transmitted via the air from the nasal respiratory epithelium of ferrets. *Nature Communications* **11**, 766 (2020). <https://doi.org/10.1038/s41467-020-14626-0>
- 43 Fehr, A. R. & Perlman, S. 1-23 (Springer New York, 2015).
- 44 Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224 (2020). <https://doi.org/10.1038/s41586-020-2179-y>

- 45 Hulswit, R. J. G. *et al.* Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proceedings of the National Academy of Sciences* **116**, 2681-2690 (2019). <https://doi.org:10.1073/pnas.1809667116>
- 46 Mathewson, A. C. *et al.* Interaction of severe acute respiratory syndrome-coronavirus and NL63 coronavirus spike proteins with angiotensin converting enzyme-2. *Journal of General Virology* **89**, 2741-2745 (2008). <https://doi.org:https://doi.org/10.1099/vir.0.2008/003962-0>
- 47 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e286 (2020). <https://doi.org:https://doi.org/10.1016/j.cell.2020.02.058>
- 48 Hou, Y. J. *et al.* SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* **182**, 429-446.e414 (2020). <https://doi.org:10.1016/j.cell.2020.05.042>
- 49 Kawase, M., Shirato, K., van der Hoek, L., Taguchi, F. & Matsuyama, S. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J Virol* **86**, 6537-6545 (2012). <https://doi.org:10.1128/jvi.00094-12>
- 50 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e278 (2020). <https://doi.org:https://doi.org/10.1016/j.cell.2020.02.052>
- 51 Huang, I. C. *et al.* SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. *J Biol Chem* **281**, 3198-3203 (2006). <https://doi.org:10.1074/jbc.M508381200>
- 52 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. *Molecular Cell* **78**, 779-784.e775 (2020). <https://doi.org:https://doi.org/10.1016/j.molcel.2020.04.022>
- 53 Johnson, B. A. *et al.* Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* (2021). <https://doi.org:10.1038/s41586-021-03237-4>
- 54 Böttcher-Friebertshäuser, E., Garten, W., Matrosovich, M. & Klenk, H. D. The hemagglutinin: a determinant of pathogenicity. *Curr Top Microbiol Immunol* **385**, 3-34 (2014). [https://doi.org:10.1007/82\\_2014\\_384](https://doi.org:10.1007/82_2014_384)
- 55 Plante, J. A. *et al.* The variant gambit: COVID-19's next move. *Cell host & microbe* **29**, 508-515 (2021). <https://doi.org:10.1016/j.chom.2021.02.020>
- 56 Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812-827.e819 (2020). <https://doi.org:10.1016/j.cell.2020.06.043>
- 57 Yurkovetskiy, L. *et al.* Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **183**, 739-751.e738 (2020). <https://doi.org:10.1016/j.cell.2020.09.032>
- 58 Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* (2020). <https://doi.org:10.1038/s41586-020-2895-3>

- 59 Hou, Y. J. *et al.* SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **370**, 1464-1468 (2020).  
<https://doi.org/10.1126/science.abe8499>
- 60 Zhang, L. *et al.* SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nature Communications* **11**, 6013 (2020).  
<https://doi.org/10.1038/s41467-020-19808-4>
- 61 *Confirmed Cases of COVID-19 Variants Identified in UK*, <[www.gov.uk](http://www.gov.uk)> (2021).
- 62 Goncalves Cabecinhas, A. R. *et al.* SARS-CoV-2 N501Y Introductions and Transmissions in Switzerland from Beginning of October 2020 to February 2021- Implementation of Swiss-Wide Diagnostic Screening and Whole Genome Sequencing. *Microorganisms* **9** (2021).  
<https://doi.org/10.3390/microorganisms9040677>
- 63 Liu, Y. *et al.* The N501Y spike substitution enhances SARS-CoV-2 transmission. *bioRxiv* (2021). <https://doi.org/10.1101/2021.03.08.434499>
- 64 Peacock, T. P. *et al.* The SARS-CoV-2 variants associated with infections in India, B.1.617, show enhanced spike cleavage by furin. *bioRxiv*, 2021.2005.2028.446163 (2021). <https://doi.org/10.1101/2021.05.28.446163>
- 65 Vu, M. N. & Menachery, V. D. Binding and entering: COVID finds a new home. *PLoS Pathogens* **17**, e1009857 (2021).  
<https://doi.org/10.1371/journal.ppat.1009857>
- 66 Vu, M. N. & Menachery, V. D. Binding and entering: COVID finds a new home. *PLoS Pathog* **17**, e1009857 (2021). <https://doi.org/10.1371/journal.ppat.1009857>
- 67 Coutard, B. *et al.* The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* **176**, 104742 (2020). <https://doi.org/10.1016/j.antiviral.2020.104742>
- 68 Peacock, T. P. *et al.* The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat Microbiol* **6**, 899-909 (2021).  
<https://doi.org/10.1038/s41564-021-00908-w>
- 69 Liu, Z. *et al.* Identification of Common Deletions in the Spike Protein of Severe Acute Respiratory Syndrome Coronavirus 2. *J Virol* **94** (2020).  
<https://doi.org/10.1128/JVI.00790-20>
- 70 Lau, S.-Y. *et al.* Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerging Microbes & Infections* **9**, 837-842 (2020).  
<https://doi.org/10.1080/22221751.2020.1756700>
- 71 Ogando, N. S. *et al.* SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. *Journal of General Virology* **101**, 925-940 (2020). <https://doi.org/10.1099/jgv.0.001453>
- 72 Zou, W. *et al.* The SARS-CoV-2 Transcriptome and the Dynamics of the S Gene Furin Cleavage Site in Primary Human Airway Epithelia. *mBio* **12**, e01006-01021 (2021). <https://doi.org/doi:10.1128/mBio.01006-21>
- 73 Jaworski, E. *et al.* Tiled-ClickSeq for targeted sequencing of complete coronavirus genomes with simultaneous capture of RNA recombination and minority variants. *Elife* **10** (2021). <https://doi.org/10.7554/eLife.68479>

- 74 Colson, P. *et al.* Occurrence of a substitution or deletion of SARS-CoV-2 spike amino acid 677 in various lineages in Marseille, France. *Virus Genes* **58**, 53-58 (2022). <https://doi.org:10.1007/s11262-021-01877-2>
- 75 Liu, L.-T. *et al.* Isolation and Identification of a Rare Spike Gene Double-Deletion SARS-CoV-2 Variant From the Patient With High Cycle Threshold Value. *Frontiers in Medicine* **8** (2022). <https://doi.org:10.3389/fmed.2021.822633>
- 76 Xie, X. *et al.* Engineering SARS-CoV-2 using a reverse genetic system. *Nat Protoc* **16**, 1761-1784 (2021). <https://doi.org:10.1038/s41596-021-00491-8>
- 77 Xie, X. *et al.* An Infectious cDNA Clone of SARS-CoV-2. *Cell Host Microbe* **27**, 841-848 e843 (2020). <https://doi.org:10.1016/j.chom.2020.04.004>
- 78 Imai, M. *et al.* Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 16587-16595 (2020). <https://doi.org:10.1073/pnas.2009799117>
- 79 Sanda, M., Morrison, L. & Goldman, R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. *Anal Chem* **93**, 2003-2009 (2021). <https://doi.org:10.1021/acs.analchem.0c03173>
- 80 Gao, C. *et al.* SARS-CoV-2 Spike Protein Interacts with Multiple Innate Immune Receptors. *bioRxiv*, 2020.2007.2029.227462 (2020). <https://doi.org:10.1101/2020.07.29.227462>
- 81 Holmes, E. C. *et al.* The origins of SARS-CoV-2: A critical review. *Cell* **184**, 4848-4856 (2021). <https://doi.org:10.1016/j.cell.2021.08.017>
- 82 Laporte, M. *et al.* The SARS-CoV-2 and other human coronavirus spike proteins are fine-tuned towards temperature and proteases of the human airways. *PLOS Pathogens* **17**, e1009500 (2021). <https://doi.org:10.1371/journal.ppat.1009500>
- 83 Lau, S. K. P. *et al.* Fatal Pneumonia Associated With a Novel Genotype of Human Coronavirus OC43. *Frontiers in Microbiology* **12** (2022). <https://doi.org:10.3389/fmicb.2021.795449>
- 84 Harcourt, J. *et al.* Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with Coronavirus Disease, United States. *Emerging infectious diseases* **26**, 1266-1273 (2020). <https://doi.org:10.3201/eid2606.200516>
- 85 Josset, L. *et al.* Cell Host Response to Infection with Novel Human Coronavirus EMC Predicts Potential Antivirals and Important Differences with SARS Coronavirus. *mBio* **4**, e00165-00113 (2013). <https://doi.org:doi:10.1128/mBio.00165-13>
- 86 Routh, A., Head, S. R., Ordoukhanian, P. & Johnson, J. E. ClickSeq: Fragmentation-Free Next-Generation Sequencing via Click Ligation of Adaptors to Stochastically Terminated 3'-Azido cDNAs. *Journal of Molecular Biology* **427**, 2610-2616 (2015). <https://doi.org:https://doi.org/10.1016/j.jmb.2015.06.011>
- 87 Menachery, V. D. *et al.* Attenuation and Restoration of Severe Acute Respiratory Syndrome Coronavirus Mutant Lacking 2'&#x2032;-O-Methyltransferase Activity. *Journal of Virology* **88**, 4251-4264 (2014). <https://doi.org:doi:10.1128/JVI.03571-13>
- 88 Routh, A. *et al.* Poly(A)-ClickSeq: click-chemistry for next-generation 3'-end sequencing without RNA enrichment or fragmentation. *Nucleic Acids Res* **45**, e112-e112 (2017). <https://doi.org:10.1093/nar/gkx286>

- 89 Elrod, N. D., Jaworski, E. A., Ji, P., Wagner, E. J. & Routh, A. Development of Poly(A)-ClickSeq as a tool enabling simultaneous genome-wide poly(A)-site identification and differential expression analysis. *Methods* **155**, 20-29 (2019). <https://doi.org/10.1016/j.ymeth.2019.01.002>
- 90 Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res* **12**, 996-1006 (2002). <https://doi.org/10.1101/gr.229102>
- 91 Su, Z. *et al.* Annexin A2 depletion exacerbates the intracerebral microhemorrhage induced by acute rickettsia and Ebola virus infections. *PLOS Neglected Tropical Diseases* **14**, e0007960 (2020). <https://doi.org/10.1371/journal.pntd.0007960>
- 92 Bern, M., Kil, Y. J. & Becker, C. Byonic: Advanced Peptide and Protein Identification Software. *Current Protocols in Bioinformatics* **40**, 13.20.11-13.20.14 (2012). <https://doi.org/10.1002/0471250953.bi1320s40>
- 93 Lin, Y.-H., Franc, V. & Heck, A. J. R. Similar Albeit Not the Same: In-Depth Analysis of Proteoforms of Human Serum, Bovine Serum, and Recombinant Human Fetuin. *Journal of Proteome Research* **17**, 2861-2869 (2018). <https://doi.org/10.1021/acs.jproteome.8b00318>
- 94 Tian, Y., Parsons, L. M., Jankowska, E. & Cipollo, J. F. Site-Specific Glycosylation Patterns of the SARS-CoV-2 Spike Protein Derived From Recombinant Protein and Viral WA1 and D614G Strains. *Frontiers in Chemistry* **9** (2021). <https://doi.org/10.3389/fchem.2021.767448>
- 95 Vu, M. N. *et al.* QTQTN motif upstream of the furin-cleavage site plays a key role in SARS-CoV-2 infection and pathogenesis. *Proceedings of the National Academy of Sciences* **119**, e2205690119 (2022). <https://doi.org/10.1073/pnas.2205690119>
- 96 Callaway, E. Heavily mutated Omicron variant puts scientists on alert. *Nature* **600**, 21 (2021). <https://doi.org/10.1038/d41586-021-03552-w>
- 97 Coutard, B. *et al.* The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Research* **176**, 104742 (2020). <https://doi.org/10.1016/j.antiviral.2020.104742>
- 98 Liu, Y. *et al.* Delta spike P681R mutation enhances SARS-CoV-2 fitness over Alpha variant. *bioRxiv*, 2021.2008.2012.456173 (2021). <https://doi.org/10.1101/2021.08.12.456173>
- 99 Escalera, A. *et al.* Mutations in SARS-CoV-2 variants of concern link to increased spike cleavage and virus transmission. *Cell Host & Microbe* **30**, 373-387.e377 (2022). <https://doi.org/10.1016/j.chom.2022.01.006>
- 100 Liu, Y. *et al.* Delta spike P681R mutation enhances SARS-CoV-2 fitness over Alpha variant. *Cell Reports* **39** (2022). <https://doi.org/10.1016/j.celrep.2022.110829>
- 101 Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121-4123 (2018). <https://doi.org/10.1093/bioinformatics/bty407>
- 102 Xie, X. *et al.* Engineering SARS-CoV-2 using a reverse genetic system. *Nature Protocols* **16**, 1761-1784 (2021). <https://doi.org/10.1038/s41596-021-00491-8>

- 103 Peacock, T. P. *et al.* The altered entry pathway and antigenic distance of the SARS-CoV-2 Omicron variant map to separate domains of spike protein. *bioRxiv*, 2021.2012.2031.474653 (2022). <https://doi.org/10.1101/2021.12.31.474653>
- 104 Meng, B. *et al.* Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. *Nature* **603**, 706-714 (2022). <https://doi.org/10.1038/s41586-022-04474-x>
- 105 Phadke, K. S., Higdon, N. B. A. & Bellaire, B. H. *In vitro* comparison of SARS-CoV-2 variants. *bioRxiv*, 2023.2003.2011.532212 (2023). <https://doi.org/10.1101/2023.03.11.532212>
- 106 Zhao, H. *et al.* SARS-CoV-2 Omicron variant shows less efficient replication and fusion activity when compared with Delta variant in TMPRSS2-expressed cells. *Emerging Microbes & Infections* **11**, 277-283 (2022). <https://doi.org/10.1080/22221751.2021.2023329>
- 107 Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501-1505 (2020). <https://doi.org/doi:10.1126/science.abd0826>
- 108 Mustafa, Z., Kalbacher, H. & Burster, T. Occurrence of a novel cleavage site for cathepsin G adjacent to the polybasic sequence within the proteolytically sensitive activation loop of the SARS-CoV-2 Omicron variant: The amino acid substitution N679K and P681H of the spike protein. *PLOS ONE* **17**, e0264723 (2022). <https://doi.org/10.1371/journal.pone.0264723>
- 109 Yuan, S. *et al.* Pathogenicity, transmissibility, and fitness of SARS-CoV-2 Omicron in Syrian hamsters. *Science* **377**, 428-433 (2022). <https://doi.org/doi:10.1126/science.abn8939>
- 110 McCormack, C. P. *et al.* Modelling the viral dynamics of the SARS-CoV-2 Delta and Omicron variants in different cell types. *bioRxiv*, 2023.2003.2015.529513 (2023). <https://doi.org/10.1101/2023.03.15.529513>
- 111 Hui, K. P. Y. *et al.* SARS-CoV-2 Omicron variant replication in human bronchus and lung ex vivo. *Nature* **603**, 715-720 (2022). <https://doi.org/10.1038/s41586-022-04479-6>
- 112 Carabelli, A. M. *et al.* SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nature Reviews Microbiology* **21**, 162-177 (2023). <https://doi.org/10.1038/s41579-022-00841-7>
- 113 Papanikolaou, V. *et al.* From delta to Omicron: S1-RBD/S2 mutation/deletion equilibrium in SARS-CoV-2 defined variants. *Gene* **814**, 146134 (2022). <https://doi.org/https://doi.org/10.1016/j.gene.2021.146134>
- 114 Shrestha, L. B., Foster, C., Rawlinson, W., Tedla, N. & Bull, R. A. Evolution of the SARS-CoV-2 omicron variants BA.1 to BA.5: Implications for immune escape and transmission. *Reviews in Medical Virology* **32**, e2381 (2022). <https://doi.org/https://doi.org/10.1002/rmv.2381>
- 115 Johnson, B. A. *et al.* Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis. *PLOS Pathogens* **18**, e1010627 (2022). <https://doi.org/10.1371/journal.ppat.1010627>
- 116 Chen, Y.-W. *et al.* TMPRSS2, a Serine Protease Expressed in the Prostate on the Apical Surface of Luminal Epithelial Cells and Released into Semen in

Prostasomes, Is Misregulated in Prostate Cancer Cells. *The American Journal of Pathology* **176**, 2986-2996 (2010). <https://doi.org/10.2353/ajpath.2010.090665>

## VITA

Michelle Vu was born on November 26, 1995 in Pomona, California to Thomas and Tina Vu and the younger sister of Bryant Vu. She married to Charles Luu in December 2022. She was naturally curious about how the world functioned since she was a child and sought to understand how biological functions worked. Michelle graduated from Fountain Valley high school in 2013 and entered the University of California – Irvine under the Biological Sciences program. There she learned more about the wonderful world of viruses and decided to pursue virology as a career. She joined Dr. Bert Semler’s group studying picornavirus-host interactions as an undergraduate researcher. She continued on in the lab as a research assistant upon graduation to build a strong foundation in molecular virology. Michelle began her doctorate training in the Microbiology and Immunology graduate program at the University of Texas Medical Branch at Galveston. She worked under the mentorship of Dr. Vineet Menachery studying the importance of the coronavirus spike C-terminal of the S1 subunit to pathogenicity. Michelle plans to pursue a career in research dissecting the interactions between viruses and hosts.

## EDUCATION

B.S., May 2017, University of California – Irvine, Irvine, California

## PUBLICATIONS

1. Schindewolf C, Lokugamage K, **Vu MN**, Johnson BA, Scharton D, Plante JA, et al. SARS-CoV-2 Uses Nonstructural Protein 16 To Evade Restriction by IFIT1 and IFIT3. *Journal of Virology*. 2023;97(2):e01532-22. doi: doi:10.1128/jvi.01532-22.
2. Bouin A, **Vu MN**, Al-Hakeem A, Tran GP, Nguyen JHC, Semler BL. Enterovirus-Cardiomyocyte Interactions: Impact of Terminally Deleted Genomic RNAs on Viral and Host Functions. *Journal of Virology*. 0(0):e01426-22. doi: doi:10.1128/jvi.01426-22.
3. **Vu MN**, Lokugamage KG, Plante JA, Scharton D, Bailey AO, Sotcheff S, Swetnam DM, Johnson BA, Schindewolf C, Alvarado RE, Crocquet-Valdes PA, Debbink K, Weaver SC, Walker DH, Russell WK, Routh AL, Plante KS, Menachery VD. QTQTN motif upstream of the furin-cleavage site plays a key role in SARS-CoV-2 infection and pathogenesis. *Proceedings of the National Academy of Sciences*. 2022;119(32):e2205690119. doi: doi:10.1073/pnas.2205690119.
4. Johnson BA, Zhou Y, Lokugamage KG, **Vu MN**, Bopp N, Crocquet-Valdes PA, Kalveram B, Schindewolf C, Liu Y, Scharton D, Plante JA, Xie X, Aguilar P, Weaver SC, Shi P-Y, Walker DH, Routh AL, Plante KS, Menachery VD. Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis. *PLOS Pathogens*. 2022;18(6):e1010627. doi: 10.1371/journal.ppat.1010627.

5. **Vu, MN** and Menachery, VD. Binding and entering: COVID finds a new home. *PLOS Pathogens* 17, e1009857, doi:10.1371/journal.ppat.1009857 (2021).
6. Johnson, K., **Vu, M.** & Freiberg, A. N. Recent advances in combating Nipah virus. *Fac Rev* 10, 74, doi:10.12703/r/10-74 (2021).
7. Muruato A, **Vu MN**, Johnson BA, Davis-Gardner ME, Vanderheiden A, Lokugamage K, Schindewolf C, Crocquet-Valdes PA, Langsjoen RM, Plante JA, Plante KS, Weaver SC, Debbink K, Routh AL, Walker D, Suthar MS, Shi P-Y, Xie X, Menachery VD. Mouse-adapted SARS-CoV-2 protects animals from lethal SARS-CoV challenge. *PLOS Biology*. 2021;19(11):e3001284. doi: 10.1371/journal.pbio.3001284.
8. Matchett WE, Joag V, Stolley JM, Shepherd FK, Quarnstrom CF, Mickelson CK, Wijeyesinghe S, Soerens AG, Becker S, Thiede JM, Weyu E, O'Flanagan SD, Walter JA, **Vu MN**, Menachery VD, Bold TD, Vezys V, Jenkins MK, Langlois RA, Masopust D. Cutting Edge: Nucleocapsid Vaccine Elicits Spike-Independent SARS-CoV-2 Protective Immunity. *J Immunol*. 2021 Jul 15;207(2):376-379. doi: 10.4049/jimmunol.2100421. Epub 2021 Jun 30. PMID: 34193597; PMCID: PMC8516699.
9. Xie X, Lokugamage KG, Zhang X, **Vu MN**, Muruato AE, Menachery VD, Shi P-Y. Engineering SARS-CoV-2 using a reverse genetic system. *Nature Protocols*. 2021;16(3):1761-84. doi: 10.1038/s41596-021-00491-8.
10. Johnson BA, Xie X, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich T, Smith JK, Zhang L, Bopp N, Schindewolf C, **Vu M**, Vanderheiden A, Winkler ES, Swetnam D, Plante JA, Aguilar P, Plante KS, Popov V, Lee B, Weaver SC, Suthar MS, Routh AL, Ren P, Ku Z, An Z, Debbink K, Diamond MS, Shi PY, Freiberg AN, Menachery VD. Loss of furin cleavage site attenuates SARS-CoV-2

- pathogenesis. *Nature*. 2021. Epub 2021/01/26. doi: 10.1038/s41586-021-03237-4. PubMed PMID: 33494095.
11. Sandler ZJ, Firpo MR, Omoba OS, **Vu MN**, Menachery VD, Mounce BC. Novel Ionophores Active against La Crosse Virus Identified through Rapid Antiviral Screening. *Antimicrob Agents Chemother*. 2020;64(6). Epub 2020/04/15. doi: 10.1128/aac.00086-20. PubMed PMID: 32284379; PMCID: PMC7269492.
  12. Bouin A, Gretteau PA, Wehbe M, Renois F, N'Guyen Y, Lévêque N, **Vu MN**, Tracy S, Chapman NM, Bruneval P, Fornes P, Semler BL, Andreoletti L. Enterovirus Persistence in Cardiac Cells of Patients With Idiopathic Dilated Cardiomyopathy Is Linked to 5' Terminal Genomic RNA-Deleted Viral Populations With Viral-Encoded Proteinase Activities. *Circulation*. 2019;139(20):2326-38. Epub 2019/02/14. doi: 10.1161/circulationaha.118.035966. PubMed PMID: 30755025; PMCID: PMC6517084.

Permanent address: 11142 Lavender Ave, Fountain Valley, CA 92708

This dissertation was typed by Michelle N. Vu