SARS-CoV-2 Variant of Concern Substitutions Alter Spike Glycoprotein Receptor Binding Domain Structure and Stability

Daniel L. Moss[1*] and Jay Rappaport[1,2]

[1]Divison of Comparative Pathology, Tulane National Primate Research Center, Covington, LA, USA

[2]Department of Microbiology and Immunology, Tulane University School of Medicine, New Orleans, LA, USA.

*Author to whom correspondence should be addressed:

Daniel L. Moss

Division of Comparative Pathology, Tulane National Primate Research Center,

18703 Three Rivers Rd., Covington, LA, 70433-8915 USA

dmoss2@tulane.edu

Abstract

The emergence of severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) and the subsequent COVID-19 pandemic has significantly impacted the world not just with disease and death but also economic turmoil. The rapid development and deployment of extremely effective vaccines against SARS-CoV-2 has made the end of the pandemic a reality within reach. However, as the virus spreads it has acquired mutations; and thus, variants of concern have emerged that are more infectious and reduce the efficacy of existing vaccines. While promising efforts are underway to combat these variants, the evolutionary pressures leading to these variants are poorly understood. To that end, here we have studied the effects of three amino-acid substitutions on the structure and function of the SARS-CoV-2 spike glycoprotein receptor-binding domain found in several variants of concern such as B.1.1.7, B.1.351 and P.1 that are now circulating. We found that these substitutions alter the RBD structure and stability, as well as its ability to bind to ACE2, which may have opposing and compensatory effects. These findings provide new insights into how these Variants of Concern (VOC) may have been selected to optimize infectivity while maintaining the structure and stability of the receptor binding domain.

Introduction

The emergence of SARS-CoV-2 in late 2019 and its subsequent spread around the world has caused the deadliest pandemic since the 1918 influenza pandemic nearly a century ago[1]. The international scientific community has risen to the challenge of combating SARS-CoV-2 and COVID-19. The year 2020 ended with the fastest development of vaccine candidates, starting with the genetic sequence of the virus being reported[2] to human trials of novel mRNA based vaccines within three months. Now there are three SARS-CoV-2 vaccines approved for emergency use within the United States and many more next generation and pan-coronavirus vaccines currently in development. These advances have made substantial contributions to the control of the COVID-19 pandemic within the United States. Despite multiple manufacturers receiving emergency use authorization, and an unprecedented vaccination campaign, significant challenges remain including uncertainty regarding durability, vaccination hesitancy, limited access to healthcare among disadvantaged persons, as well as the continued emergence of VOC. Our ultimate success in quelling this pandemic may ultimately lie in our ability, not only to characterize new variants, but to be able to predict the emergence of new variants. Suh advances will require an increased understanding of evolutionary pressures and constraints on viral variation.

Three SARS-CoV-2 lineages B.1.1.7 (or 501Y.V1) first identified within the United Kingdom, B.1.351 (or 501Y.V2) identified in South Africa, and P.1 (or 501Y.V3) identified in Brazil, have been demonstrated to possess increased infectivity[3] and in the case of B.1.351 and P.1 exhibit reduced neutralization by antibodies reacting with the cognate regions of the spike protein within the original Wuhan strain of SARS-CoV-2[4–6]. B.1.1.7 possesses the N501Y substitution within the spike glycoprotein receptor binding domain (RBD) which has been shown to enhance binding to ACE2, the entry receptor for SARS-CoV-2[7–9]. B.1.351 and P.1 possess N501Y as well as substitutions at two other sites within the RBD, E484K and K417N in B.1.351 and K417T in P.1[10]. These RBD substitutions present in the spike protein of the B.1.351 and P.1 variants have been shown to reduce the binding and neutralization of mRNA vaccine induced antibodies as well as potent human monoclonal antibodies[11].

The consequences of the K417N, E484K and N501Y substitutions on RBD-ACE2 interactions have also been examined, with the increased infectivity of the B.1.1.7 variant resulting from

enhanced binding to ACE2 when the RBD N501Y substitution is present[9]. The E484K substitution has been shown to enhance ACE2 binding[12] and reduce the efficacy of neutralizing antibodies[13]. A recent study examined the effects of the K417N substitution on ACE2 binding and antibody interactions using molecular dynamics and found that K417N disrupts RBD-ACE2 interactions, as well as interactions with a monoclonal antibody[14]. However, the effects of these substitutions on the structure of the RBD itself have not been examined. Based on the nature of these substitutions, including residue changes in charge or polar to non-polar substitutions, we hypothesized that the K417N, E484K and N501Y substitutions alter RBD structure and stability as well as ACE2 binding interactions. We studied those changes in single-substitution RBD variants as well as in the RBD containing all three substitutions using molecular dynamics and biophysical approaches. Our data suggest that these VOC substitutions significantly alter RBD structure and stability, with consequences for ACE2 binding and proteolytic susceptibility, having potentially opposing consequences for the evolutionary success of new variants. These findings have implications for viral evolution as well as the design of subunit vaccine candidates.

Results

RBD Variant of Concern Substitutions Alter RBD Structure *in silico*

We began our studies on RBD variant of concern substitutions with molecular dynamics to investigate whether single substitutions or all three substitutions found within the B.1.351 lineage would alter RBD structure. We used homology modeling of the RBD, spike glycoprotein residues 319 – 541, to generate structures for analysis by molecular dynamics simulation using the whole spike glycoprotein as a template (PDB 6VXX)[15,16]. The model we generated of the wild-type/Wuhan strain RBD was used as a template for further modeling of RBD variants with single substitutions as well as all three substitutions (K417N/E484K/N501Y) found within the B.1.351 lineage. Molecular dynamic simulations were performed with the GROMACS 2020.5 package [17] using the CHARMM36 force field[18]. After solvation and neutralization, systems were minimized and equilibrated before undergoing a 25-nanosecond NPT equilibration run, followed by three separate 10-nanosecond production runs. From these trajectories we observed no large differences in root mean square deviation (RMSD) of the entire protein structure between RBD variants (Figure 1A). K417N and K417N/E484K/N501Y exhibited initial jumps in RMSD but these converged with the other RBD variants after several nanoseconds had elapsed. Next, we calculated the radius of gyration, a measure of the compactness of the molecule, at each frame of the 10 ns production run. We observed an increase in the radius of gyration for the N501Y and K417N/E484K/N501Y RBD variants when compared to wild-type, suggesting these substitutions result in a less compact structure, while K417N resulted in a marked decrease in the radius of gyration, suggesting a more compact protein structure (Figure 1B). These changes in radius of gyration were accompanied by changes in the overall hydrogen bond content of RBD variants. The K417N substitution resulted in a noticeable increase in the overall number of hydrogen bonds within the molecule, which became more apparent after about 5 ns into the simulations, while the E484K and the K417N/E484K/N501Y variants exhibited a decrease in hydrogen bond content (Figure 1C). Overall, these RBD substitutions had a minimal effect on residue specific fluctuations with the exception of a region between residues 468 and 488 at the ACE2 binding interface (Figure 1D and 2A). The N501Y substitution appears to have altered the flexibility of this region although these differences were not significant when compared to the wild-type RBD. However, the K417N substitution significantly reduced the flexibility of this region when compared to all other variants, suggesting potential consequences for ACE2 binding.

The 468 – 488 loop of the RBD comprises part of the ACE2 binding interface and contains the E484K substitution found in the P.1 and B.1.351 lineages. This region of the RBD undergoes a conformational change when bound to ACE2 to form contacts with ACE2 residues (Figure 2B). Our simulation results suggest that the K417N substitution significantly alters the flexibility of the 468-488 region as well as the overall structure of the RBD alone by altering hydrogen bond content (Figure 1C). Frames extracted from the trajectories of the wild-type and K417N simulations show the 468 – 488 loop in different orientations for the two RBD variants (Figure 2C). Closer examination of the simulation coordinates shows E484 in close proximity to R403 in the K417N variant while in the wild-type RBD E484 is more than 20 angstroms away (Figure 2D). The proximity of E484 and R403 in K417N suggest the formation of a hydrogen bond interaction between these two residues that is not present in the wild-type RBD trajectory. The distances between these two residues are more than doubled in the trajectories generated from simulation of the wild-type and E484K RBD structures (Figure 2E). These sidechain interactions are not present in any of the wild-type or E484K RBD trajectories and may be responsible for the differences in the flexibility of the 468 – 488 loop we have observed here. Analysis of the RMSD for the 468 – 488 loop from the wild-type, K417N and E484K trajectories suggests that the E484-R403 interaction significantly reduces the flexibility of this loop region (Figure 2F). The 468 – 488 loop remains relatively static throughout the trajectory in the K417N variant while in the wild-type and E484K RBD this loop region samples more orientations (Figure 2F). This hydrogen bond interaction between R403 and E484 may disfavor the necessary conformational change in this region that facilitates ACE2 binding.

RBD Variant of Concern Substitutions Alter RBD Stability and ACE2 Binding Affinity

Molecular dynamic simulations of RBD variants suggested that variant of concern substitutions alter RBD structure and hydrogen bonding. Therefore, we hypothesized that these RBD substitutions alter RBD stability and resistance to unfolding. To test this we measured guanidine-induced unfolding of RBD variants using a fluorescence based unfolding assay[19,20]. Unfolding curves generated by measuring fluorescent dye binding in response to increasing guanidine-HCl concentration were fit by non-linear regression to estimate free energy of unfolding[21]. We observed that the E484K substitution significantly destabilized the RBD while the N501Y and K417N substitutions significantly stabilized the RBD relative to wild-type (Figure 3A and B). We observed no difference in unfolding energy when all three variant of concern substitutions were present in the RBD compared to the wild-type protein.

As discussed above, it has been reported that the substitutions to the spike protein observed in the B.1.1.7 and B.1.351 lineages exhibit higher binding affinity for ACE2. We sought to replicate these observations and examine the effects of the individual K417N and E484K substitutions on ACE2 binding affinity. It has also been reported that K417N reduces RBD binding affinity for ACE2 by both computational and experimental investigation[14,22]. To test the effects of RBD substitutions on ACE2 binding we utilized a binding competition assay where RBD variants compete with horseradish peroxidase-labeled wild-type RBD (HRP-RBD) for binding to ACE2 adsorbed to a microplate. As RBD variants are diluted HRP-RBD can outcompete for binding to ACE2 resulting in increased absorbance at 450 nm after plate development (Figure 4A). These data were analyzed by non-linear regression to calculate LogIC50 values. We found that the K417N substitution significantly reduced ACE2 binding as less dilution was needed for HRP-RBD to outcompete for ACE2 binding (Figure 4B). E484K exhibited a slightly higher affinity for ACE2 when compared to the wild-type RBD. The N501Y substitution produced a marked increase in ACE2 affinity consistent with previous reports and the presence of all three variant of concern

substitutions exhibited the greatest increase in ACE2 affinity compared to wild-type RBD and bound ACE2 more strongly than the N501Y substitution alone (Figure 4B). These results were consistent with previously reported observations of the effects of variant of concern substitutions on RBD-ACE2 binding[8,9,14,22].

RBD Variant of Concern Substitutions Alter RBD Proteolytic Susceptibility

Molecular dynamic simulations and results from unfolding studies presented here lead us to the conclusion that RBD variant of concern substitutions alter RBD structure and stability. Changes in protein structure and stability are associated with changes in proteolytic resistance[19,20,23,24]. Based on these previous observations and those we have reported here thus far, we hypothesized that RBD substitutions alter proteolytic susceptibility in accordance with changes in stability. Substitutions that stabilize the RBD increase proteolytic resistance, while destabilizing substitutions will decrease proteolytic resistance. To test this, we performed limited proteolysis experiments using the lysosomal protease cathepsin S at pH 5.6. Proteolysis reactions were sampled after 0, 15, 30 and 60 minutes of incubation and analyzed by SDS-PAGE and Coomassie staining (Figure 5A). Gels were imaged and intensities of the RBD bands were measured with gel analysis software. Band intensities were normalized to the 0 minutes timepoint and compared by two-way ANOVA. We observed that the K417N substitution significantly increased proteolytic resistance at all time points compared to the wild-type RBD (Figure 5B). The E484K substitution significantly decreased proteolytic resistance compared to the wild-type RBD except after 60 minutes of incubation. The N501Y substitution exhibited increased resistance to proteolysis by cathepsin S only after 60 minutes of incubation relative to the wild-type RBD. The presence of all three substitutions resulted in a large increase in cathepsin S proteolytic susceptibility after 15 and 30 minutes but degradation was similar to that seen for wild-type RBD after 60 minutes (Figure 5B). Taken together these results demonstrate that RBD variant of concern substitutions significantly alter RBD proteolytic susceptibility.

Discussion

In this study we tested the hypothesis that SARS-CoV-2 variant of concern substitutions within the spike glycoprotein RBD alter the RBD structure, stability and ACE2 binding affinity. We studied the RBD by molecular dynamics simulation and found that the K417N substitution alone may make the RBD more compact, as measured by a decrease in the radius of gyration and increase its hydrogen bond content and alter the flexibility of a loop spanning residues 468 – 488 that contacts the viral entry receptor ACE2. The presence of all three substitutions (K417N/E484K/N501Y), found in the B.1.351 variant, appeared to have the opposite effect as K417N alone with an increase in the radius of gyration and a decrease in hydrogen bond content making the molecule less compact and more flexible, while leaving the ACE2 binding loop contained in residues 468 – 488 unchanged relative to the wild-type RBD. A similar molecular dynamic simulation analysis of this region was reported for the wild-type RBD and the E484K, N501Y and the K417N/E484K/N501Y substitutions, albeit with nearly opposite results to those reported here[25]. The authors observed that the E484K and N501Y substitutions alone increased flexibility relative to wild-type while both E484K and N501Y together along with all three substitutions (K417N/E484K/N501Y) did not alter flexibility compared to the wild-type RBD. These differences in computed RMSF values may be due to the use of different forcefields for molecular dynamic as well as equilibration conditions.

In a binding competition assay, the N501Y and K417N/E484K/N501Y variants all exhibited higher binding affinity for ACE2 while the K417N substitution alone exhibited reduced binding affinity.

The E484K substitution caused a modest but significant increase in affinity as measured by the competition assay employed here but other studies have reported the E484K substitution results in a negligible increase in affinity[26]. Nonetheless our observations are all in agreement with numerous previous reports on the effects of these substitutions on RBD-ACE2 binding[9,12,14,22,27–30]. Reduced binding to ACE2 in K417N may result from a decrease in the flexibility of the 468 – 488 region that undergoes a conformational change when bound to ACE2 in the RBD variant containing only the K417N substitution. Analysis of the simulation trajectories suggests that glutamate 484 forms a hydrogen bond interaction with arginine 403 when only the K417N substitution is present as these residues are much further apart in simulation data generated from the wild-type and E484K RBD variant. The presence of the hydrogen bond between E484 and R403 results in a drastic decrease in the RMSD of the loop spanning residues 468 – 488 in K417N when compared to the wild-type and E484K RBD variant. The 468 – 488 loop region has been reported to undergo a structural transition that is important for ACE2 binding[31]. Therefore, a hydrogen bond between E484 and R403 that restricts the flexibility of the 468 – 488 loop may be responsible for the reduction in ACE2 binding affinity we observed here but more experiments will be necessary to determine this. More extensive simulation experiments in a previous study have shown that the K417N substitution may abolish a salt bridge of K417 with an aspartate residue at position 30 within ACE2 and therefore reduce the strength of binding interactions[14]. If this alone is enough to explain the observed reduction in ACE2 binding by the K417N variant or if it simply contributes remains to be determined.

The utility of the K417N substitution to the virus remains unclear. It has not been reported alone in other variants and this may be due to its negative effects on ACE2 binding reported here and by others[14]. K417N also appears to reduce the strength of antibody-RBD interactions and may have been selected for immune evasion despite its potential negative effects on transmissibility. Our data suggest that K417N alone alters the RBD structure in such a way that conformational changes necessary for ACE2 binding may be disfavored relative to the wild-type or other variants. We also observed that the E484K substitution alone significantly destabilizes the RBD by both denaturant induced unfolding and limited proteolysis. Therefore, the stabilizing effects of the K417N substitution may be necessary to offset the negative structural effects of the E484K substitution. The N501Y substitution also stabilized the RBD in our studies and the variant B.1.526 has been identified in New York that possesses both the N501Y and E484K substitutions[32]. The K417N and E484K substitutions have not yet been reported to exist alone in a SARS-CoV-2 variant and its possible their beneficial effects for immune evasion in the case of both K417N and E484K or ACE2 binding in the case of E484K are outweighed by negative structural effects. However, the caveat here is that the RBD does not exist in isolation but rather as part of the much larger spike glycoprotein and therefore the effects of these substitutions on structure and stability will likely not translate to the larger molecule but may still be relevant for understanding the evolution of SARS-CoV-2.

The presence of K417N alone and N501Y alone significantly stabilized the RBD relative to wild-type and it is unclear from the simulation data how this was achieved for the N501Y substitution found in the B.1.1.7 variant. Surprisingly the E484K substitution alone significantly destabilized the RBD relative to wild-type. Our simulation data suggest that E484K alone reduces overall hydrogen bond content in the RBD. Reduced hydrogen bond content may explain the observed decrease in stability and suggests that the elimination of specific interactions between residue 484 and other RBD residues alone cannot explain the observed reduction in RBD stability. Finally, we tested the stability of RBD variants by limited proteolysis with cathepsin S. Protein

susceptibility to limited proteolysis is a robust measure of protein stability and has also been linked to protein immunogenicity[19,20,23,24]. We observed that the K417N substitution increased proteolytic resistance at all time points while the E484K and all three substitutions together (K417N/E484K/N501Y) decreased resistance after 15 and 30 minutes of incubation. The N501Y variant exhibited increased proteolytic resistance only after 60 minutes of incubation. Our limited proteolysis results correlated with predictions from molecular dynamics as well as unfolding experiments.

In summary we have reported here that RBD variant of concern substitutions alter RBD structure and stability with K417N and N501Y increasing stability while E484K reduces stability. All three substitutions together, as found in the B.1.351 variant, exhibit similar stability to the wild-type RBD albeit with a possibly more open conformation and significantly higher ACE2 binding affinity that is greater than the sum of its parts. Taken together our findings suggest that the evolution of the SARS-CoV-2 RBD has been guided not only by pressure for increased ACE2 binding and immune evasion but also the need to compensate for potentially destabilizing substitutions in order to maintain RBD structure for optimal interactions with the ACE2 receptor.

Materials and Methods

Molecular Dynamic Simulations

Molecular dynamic simulations were performed with the GROMACS 2020.5 package[17,33] with the CHARMM36 all-atom force field[18] and the TIP3P water model[34]. The wild-type SARS-CoV-2 spike glycoprotein receptor binding domain (residues 319 – 541) and variants were modeled from the full length spike glycoprotein Cryo-EM structure PDB entry 6VXX[16] using SWISS-MODEL[15]. RBD models were solvated in a dodecahedral box with a minimum protein to edge distance of 1.5 nm and 37162 water molecules. The system charge was neutralized with approximately 105 Na^+ and 113 Cl^- ions at a concentration of 0.15 M. Assembled systems were minimized for 10 ps and then equilibrated for 100 ps in the *NVT* ensemble followed by further equilibration for 1000 ps with harmonic position restraints on heavy protein atoms (1000 kJ mol^{-1} nm^{-2}) and Berendsen coupling[35] to maintain temperature and pressure (P ~ 1 bar, and T ~ 300 K). Systems were further equilibrated for 25 ns in the *NPT* ensemble followed by three individual 10 ns production runs for each variant also in the *NPT* ensemble. Trajectory data were saved every 1 ps. Analysis was performed with GROMACS, VMD 1.9.3[36], PyMOL and GraphPad Prism. Molecular dynamics data are shown as the mean of the three independent production runs.

Protein Expression and Purification

The vector pCAGGS containing the receptor binding domain (RBD, residues 319-541) of the spike glycoprotein from SARS-CoV-2, Wuhan-Hu-1 was obtained from BEI resources. Individual K417N, E484K and N501Y substitutions were introduced to the by site-directed mutagenesis. The RBD variant containing three substitutions (K417N/E484K/N5101Y) was synthesized and cloned by Genscript into the pcDNA3.1 vector and subcloned into pCAGGS. Vectors were used to transform DH5α bacteria for transfection DNA prep. DNA was purified using the PureLink plasmid midiprep kit (Invitrogen, CA), filtered through a 0.22 µm filter and stored at -20 ˚C. The SARS-CoV-2 spike glycoprotein RBD was expressed in Freestyle 293 cells (ThermoFisher, MA) grown in Freestyle 293 media at 37 ˚C with 8% CO_2 and shaking at 135 rpm in a humidified incubator. Cells were transfected with plasmid DNA and 25 kDa polyethylenimine (Polysciences, PA) at 1 µg/mL in Freestyle 293 media. Media was harvested 4 days after transfection and clarified by centrifugation at 4000 xg for 10 minutes. His-tagged RBD protein was purified from cell culture media by Ni-NTA chromatography and an AKTA pure system. Columns were washed

with Buffer A (20 mM Sodium Phosphate pH 7.2, 500 mM NaCl) and bound protein was eluted with Buffer B (Buffer A + 250 mM Imidazole). Protein was then desalted using a HiPrep 26/10 desalting column and phosphate buffered saline (PBS) pH 7.4 and concentrated using a 10 kDa centrifugal concentrator (Milipore-Sigma, MA). Protein concentration was determined by absorbance at 280 nm and an estimated extinction coefficient of 33850 M^{-1} cm^{-1}. Concentrated protein was aliquoted, snap frozen in liquid nitrogen and stored at -80 ˚C.

Unfolding Experiments

Guanine unfolding experiments were performed by fluorescent dye assay using bis-ANS (4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid, Tocris, Bristol, UK) as described previously[20] with minor modifications. Protein (5 µM) was mixed with bis-ANS (10 µM) in PBS and guanidine HCl (prepared in PBS) ranging from 0 M to 2.5 M in a 96-well black plate in duplicate. Mixtures were incubated for 1 hour at room temperature before fluorescence readings were taken using a GloMax Explorer (Promega, WI) plate reader with an excitation filter at 405 nm and emission filter at 500 – 550 nm. Fluorescence data were used to calculate the fraction folded at a given guanidine concentration and analyzed by non-linear regression [21] to estimate free energy of unfolding and best-fit values were compared by one-way ANOVA. Data presented are representative of three independent experiments.

RBD-ACE2 Binding Competition Assay

RBD-ACE2 binding competition assay was developed using the SARS-CoV-2 surrogate virus neutralization test kit (Genscript, NJ). First a 5-fold dilution series of RBD variant starting at 10 µM was prepared in sample dilution buffer in duplicate. Next, the serum/antibody incubation step was replaced by incubating horseradish peroxidase-labeled RBD (RBD-HRP) 1:1 with diluted RBD variant and incubated at 37 ˚C for 30 minutes. Then the kit procedure was followed as described in the manufacturer's instructions. After development absorbance at 450 nm was measured with the GloMax Explorer plate reader. Absorbance values were plotted as a function of RBD variant concentration and analyzed by non-linear regression to estimate LogIC50 values. Best-fit values were compared by one-way ANOVA. Data presented are representative of three independent experiments.

Limited Proteolysis

Limited proteolysis experiments were conducted with recombinant human cathepsin S (Milipore-Sigma) diluted in phosphate-citrate buffer at pH 5.6 and 37 ˚C[20]. Proteolysis reactions were prepared in 150 µL phosphate-citrate buffer pH 5.6 containing 0.5 µg/µL of RBD variant, 0.025 µg/µL Cathepsin S and 2 mM dithiothreitol. Reactions were prepared in triplicate and sampled after 0, 15, 30 and 60 minutes of incubation. Proteolysis was halted by mixing with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer containing 150 mM 2-mercaptoethanol and incubation at 95 ˚C for 5 minutes. Proteolysis reactions were analyzed by SDS-PAGE and staining with Coomassie blue. Gels were imaged with a Chemidoc MP imaging system (Bio-Rad, CA) and analyzed with ImageJ software. Band intensities of the intact RBD were extracted from images and normalized to the 0-minute band intensity before analysis by two-way ANOVA.

Acknowledgments

The following reagent was produced under HHSN272201400008C and obtained through BEI Resources, NIAID, NIH: Vector pCAGGS Containing the SARS-Related Coronavirus 2, Wuhan-Hu-1 Spike Glycoprotein Receptor Binding Domain (RBD), NR-52309.

Competing Interests

The authors have no conflicts of interest to declare.

References

1. Feehan, J. & Apostolopoulos, V. Is COVID-19 the worst pandemic? *Maturitas* (2021) doi:10.1016/j.maturitas.2021.02.001.
2. Wu, A. *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* **27**, 325–328 (2020).
3. Frampton, D. *et al.* Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *The Lancet Infectious Diseases* **0**, (2021).
4. Dejnirattisai, W. *et al.* Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* (2021) doi:10.1016/j.cell.2021.03.055.
5. Yang, Z. *et al.* Evasion of antibody neutralization in emerging severe acute respiratory syndrome coronaviruses. *PNAS* **102**, 797–801 (2005).
6. Zhou, D. *et al.* Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* (2021) doi:10.1016/j.cell.2021.02.037.
7. Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* **5**, 562–569 (2020).
8. Ali, F., Kasry, A. & Amin, M. The new SARS-CoV-2 strain shows a stronger binding affinity to ACE2 due to N501Y mutant. *Med Drug Discov* **10**, 100086 (2021).
9. Luan, B., Wang, H. & Huynh, T. Enhanced binding of the N501Y-mutated SARS-CoV-2 spike protein to the human ACE2 receptor: insights from molecular dynamics simulations. *FEBS Lett* (2021) doi:10.1002/1873-3468.14076.
10. Hoffmann, M. *et al.* SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell* (2021) doi:10.1016/j.cell.2021.03.036.

11. Wang, Z. *et al.* mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* 1–7 (2021) doi:10.1038/s41586-021-03324-6.
12. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).
13. Wibmer, C. K. *et al.* SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nature Medicine* **27**, 622–625 (2021).
14. Luan, B. & Huynh, T. Insights into SARS-CoV-2's Mutations for Evading Human Antibodies: Sacrifice and Survival. *J Med Chem* (2021) doi:10.1021/acs.jmedchem.1c00311.
15. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296–W303 (2018).
16. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* (2020) doi:10.1016/j.cell.2020.02.058.
17. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
18. Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **34**, 2135–2145 (2013).
19. Moss, D. L., Mettu, R. R. & Landry, S. J. The Serpin-like Loop Insertion of Ovalbumin Increases the Stability and Decreases the OVA 323–339 Epitope Processing Efficiency. *Biochemistry* (2021) doi:10.1021/acs.biochem.1c00095.
20. Moss, D. L., Park, H.-W., Mettu, R. R. & Landry, S. J. Deimmunizing substitutions in pseudomonas exotoxin domain III perturb antigen processing without eliminating T-cell epitopes. *J. Biol. Chem.* jbc.RA118.006704 (2019) doi:10.1074/jbc.RA118.006704.
21. Manyusa, S. & Whitford, D. Defining Folding and Unfolding Reactions of Apocytochrome b5 Using Equilibrium and Kinetic Fluorescence Measurements. *Biochemistry* **38**, 9533–9540 (1999).

22. Tian, F. *et al.* Mutation N501Y in RBD of Spike Protein Strengthens the Interaction between COVID-19 and its Receptor ACE2. *bioRxiv* 2021.02.14.431117 (2021) doi:10.1101/2021.02.14.431117.
23. Delamarre, L., Couture, R., Mellman, I. & Trombetta, E. S. Enhancing immunogenicity by limiting susceptibility to lysosomal proteolysis. *Journal of Experimental Medicine* **203**, 2049–2055 (2006).
24. Machado, Y. *et al.* Fold stability during endolysosomal acidification is a key factor for allergenicity and immunogenicity of the major birch pollen allergen. *Journal of Allergy and Clinical Immunology* **137**, 1525–1534 (2016).
25. Khan, A. *et al.* Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *Journal of Cellular Physiology* **n/a**, (2021).
26. Laffeber, C., Koning, K. de, Kanaar, R. & Lebbink, J. H. Experimental evidence for enhanced receptor binding by rapidly spreading SARS-CoV-2 variants. *bioRxiv* 2021.02.22.432357 (2021) doi:10.1101/2021.02.22.432357.
27. Ahmed, W., Philip, A. M. & Biswas, K. H. Stable Interaction Of The UK B.1.1.7 lineage SARS-CoV-2 S1 Spike N501Y Mutant With ACE2 Revealed By Molecular Dynamics Simulation. *bioRxiv* 2021.01.07.425307 (2021) doi:10.1101/2021.01.07.425307.
28. Nelson, G. *et al.* Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escape mutant. *bioRxiv* 2021.01.13.426558 (2021) doi:10.1101/2021.01.13.426558.
29. Rezaei, S., Sefidbakht, Y. & Uskoković, V. Comparative molecular dynamics study of the receptor-binding domains in SARS-CoV-2 and SARS-CoV and the effects of mutations on the binding affinity. *Journal of Biomolecular Structure and Dynamics* **0**, 1–20 (2020).

30. Wang, W. B. *et al.* E484K mutation in SARS-CoV-2 RBD enhances binding affinity with hACE2 but reduces interactions with neutralizing antibodies and nanobodies: Binding free energy calculation studies. *bioRxiv* 2021.02.17.431566 (2021) doi:10.1101/2021.02.17.431566.
31. Yesudhas, D., Srivastava, A., Sekijima, M. & Gromiha, M. M. Tackling Covid-19 using disordered-to-order transition of residues in the spike protein upon angiotensin-converting enzyme 2 binding. *Proteins: Structure, Function, and Bioinformatics* **n/a**, (2021).
32. West, A. P. *et al.* Detection and characterization of the SARS-CoV-2 lineage B.1.526 in New York. *bioRxiv* 2021.02.14.431043 (2021) doi:10.1101/2021.02.14.431043.
33. Lindahl, Abraham, Hess & van der Spoel. *GROMACS 2020.5 Source code*. (Zenodo, 2021). doi:10.5281/zenodo.4420785.
34. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
35. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
36. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).

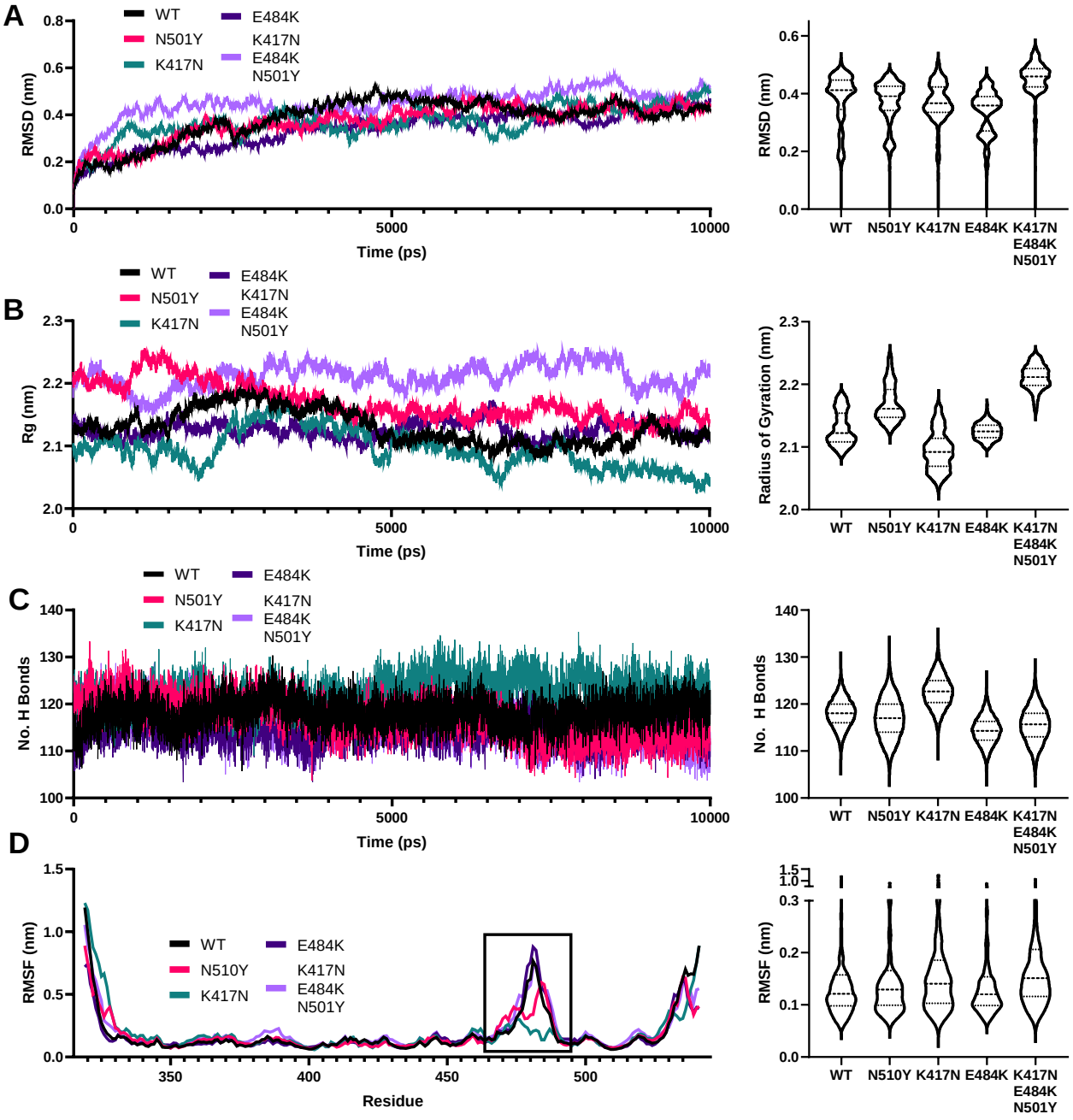

Figure 1: Molecular Dynamic Simulations of SARS-CoV-2 Receptor Binding Domain Suggests that Variant-of-Concern Substitutions Alter RBD Structure. A. Root mean square deviation plotted as a function of time (left) and averaged as violin plot (right) for each RBD variant simulated. Data shown are the average of three separate production runs for each variant (n = 3). B. Radius of gyration plotted as a function of time (left) and averaged as violin plot (right) for each RBD variant (n = 3). C. Total number of hydrogen bonds within the RBD for each variant simulated plotted as a function of time (left) and averaged and plotted in a violin plot (right). D. Root mean square fluctuation plotted as a function of time (left) or averaged and shown as a violin plot (right). Boxed area covers residues 468 – 488.

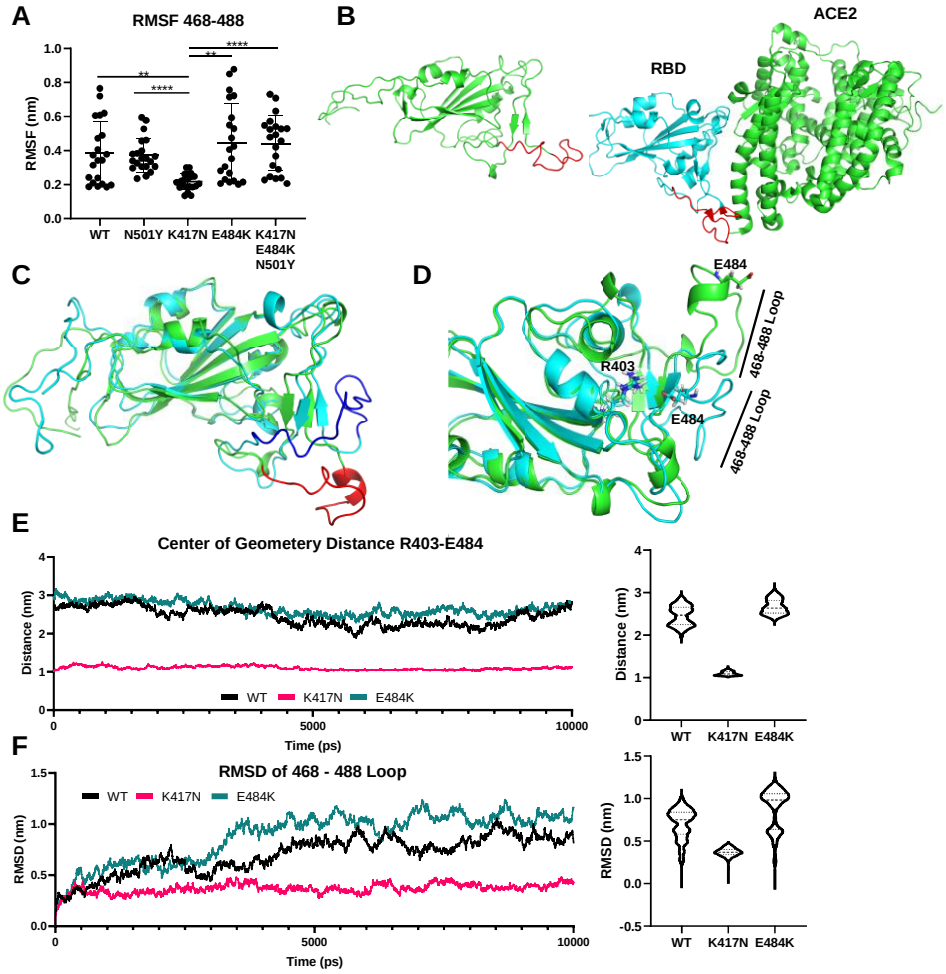

Figure 2: The K417N Substitution Alters Flexibility of an ACE2 Binding Loop. A. RMSF values from figure 1D for the loop region covering residues 468 – 488. Mean values were compared by Brown-Forsythe ANOVA and Dunnett's test for multiple comparisons with p values < 0.01 indicated by ** and p values < 0.0001 indicated by ****. Error bars show standard deviation and n = 21. B. Crystallographic models of the RBD (left) and RBD(cyan)-ACE2(green) complex (right). The RBD structure was generated by homology modeling using the cryo-EM structure of the whole spike as a template (PDB 6VXX) while the RBD-ACE2 complex model was produced using the PDB structure 6M0J. In both models the 468 – 488 loop is colored red. C. Snapshot of RBD simulation trajectories for the wild-type (green) and K417N (cyan) models. Coordinates were aligned and the 468 – 488 loop is colored in red for wild-type and blue for K417N. D. Aligned coordinates shown in C rotated to show the positions of arginine 403 and glutamate 484 in the wild-type (green) and K417N (cyan) models. The 468 – 488 loop is indicated. E. Center of geometry distances plotted as a function of time (left) and averaged shown as a violin plot (right) from the simulation data of the wild-type, K417N and E484K RBD models. Distance data was computed for three separate production runs for each variant and shown as an average. F. RMSD of residues 468 – 488 plotted as a function of time (left) and averaged in a violin plot (right). RMSD was calculated for each separate production run and plotted as an average for all three RBD variants shown.

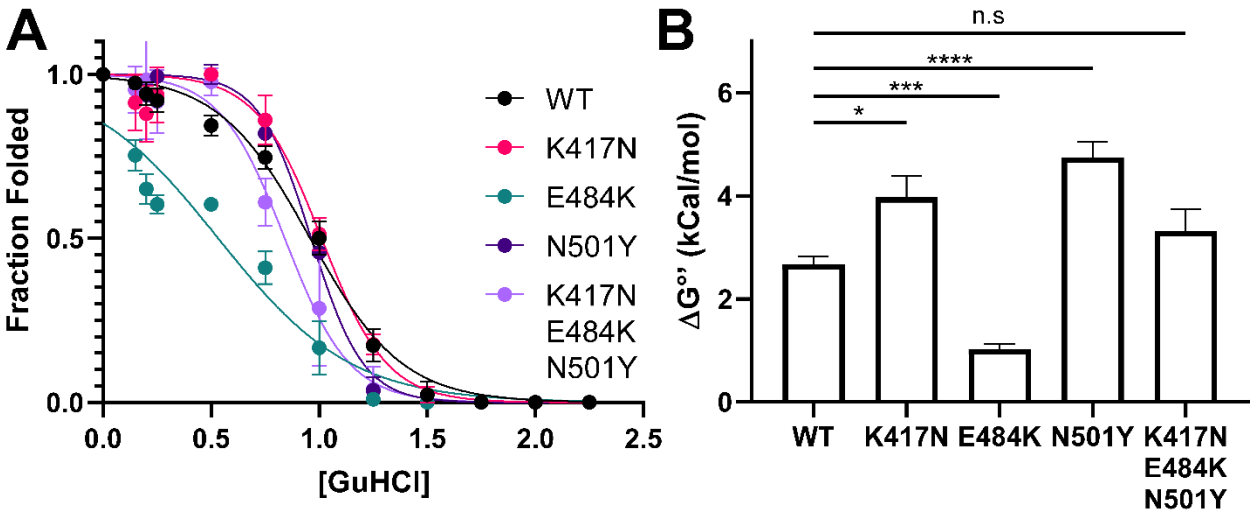

Figure 3: RBD Variant of Concern Substitutions Alter RBD Stability. A. Unfolding curves for the indicated RBD variant generated by monitoring Bis-ANS fluorescence as a function of guanidine-HCl concentration. Non-linear regression was used to analyze unfolding data. B. Calculated ΔG'° values from non-linear regression analysis for each RBD variant were compared by one-way ANOVA and Tukey's test for multiple comparisons. P value less than 0.05 is indicated by *, p value less than 0.01 is indicated by *** and less than 0.0001 is indicated by ****. Data here are representative of three separate unfolding experiments.

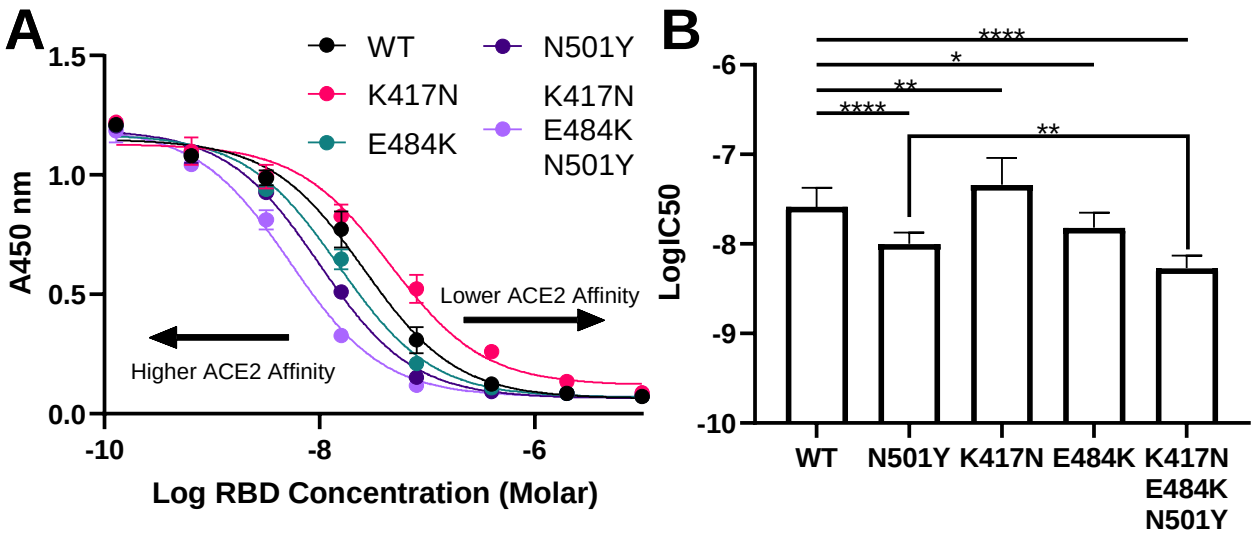

Figure 4: RBD Variant of Concern Substitutions Alter ACE2 Binding Affinity. A. Inhibition curves for each indicated RBD variant. As the RBD variant is diluted out HRP labeled wild-type RBD can outcompete for binding to ACE2 resulting in an increase in absorbance at 450 nm. Data were analyzed by non-linear regression. B. LogIC50 values from A were compared by one-way ANOVA and Tukey's test for multiple comparisons with asterisks indicating p value: * for < 0.05, ** for < 0.01, *** for < 0.001 and **** for < 0.0001. Data shown are representative of three independent experiments.

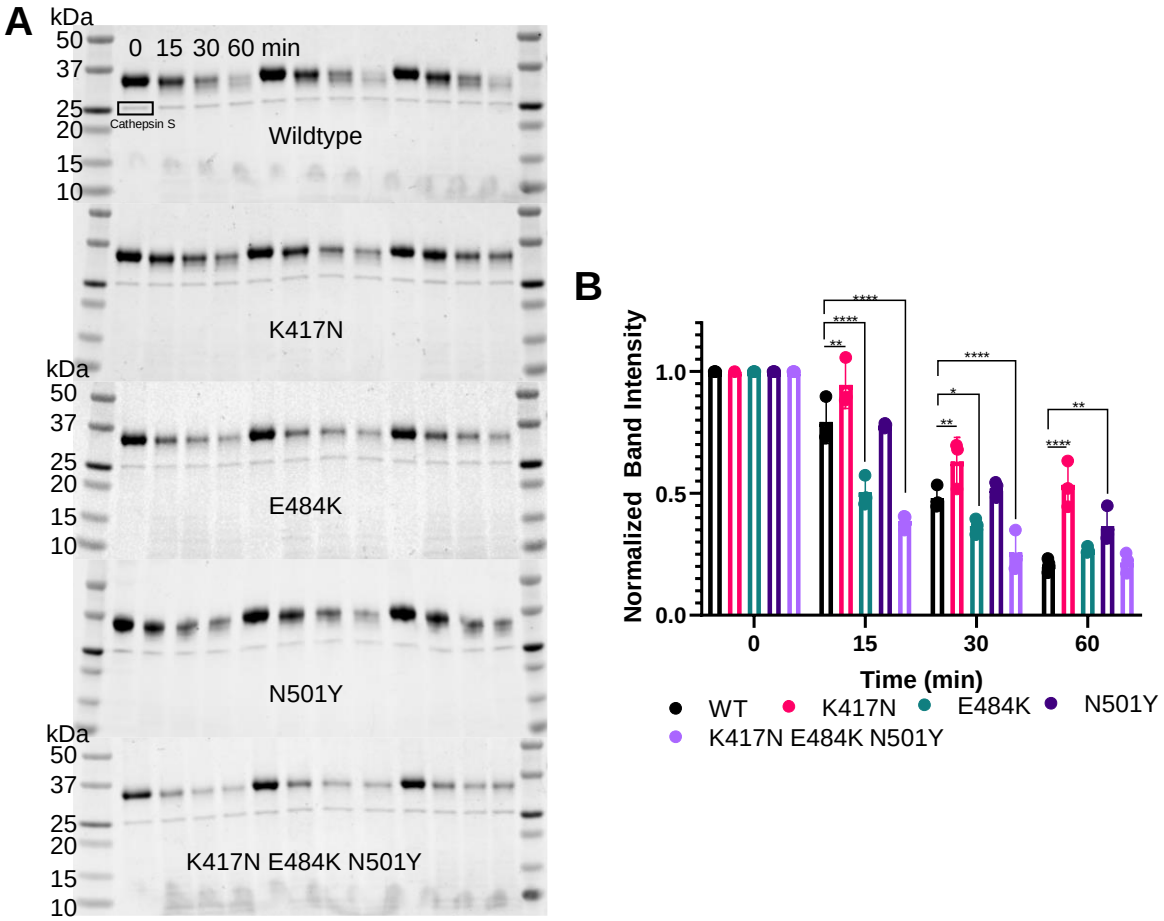

Figure 5: RBD Variant of Concern Substitutions Alter Susceptibility to Limited Proteolysis by Cathepsin S. A. SDS-PAGE and Coomassie staining of limited proteolysis reactions for each RBD variant indicated run in triplicate. B. Band intensities were calculated from A and normalized to the 0 min time point. Mean intensities for each replicate are shown, error bars indicate standard deviation and n = 3. Mean values were compared by two-way ANOVA and Dunnett's test for multiple comparisons asterisks indicating p value: * for < 0.05, ** for < 0.01 and **** for < 0.0001.