Interplay between protein stability, binding to ACE2 and escape from neutralizing antibodies determines the natural selection of SARS-CoV-2 receptor binding domain variants

Vaibhav Upadhyay, Alexandra Lucas, Sudipta Panja and Krishna M.G. Mallela*

Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Anschutz Medical Campus, Aurora, CO 80045.

*Corresponding author

Krishna Mallela, PhD
Department of Pharmaceutical Sciences
University of Colorado Anschutz Medical Campus
12850 E. Montview Blvd, MS C238-V20
Aurora, CO 80045.
Phone: 1-303-724-3576
Fax: 1-303-724-7266
E-mail: krishna.mallela@cuanschutz.edu

Abbreviations

ACE2, angiotensin converting enzyme 2; CD, circular dichroism; COVID-19, coronavirus disease 2019; Fab, antigen binding fragment; HEK, human embryonic kidney; ITC, isothermal titration calorimetry; MERS-CoV, Middle East respiratory syndrome coronavirus; Ni-NTA, nickel nitrilotriacetic acid; PEI, polyethylenimine; RBD, receptor binding domain; RBM, receptor binding motif; SARS-CoV, severe acute respiratory syndrome coronavirus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; ScFv, single chain variable fragment; Tm, midpoint temperature of thermal denaturation; VOC, variant of concern.

Abstract

Emergence of new SARS-CoV-2 variants has raised concerns at the effectiveness of vaccines and antibody therapeutics developed against the unmutated wild-type virus. It is thus important to understand the emergence of mutants from an evolutionary viewpoint to be able to devise effective countermeasures against new variants. We examined the effect of 12 most commonly occurring mutations in the receptor binding domain (RBD) on its expression, stability, activity, and antibody escape potential- some of the factors that may influence the natural selection of mutants. Recombinant proteins were expressed in human cells. Stability was measured using thermal denaturation melts. Activity and antibody escape potential were measured using isothermal titration calorimetry in terms of binding of RBD variants to ACE2 and to a neutralizing human antibody CC12.1, respectively. Our results show that variants differ in their expression levels, suggesting that mutations can impact the availability of proteins for virus assembly. All variants have similar or higher stability than the wild-type, implying that increased RBD stability might be another important factor in virus evolution. In terms of ACE2 binding, when compared to the wild-type, only 3 out of 7 expressed variants show stronger affinity, 2 have similar affinity, whereas the other 2 have weaker affinity, indicating that increased affinity towards ACE2 is an important but not the sole factor in the natural selection of variants. In terms of CC12.1 binding, when compared to the wild-type, 4 out of 7 variants have weaker affinity, 2 have a similar affinity, and 1 variant has a stronger affinity. Taken together, these results indicate that multiple factors contribute towards the natural selection of variants, and all of these factors have to be considered in order to understand the natural selection of SARS-CoV-2 variants. In addition, since not all variants can escape a given neutralizing antibody, antibodies to treat new variants can be chosen based on the specific mutations in that particular variant.

Introduction

Corona virus disease 2019 (COVID-19) pandemic has emerged as a global threat in December 2019. As of May 2021, it has infected over 167 million people and claimed 3.5 million lives (https://coronavirus.jhu.edu/map.html). The causative agent for COVID-19 is severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a single stranded RNA virus, that belongs to sarbecovirus subgenus of betacoronaviruses.[1] It shares close genomic similarity to SARS-CoV (79% identity) and Middle East respiratory syndrome coronavirus (MERS-CoV, 50% identity) that were responsible for SARS outbreak in 2002 and MERS outbreak in 2012 respectively.[2-6] These viruses are thought to have originated in bats and transmitted to other mammals including humans.[7]

Both SARS-CoV and SARS-CoV-2 enter the human host through interaction of its spike protein with angiotensin-converting enzyme 2 (ACE2) present on the membrane of host epithelial cells.[8-10] Specifically, the receptor binding domain (RBD) of the spike protein binds with ACE2 and thus is a major determinant of the viral infectivity and evolution[11-13]. The viral evolution through accumulation of mutations in SARS-CoV-2 is slower than known for other RNA viruses like HIV and influenza.[14-16] Still, SARS-CoV-2 variants pose a major challenge for devising measures to counter the virus threat, as new variants continue to emerge, some of which are believed to be more infectious than the wild-type virus.[17-19] Among all the mutations happening in the viral genome, mutations in RBD are considered to play a significant role in infectivity due to its role in ACE2 binding. Most of the neutralizing antibody response of the host is generated against RBD,[20,21] because of which RBD is also a major target for most of the therapeutic antibodies developed.[22-24] Mutations in RBD are predicted to dictate the emergence of escape mutants and shape the evolutionary path of the virus through the process of natural selection that would favor the mutants that could evade the antibody response. Apart from this, RBD alone or as part of spike protein is also used as antigen in many prospective vaccines.[25-35] The emergence of mutations in RBD is considered to have an impact on the effectiveness of these vaccines. Lower efficacy of some of the vaccines was reported against the emerging variants in UK, South Africa and Brazil along with reports of new variants escaping the antibodies approved for emergency use.[36-39]

In this work, we examined the effect of new variants on RBD protein expression, stability, its binding to ACE2, and antibody escape using a naturally occurring human neutralizing antibody CC12.1. We hypothesize that all these factors act in conjunction and can determine the virus evolution and natural selection of new variants. To test, we selected 12 most frequently occurring mutations in RBD as of January 2021 (https://www.gisaid.org/hcov19-mutation-dashboard). These include 9 single site mutations K417N, N439K, Y453F, S477N, S477I, T478I, E484K, S494P and N501Y (UK variant), a double mutant (E484K, N501Y), and two triple mutants (K417N, E484K, N501Y (South Africa variant); K417T, E484K, N501Y (Brazil variant)). Location of these residues in RBD bound to ACE2 and CC12.1 are shown in Figs. 1A and 1B. All these mutating residues are in the receptor binding motif (RBM) of RBD. Our results presented below indicate that multiple factors contribute towards the natural selection of variants, and all of these factors have to be considered in order to understand the natural selection of SARS-CoV-2 variants.

Methods

Cloning and expression of RBD, RBD mutants, ACE2 and CC12.1 ScFv

Amino acid sequences of SARS-CoV-2 RBD and human-ACE2 protein were obtained from Uniprot (Uniprot ID-P0DTC2 and Q9BYF1 respectively). The sequence of heavy and light chain variable regions (V_H and V_L respectively) of CC12.1 neutralizing antibody was obtained from RCSB-PDB (PDB ID 6XC2). The single chain variable fragment (ScFv) for CC12.1 was designed as V_H-$(GGGGS)_3$-V_L. The protein sequences were codon-optimized for expression in human cells and synthesized by Twist Biosciences. The RBD variants were generated by site directed mutagenesis using mutagenic primers. The synthesized and mutated genes were cloned into pcDNA 3.4 Topo vector, modified by including a signal sequence of human immunoglobulin heavy chain, His-Tag and SUMOstar protein. The expression vectors were transfected transiently into Expi293 (modified HEK293) cells using polyethylenimine (PEI) and the protein was recovered from the culture supernatant after 5 days. The expression levels for the proteins were compared after running the culture supernatants on SDS-PAGE and quantifying the band intensities corresponding to the target protein using the ImageLab software from Bio-Rad.

Protein Purification

The supernatant of the culture was filtered through 0.22 μm filter and purified using Nickel-nitrilotriacetic acid (Ni-NTA) chromatography. The eluted protein was dialyzed in buffer containing 50 mM of Tris-HCl, 20 mM NaCl, pH 8.0 and digested using SUMOstar protease overnight to cleave the target protein from SUMOstar and His-Tag. The digested protein was passed again through the Ni-NTA column to obtain the untagged target protein in the flow-through. All proteins were dialyzed in buffer containing 50 mM Sodium Phosphate, 20 mM NaCl, pH 7.0.

Isothermal titration calorimetry (ITC)

ITC experiments were performed using MicroCal-PEAQ-ITC instrument from Malvern in buffer containing 50 mM Sodium Phosphate, 20 mM NaCl, pH 7.0 at 20° C. For ACE2-RBD interactions, ACE2 at a concentration of 12 μM was taken in the cell and RBD variants at concentration of 120 μM were taken in the syringe. For RBD-CC12.1 ScFv interactions, RBD or its variants were taken in the cell at a concentration of 20 μM and CC12.1 ScFv was taken in the syringe at 200 μM. The syringe contents were titrated into the cell as 18 injections of 2 μl each with 150 s spacing between the injections. ITC data analysis was done using MicroCal PEAQ-ITC Analysis Software from Malvern.

Circular dichroism (CD) spectroscopy

CD spectra of RBD variants were recorded on an Applied Photophysics Chirascan Plus spectrometer at a protein concentration of 5 μM in buffer containing 10 mM Sodium Phosphate, 4 mM NaCl, pH 7.0 in a 1 mm pathlength cuvette. The data was collected at an interval of 1 nm and averaged for 2 s at each wavelength.

The thermal melts for RBD variants were recorded at a protein concentration of 2 μM in buffer containing 50 mM Sodium Phosphate, 20 mM NaCl, pH 7.0 in 1 cm pathlength cuvette. The temperature scan rate was 1° C/min. CD signal at 222 nm averaged for 2 s was plotted against the temperature and fitted to a two-state unfolding model using the equation,

$$S_T = \frac{(S_N+m_NT)+(S_U+m_UT)e^{-\left(\frac{\Delta H_m}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right)}}{1+e^{-\left(\frac{\Delta H_m}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right)}} \quad \text{Eq. 1}$$

where S_T is the measured signal as a function of temperature T, S_N and S_U are the signals corresponding to the native and unfolded baselines, m_N and m_U are the slopes of linear dependence of S_N and S_U, ΔH_m is the enthalpy change at T_m, R is the universal gas constant and T is the absolute temperature in Kelvin.

Results

RBD mutations affect protein expression

Protein expression was performed in human embryonic kidney (HEK) cells, in which the protein transport through the secretory pathway, the post-translational modifications and quality control mechanisms ultimately decide the secreted protein levels.[40] Expression in HEK cells closely match the natural infection scenario, where the virus uses the host cell machinery to synthesize its structural proteins. The protein expression levels have a direct bearing on the yield of the viruses, as the virus yield is directly proportional to the amount of the proteins available for virus assembly.[41] The infectivity of a particular variant can thus be dependent on the protein expression levels. It should be noted however, that this study only represents the expression level of RBD and in natural scenario the expression level of the complete spike protein would ultimately decide the virus yield and thus infectivity. The relative expression of the wild-type RBD along with the mutants were compared using SDS-PAGE after 3 days of expression (Figs. 2A and 2B). Among 9 single site mutants, 2 mutants (N439K and S477I) did not express very well and appreciable amount of protein needed for binding studies could not be obtained for these mutants. The levels of expression were also low for 2 other single site mutants- T478I and E484K. Expression levels comparable to or higher than the wild type was obtained for 5 single site mutants- K417N, Y453F, S477N, S494P, and N501Y. Apart from the single site mutations, clone carrying double mutations- E484K and N501Y did not express. The clone carrying triple mutations (K417N, E484K and N501Y) corresponding to the South Africa variant also could not be expressed, but the other clone carrying triple mutations (K417T, E484K and N501Y) corresponding to the Brazil variant showed high expression. These results suggest that RBD mutations can impact the overall protein expression levels, which can possibly affect the virus yield and infectivity. Further studies on the mutant proteins were thus carried out on 6 single site mutants (K417N, Y453F, S477N, E484K, S494P, and N501Y) and one triple mutant corresponding to the Brazil variant, which were purified to homogeneity along with ACE2 and CC12.1 ScFv (Figs. 3A and 3B).

Mutations do not significantly affect the global protein structure

Most of the random mutations in proteins do not affect the protein structure and are thus considered neutral. Some mutations though, can bring significant structural changes, and can have either a beneficial or a deleterious effect on virus fitness. In the absence of other selection pressures, the deleterious mutations are lost but the beneficial mutations get selected and prevail. Here we investigated the effect of 7 frequently occurring mutations on the RBD structure using far-UV CD spectroscopy (Fig. 3C). The CD spectra for the wild-type protein showed a major negative band at 208 nm and a positive band at 192 nm. The CD spectra was similar to previously reported spectra for wild-type RBD.[42] The far-UV CD spectra of all 7 mutants were similar to that of the wild-type protein with most significant difference observed for the Brazil

triple mutant. Deconvolution using BeStSel web software[43] reveal a low percentage alpha-helix along with high beta-sheet content for all the variants. The proportion of the random coil structure is substantial for RBD accounting for about 50% (Table 1), primarily originating from the RBM (Fig. 1). These CD results suggest that the mutations do not significantly affect the global structure of RBD, and all the mutants adopt a similar fold like the wild-type RBD. This is also consistent with the soluble expression of variants (Fig. 3A). Proteins that are expressed as soluble in general adopt well-folded structures. This shows that only those mutants that do not have any deleterious effect on the protein structure are getting naturally selected. Since software used for deconvoluting CD spectra in general accounts only for regular secondary structures, minor differences in the CD spectra of the Brazil variant must be due to changes in short, irregular and non-repeating secondary structures.

RBD mutants have similar or higher thermal stability than wild-type protein

Protein stability can be an important factor in protein evolution. More stable proteins can accommodate wide range of mutations and determine the evolvability of the proteins.[44] It is thus important to determine the stability of protein to be able to determine its evolutionary path. The stability of the wild-type RBD and its mutants were assessed with thermal denaturation melts using far-UV CD spectroscopy (Fig. 4). Thermal denaturation of the wild-type protein shows a co-operative transition with slopy native and denatured baselines. The transition could be fitted well to a two-state unfolding model, giving T_m value of 55.5 ± 1.0 °C (Table 2). The thermal denaturation curves of the mutant proteins were similarly obtained and were found to be cooperative, which fitted well to a two-state unfolding model. All the mutants showed T_m values which were comparable to or higher than the wild-type protein, ranging from 54.7 ± 0.8 - 58.2 ± 0.7 °C. These results suggest that the virus is evolving to increased RBD stability, which may have fitness advantages for the virus. It also points toward the possibility of increased occurrence of mutations in variants that have higher stability.

Some but not all RBD mutants show increased binding affinity to ACE2

SARS-CoV-2 RBD mediates the interaction of the virus spike protein with the ACE2 receptor on the host cell surface and is thus, a major determinant of the viral entry into the host cell. Mutations in the RBD can impact its interaction with ACE2 and can have an important role in determining the infectivity of the virus, with higher affinity interactions contributing to increased infectivity. The binding interactions of the wild-type and mutant RBD proteins with ACE2 were investigated using ITC (Fig. 5 and Table 3). The wild-type RBD interacts with ACE2 in a stoichiometric ratio of 1:1, with K_d value of 10.0 ± 3.1 nM and enthalpy of interaction (ΔH value) of 11.8 ± 0.2 kcal/mol. The measured K_d value is consistent with previously published studies on ACE2-RBD interaction using surface plasmon resonance (SPR) with immobilized protein.[45] All the mutants studied interacted with ACE2 in the same stoichiometric ratio of 1:1. Two of the 7 mutants, Y453F and S494P, did not show significant difference in their binding interaction with ACE2, with K_d and ΔH values similar to the wild-type protein (Table 3). For two other mutants K417N and E484K, the K_d value obtained was higher than the wild-type protein (Table 3), indicating weaker affinity for ACE2. The corresponding ΔH values did not show any significant difference for K417N mutant but showed a decreased value of 9.5 ± 0.2 kcal/mol for E484K mutant. The other two single site mutants (S477N and N501Y) and the Brazil triple mutant showed increased affinity for ACE2 binding with K_d values of 5.0 ± 1.4 nM, 3.0 ± 2.1 nM and 1.6 ± 1.5 nM respectively (Table 3). There was no difference in ΔH values for

N501Y mutant and the triple mutant but an increased ΔH value of 16.4 ± 0.2 kcal/mol was obtained for S477N mutant, which may indicate increased interactions between RBD and ACE2 upon mutation. These ITC results indicate that the ACE2 binding is an important but not the sole parameter for virus evolution.

Some but not all RBD mutants show CC12.1 antibody escape

CC12.1 antibody is a monoclonal antibody isolated from the convalescent plasma of a COVID-19 survivor.[46,47] This antibody represents the class of antibodies most elicited by SARS-CoV-2 infection. The circulating antibodies can play a major role in shaping the evolutionary path of RBD. The mutants that can escape the antibody recognition are naturally selected and represented more in the viral pool. We investigated the binding of the wild type RBD and its mutants to the ScFv of CC12.1 antibody through ITC (Fig. 6 and Table 4). The CC12.1 ScFv interacts with wild type RBD at a stoichiometric ratio of 1:1 with a K_d value of 28 ± 8.8 nM and ΔH value of 4.8 ± 0.1 kcal/mol. Compared to ACE2 interaction which is primarily enthalpically driven (Table 3), RBD interaction with CC12.1 has a significant entropic component (Table 4). Out of the 7 mutants studied, 2 single site mutations (S477N and S494P) did not have any impact on the binding affinity of RBD towards CC12.1 ScFv (Table 4). E484K mutant however showed increased affinity towards CC12.1 ScFv with a K_d value of 1.7 ± 4.6 nM (Table 4), which suggests that CC12.1 may be able to neutralize the E484K mutant. Other 4 variants, single site mutants K417N, Y453F, N501Y and the Brazil triple mutant, all of which showed decreased affinity towards CC12.1 binding, with K_d values of 119.0 ± 50.4 nM, 827 ± 146 nM, 62.7 ± 21.9 and 433 ± 94.9 nM respectively, representing escape from CC12.1. Another interesting observation is the decrease in ΔH value for K417N and Y453F indicating decreased strength of interactions between the mutants and CC12.1. These results show that antibody escape can be an important parameter but not the only parameter that shapes virus evolution and natural selection of mutants.

Discussion

The reports of mutations in the spike protein of SARS-CoV-2 started to appear very soon after its emergence in Wuhan, China in December 2019. Most of the mutations in the RBD were noticed during the Fall of 2020, and variant B.1.1.7 with RBD mutation N501Y became one of the first RBD variants to be labelled as a variant of concern (VOC). Since the beginning of 2021, there has been an emergence of numerous variants in different parts of the world. The rate of mutations in SARS-CoV-2 genome is lesser than that known for other RNA viruses like influenza and HIV.[14-16] The prime reason for that is the ability of the virus to proof-read errors included by RNA dependent RNA polymerase due to the presence of exonuclease activity.[48] Even then, there are a substantial number of variants occurring with high frequency reported for this virus, with the list continuing to grow. This can be explained by the fact that the appearance of mutations is also dependent on the viral population size. Higher viral population will support greater number of mutants. As this virus has infected human population globally and has emerged as a pandemic, it is not surprising to encounter the different variants. The biggest concern with the emergence of the variants is the ineffectiveness of the currently developed vaccines, which are being used to vaccinate people and the antibody-based therapeutics, which are approved for emergency use. There are reports of breakthrough infection in already vaccinated individuals,[49] and also reports of decreased neutralization ability of vaccinated individuals' serum against varaints,[36-38] which substantiate the need to study these variants in more detail.

The occurrence and persistence of a particular mutation in virus pool is dependent on a number of factors. In general, it is considered that mutations that provide a selective advantage to virus fitness are naturally selected. The most important characteristics of proteins that can affect virus fitness are its stability and activity. Mutations in viral proteins can have a neutral, stabilizing or destabilizing effect on protein stability. Stabilizing mutations offer a fitness advantage to the virus by increasing the proportion of correctly folded protein and increased resistance to protein degradation and aggregation inside the cell.[50,51] Higher stability, similar to higher expression, can thus lead to higher infectivity through increased virus yield.[52] Destabilizing mutations on the other hand do not offer fitness advantage and are mostly deleterious.[53] They lead to lower virus yield with decreased infectivity. For example, in the case of hemagglutinin protein of Influenza virus, it has been shown that the higher stability variants provide fitness advantage and tend to persist longer.[52] The mutation rate and the population size also affect the stability of the new variants. It has been shown that higher mutation rates and lower population size lead to the emergence of low stability variants and on the contrary low mutation rates and high population sizes lead to emergence of variants with higher stability.[54] Our results with SARS-CoV-2 RBD support this hypothesis. Most of the mutations that we examined have either similar or higher stability than the wild-type protein. These results should be important in understanding the persistence of the mutants and predicting which variants have the potential to become VOC. Looking at the data, the triple mutant (Brazil variant) has the highest T_m and we predict it to prevail for longer times.

Protein activity is another parameter which affects virus fitness. Unlike protein stability which can be impacted by mutations at many sites on protein, protein activity is controlled by a few key amino acid residues. In case of SARS-CoV-2 RBD, binding to ACE2 with a higher affinity provides a selective advantage towards virus fitness. Higher the affinity, higher the virus infectivity. The key residues of RBD which interact with ACE2 residues are shown in Fig. 7A. Most of the frequently occurring mutations (Fig. 1A) are at the binding interface. Our data on ACE2 binding of the RBD mutants shows that the two mutations Y453F and S494P do not impact ACE2 binding. S494 is not part of the binding interface with ACE2, and hence its mutation to proline does not bring any change to the binding affinity. Y453 also does not form any polar contacts with ACE2. Its mutation to phenylalanine is a conservative mutation and thus does not show any difference in binding affinity. Two mutations K417N and E484K show a decrease in ACE2 binding affinity. K417 residue in SARS-CoV-2 RBD has previously been shown to be responsible for increase in binding efficiency of RBD towards ACE2.[45] K417 residue in SARS-CoV-2 RBD helps in bringing a more positive charge on the RBD surface which better interacts with the negatively charged residues on ACE2. Mutation of lysine to an uncharged amino acid asparagine decreases the positive charge on the surface thereby decreasing binding affinity towards ACE2. Similarly, E484 is a negatively charged amino acid interacts with a positively charged amino acid K31 of ACE2. Mutation to lysine, a positively charged amino acid abolishes this interaction, explaining the decrease in the binding affinity. The other single site mutations (N501Y and S477N) and the Brazil triple mutation considered in this study increase the binding affinity. From this data on ACE2 binding, it appears that protein activity is a big factor in driving protein evolution, but it is not the only driving force. RBM mutants are naturally selected even though they do not show increased affinity for ACE2.

Neutral mutations which do not offer any advantage towards virus fitness and even deleterious mutations can get selected and become prevalent in the presence of other selection pressures like escape from human immune responses.[55] Neutralizing antibodies are a key

component of immune response against natural infection by viruses. Neutralizing antibodies can also be administered as recombinant monoclonal antibody therapeutics and as convalescent plasma to provide passive immunity. The circulating neutralizing antibodies thus provide a selection pressure that can drive the evolution of virus variants.[56-58] Variants which can escape the neutralizing antibodies have a selective advantage and can persist and spread to other individuals. The effectiveness with which a variant can escape the neutralizing antibodies and become prevalent depends on the exposure of that variant to neutralizing antibodies. In the course of normal infection, the levels of antibody during the virus transmission phase are negligible. Thus, the transmitting viruses do not have exposure to neutralizing antibodies and are not considered to contribute to spread of variants with antibody escape potential.[59] But reinfection in people with weak immunity increases the chances of exposure of virus to neutralizing antibodies elicited during first infection and contribute to selection of variants with escape potential.[60-62] Similarly, convalescent plasma therapy, monoclonal antibody therapy (especially single antibody) and less immunogenic vaccine candidates can increase the chances of variants coming in contact with neutralizing antibodies and drive the emergence of variants that can escape these antibodies.[63]

In this study, we tested the potential of the most prevalent variants to escape CC12.1 antibody, a naturally elicited human antibody as a response to SARS-CoV-2 infection.[46] It belongs to a class of antibodies that target RBD and is the most abundant class of antibodies naturally produced by humans.[20,47,64] Antibodies belonging to this class are encoded by the VH3-53 gene segment and represents a set of neutralizing antibodies that bind to the RBD epitope that overlaps with ACE2 binding epitope (Fig. 1).[20] These antibodies are characterized by short H3 CDR and can bind the RBD in "up" conformation. Nevertheless, among the RBD targeting antibodies, these are the most abundant and commonly found antibodies and represent the general antibody response to SARS-CoV-2 infection in humans.[47] Our study examined which of the most prevalent variants have the ability to escape natural immune responses, but it should also be considered that the neutralizing antibody response towards an infection is varied with several neutralizing antibodies generated against different epitopes of spike protein.[65,66] Also, the natural antibody responses can vary from individual to individual. Even then, neutralizing antibodies against RBD represents the most potent and widely used countermeasures against COVID-19.[67] The RBD-CC12.1 binding interface is shown in Fig. 7B. Most of the frequently mutating residues we examined are also part of the binding interface, which implies that the mutations are quite likely to impact CC12.1 binding. Our data shows that only 4 out of the 7 mutants (K417N, Y453F, N501Y, and Brazil triple mutant) we examined bind to CC12.1 with weaker affinity (or higher K_d) (Table 4). The single site mutant Y453F showed the weakest binding or the highest potential to escape CC12.1 ScFv. Interestingly, Y453 is not a part of the binding interface with CC12.1, implying that the residues which do not directly interact with its binding partner can also impact binding after mutation. Other 2 mutants- S477N and S494P did not show any difference in their binding affinity to CC12.1. More importantly, E484K binds with a higher affinity (or lower K_d) to CC12.1 (Table 4). These results show that the neutralizing antibody response is one of the driving forces for natural selection of RBD variants and should be considered in closer detail. It should also be considered that results presented here are only representing the escape towards one class of antibody, and other antibodies that are naturally occurring or administered passively would also have an impact on the emergence of these variants. For example, although E484K mutation may not escape CC12.1 (Table 4), it has been shown to escape other class of antibodies.[68] These results also stress on the need to use viral

sequencing to find the variant that has infected a patient and use of those therapeutic antibodies against which the variant does not show escape potential. To achieve this, different therapeutic neutralizing antibodies that target different epitopes of RBD should be developed, and a cocktail antibody drug may work better against new variants.

In **Summary**, we have examined the various factors that might be contributing to the natural selection of most frequent SARS-CoV-2 RBD variants, in particular protein expression, stability, activity in terms binding to ACE2, and antibody escape potential in terms of binding to a human neutralizing antibody. Table 5 summarizes our observations with red color indicating changes that favor natural selection of variants. These results show that multiple factors contribute to the natural selection of variants, and should be considered when evaluating any future variants. For example, the triple mutant (Brazil variant) poses a serious threat with all factors, increased protein expression, increased stability, increased activity and increased antibody escape potential favoring its emergence and persistence. We hope that such studies on factors determining the natural selection of variants will help in designing better countermeasures like new vaccine candidates and antibody therapeutics that work against emerging variants.

References

1 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).

2 Kim, J. M. *et al.* Identification of Coronavirus Isolated from a Patient in Korea with COVID-19. *Osong Public Health Res Perspect* **11**, 3-7, doi:10.24171/j.phrp.2020.11.1.02 (2020).

3 Zhong, N. S. *et al.* Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet* **362**, 1353-1358, doi:10.1016/s0140-6736(03)14630-2 (2003).

4 Ksiazek, T. G. *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953-1966, doi:10.1056/NEJMoa030781 (2003).

5 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814-1820, doi:10.1056/NEJMoa1211721 (2012).

6 Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574, doi:10.1016/S0140-6736(20)30251-8 (2020).

7 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

8 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286, doi:10.1016/j.cell.2020.02.058 (2020).

9 Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569, doi:10.1038/s41564-020-0688-y (2020).

10 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, doi:10.1016/j.cell.2020.02.052 (2020).

11 Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221-224, doi:10.1038/s41586-020-2179-y (2020).
12 Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu Rev Virol* **3**, 237-261, doi:10.1146/annurev-virology-110615-042301 (2016).
13 Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* **7**, 439-450, doi:10.1038/nrmicro2147 (2009).
14 Rausch, J. W., Capoferri, A. A., Katusiime, M. G., Patro, S. C. & Kearney, M. F. Low genetic diversity may be an Achilles heel of SARS-CoV-2. *Proc Natl Acad Sci U S A* **117**, 24614-24616, doi:10.1073/pnas.2017726117 (2020).
15 van Dorp, L. *et al.* Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect Genet Evol* **83**, 104351, doi:10.1016/j.meegid.2020.104351 (2020).
16 Dearlove, B. *et al.* A SARS-CoV-2 vaccine candidate would likely match all currently circulating variants. *Proc Natl Acad Sci U S A* **117**, 23652-23662, doi:10.1073/pnas.2008281117 (2020).
17 Abdool Karim, S. S. & de Oliveira, T. New SARS-CoV-2 Variants - Clinical, Public Health, and Vaccine Implications. *N Engl J Med*, doi:10.1056/NEJMc2100362 (2021).
18 Voloch, C. M. *et al.* Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *J. Virol.*, doi:10.1128/JVI.00119-21 (2021).
19 Li, M., Lou, F. & Fan, H. SARS-CoV-2 variants: a new challenge to convalescent serum and mRNA vaccine neutralization efficiency. *Signal Transduct Target Ther* **6**, 151, doi:10.1038/s41392-021-00592-6 (2021).
20 Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682-687, doi:10.1038/s41586-020-2852-1 (2020).
21 Jiang, S., Hillyer, C. & Du, L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol.* **41**, 355-359, doi:10.1016/j.it.2020.03.007 (2020).
22 Hansen, J. *et al.* Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010-1014, doi:10.1126/science.abd0827 (2020).
23 Gottlieb, R. L. *et al.* Effect of Bamlanivimab as Monotherapy or in Combination With Etesevimab on Viral Load in Patients With Mild to Moderate COVID-19: A Randomized Clinical Trial. *JAMA* **325**, 632-644, doi:10.1001/jama.2021.0202 (2021).
24 Jones, B. E. *et al.* The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in non-human primates. *Sci Transl Med*, doi:10.1126/scitranslmed.abf1906 (2021).
25 Baden, L. R. *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N Engl J Med* **384**, 403-416, doi:10.1056/NEJMoa2035389 (2021).
26 Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med* **383**, 2603-2615, doi:10.1056/NEJMoa2034577 (2020).
27 Voysey, M. *et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99-111, doi:10.1016/S0140-6736(20)32661-1 (2021).
28 Sadoff, J. *et al.* Interim Results of a Phase 1-2a Trial of Ad26.COV2.S Covid-19 Vaccine. *N Engl J Med*, doi:10.1056/NEJMoa2034201 (2021).

29 Keech, C. *et al.* Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. *N Engl J Med* **383**, 2320-2332, doi:10.1056/NEJMoa2026920 (2020).
30 Zhu, F. C. *et al.* Immunogenicity and safety of a recombinant adenovirus type-5-vectored COVID-19 vaccine in healthy adults aged 18 years or older: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **396**, 479-488, doi:10.1016/S0140-6736(20)31605-6 (2020).
31 Goepfert, P. A. *et al.* Safety and immunogenicity of SARS-CoV-2 recombinant protein vaccine formulations in healthy adults: interim results of a randomised, placebo-controlled, phase 1-2, dose-ranging study. *Lancet Infect. Dis.*, doi:10.1016/S1473-3099(21)00147-X (2021).
32 Liang, J. G. *et al.* S-Trimer, a COVID-19 subunit vaccine candidate, induces protective immunity in nonhuman primates. *Nat Commun* **12**, 1346, doi:10.1038/s41467-021-21634-1 (2021).
33 Chappell, K. J. *et al.* Safety and immunogenicity of an MF59-adjuvanted spike glycoprotein-clamp vaccine for SARS-CoV-2: a randomised, double-blind, placebo-controlled, phase 1 trial. *Lancet Infect. Dis.*, doi:10.1016/S1473-3099(21)00200-0 (2021).
34 Yang, S. *et al.* Safety and immunogenicity of a recombinant tandem-repeat dimeric RBD-based protein subunit vaccine (ZF2001) against COVID-19 in adults: two randomised, double-blind, placebo-controlled, phase 1 and 2 trials. *Lancet Infect. Dis.*, doi:10.1016/S1473-3099(21)00127-4 (2021).
35 Yang, J. *et al.* A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* **586**, 572-577, doi:10.1038/s41586-020-2599-8 (2020).
36 Huang, B. *et al.* Serum sample neutralisation of BBIBP-CorV and ZF2001 vaccines to SARS-CoV-2 501Y.V2. *Lancet Microbe*, doi:10.1016/S2666-5247(21)00082-3 (2021).
37 Madhi, S. A. *et al.* Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N Engl J Med*, doi:10.1056/NEJMoa2102214 (2021).
38 Wang, Z. *et al.* mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616-622, doi:10.1038/s41586-021-03324-6 (2021).
39 Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850-854, doi:10.1126/science.abf9302 (2021).
40 Sun, Z. & Brodsky, J. L. Protein quality control in the secretory pathway. *J. Cell Biol.* **218**, 3171-3187, doi:10.1083/jcb.201906047 (2019).
41 Perlmutter, J. D. & Hagan, M. F. Mechanisms of virus assembly. *Annu Rev Phys Chem* **66**, 217-239, doi:10.1146/annurev-physchem-040214-121637 (2015).
42 Argentinian AntiCovid, C. Structural and functional comparison of SARS-CoV-2-spike receptor binding domain produced in Pichia pastoris and mammalian cells. *Scientific reports* **10**, 21779, doi:10.1038/s41598-020-78711-6 (2020).
43 Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* **46**, W315-W322, doi:10.1093/nar/gky497 (2018).
44 Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A* **103**, 5869-5874, doi:10.1073/pnas.0510098103 (2006).
45 Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220, doi:10.1038/s41586-020-2180-5 (2020).

46 Rogers, T. F. *et al.* Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956-963, doi:10.1126/science.abc7520 (2020).
47 Yuan, M. *et al.* Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119-1123, doi:10.1126/science.abd2321 (2020).
48 Denison, M. R., Graham, R. L., Donaldson, E. F., Eckerle, L. D. & Baric, R. S. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* **8**, 270-279, doi:10.4161/rna.8.2.15013 (2011).
49 Hacisuleyman, E. *et al.* Vaccine Breakthrough Infections with SARS-CoV-2 Variants. *N Engl J Med*, doi:10.1056/NEJMoa2105000 (2021).
50 Bershtein, S., Mu, W., Serohijos, A. W., Zhou, J. & Shakhnovich, E. I. Protein quality control acts on folding intermediates to shape the effects of mutations on organismal fitness. *Mol. Cell* **49**, 133-144, doi:10.1016/j.molcel.2012.11.004 (2013).
51 Dinner, A. R., Abkevich, V., Shakhnovich, E. & Karplus, M. Factors that affect the folding ability of proteins. *Proteins* **35**, 34-40, doi:10.1002/(sici)1097-0134(19990401)35:1<34::aid-prot4>3.0.co;2-q (1999).
52 Klein, E. Y. *et al.* Stability of the Influenza Virus Hemagglutinin Protein Correlates with Evolutionary Dynamics. *mSphere* **3**, doi:10.1128/mSphereDirect.00554-17 (2018).
53 Elena, S. F., Carrasco, P., Daros, J. A. & Sanjuan, R. Mechanisms of genetic robustness in RNA viruses. *EMBO Rep* **7**, 168-173, doi:10.1038/sj.embor.7400636 (2006).
54 Wylie, C. S. & Shakhnovich, E. I. A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci U S A* **108**, 9916-9921, doi:10.1073/pnas.1017572108 (2011).
55 Weisblum, Y. *et al.* Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* **9**, doi:10.7554/eLife.61312 (2020).
56 Dowd, K. A., Netski, D. M., Wang, X. H., Cox, A. L. & Ray, S. C. Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with hepatitis C virus. *Gastroenterology* **136**, 2377-2386, doi:10.1053/j.gastro.2009.02.080 (2009).
57 Richman, D. D., Wrin, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* **100**, 4144-4149, doi:10.1073/pnas.0630530100 (2003).
58 Petrova, V. N. & Russell, C. A. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* **16**, 47-60, doi:10.1038/nrmicro.2017.118 (2018).
59 Wolfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465-469, doi:10.1038/s41586-020-2196-x (2020).
60 Van Elslande, J. *et al.* Symptomatic SARS-CoV-2 reinfection by a phylogenetically distinct strain. *Clin. Infect. Dis.*, doi:10.1093/cid/ciaa1330 (2020).
61 Larson, D. *et al.* A Case of Early Re-infection with SARS-CoV-2. *Clin. Infect. Dis.*, doi:10.1093/cid/ciaa1436 (2020).
62 To, K. K. *et al.* COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. *Clin. Infect. Dis.*, doi:10.1093/cid/ciaa1275 (2020).
63 Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014-1018, doi:10.1126/science.abd0831 (2020).

64 Fagiani, F., Catanzaro, M. & Lanni, C. Molecular features of IGHV3-53-encoded antibodies elicited by SARS-CoV-2. *Signal Transduct Target Ther* **5**, 170, doi:10.1038/s41392-020-00287-4 (2020).
65 Luchsinger, L. L. *et al.* Serological Assays Estimate Highly Variable SARS-CoV-2 Neutralizing Antibody Activity in Recovered COVID-19 Patients. *J. Clin. Microbiol.* **58**, doi:10.1128/JCM.02005-20 (2020).
66 Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442, doi:10.1038/s41586-020-2456-9 (2020).
67 DeFrancesco, L. COVID-19 antibodies on trial. *Nat. Biotechnol.* **38**, 1242-1252, doi:10.1038/s41587-020-0732-8 (2020).
68 Starr, T. N., Greaney, A. J., Dingens, A. S. & Bloom, J. D. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep Med* **2**, 100255, doi:10.1016/j.xcrm.2021.100255 (2021).
69 Bevington, P. R. & Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*. 3rd edn, (McGraw-Hill, 2003).

Figure legends

Figure 1. Structures of SARS-CoV-2 RBD (colored blue) interacting with (A) ACE2 (colored gray; PDB ID 6m0j) and (B) CC12.1 Fab (colored green; PDB ID 6xc2), showing the most frequently mutating residues in RBD- K417, N439, Y453, S477, T478, E484, S494 and N501 (colored red). RBM is shown in yellow color. The single mutants of RBD used in this study were K417N, N439K, Y453F, S477N, T478I, E484K, S494P and N501Y. A double mutant (E484K and N501Y) and triple mutants corresponding to South Africa variant (K417N, E484K and N501Y) and Brazil variant (K417T, E484K and N501Y) were also used. The position of Y453 is not visible in the surface view of RBD interacting with CC12.1 Fab as it is buried at the interface.

Figure 2. Comparison of relative expression of RBD and its mutants. (A) SDS-PAGE showing relative amounts of expressed RBD and its mutants. M represents molecular weight markers (From top to bottom: 180, 130, 100, 70, 55, 35 and 25 kDa) (B) Relative expression of mutants quantified from the band intensities in SDS-PAGE in panel A.

Figure 3. (A) Purified RBD and its mutants (B) ACE2 and CC12.1 ScFv. M represents molecular weight marker (From top to bottom: 180, 130, 100, 70, 55, 35, 25, 15 and 10 kDa). (C) Comparison of secondary structures of RBD and its mutants using far-UV CD spectroscopy.

Figure 4. Thermal denaturation melts of RBD and its mutants obtained using far-UV CD spectroscopy. The solid lines show the fits to a 2-state unfolding equation (Eq. 1 in Methods section). Table 2 lists the T_m (midpoint meting temperature) and the ΔH (enthalpy change at Tm) values of RBD variants.

Figure 5. Binding of RBD and its variants to ACE2 studied using ITC. Table 3 lists the interaction parameters from the data fit.

Figure 6. Binding of RBD and its variants to CC12.1 ScFv studied using ITC. Table 4 lists the interaction parameters from the data fit.

Figure 7. A model of SARS-CoV-2 RBD colored blue interacting with (A) ACE2 colored gray and (B) CC12.1 Fab colored green, showing the interface residues of RBD in yellow along with the side chains. The frequently mutating residues that are part of binding interface are colored red and are also shown with their side chains.

Table 1. Proportion of the secondary structural elements calculated by deconvolution of far-UV CD spectra of RBD and its mutants using BeStSel software.

Secondary structure	**Wild-type**	**K417N**	**Y453F**	**S477N**	**E484K**	**S494P**	**N501Y**	**K417T E484K N501Y**
α-Helix	13.7	12.8	15.3	13.5	15.7	14.1	11.4	10.4
Antiparallel β-sheet	25.0	17.9	19.2	22	16.3	21.4	22.9	24.3
Parallel β-sheet	0.0	4.4	1.5	2	4	2.1	2.8	2.9
Turn	13.9	14.1	13.9	13.2	13.4	13.3	13	13.1
Others	47.3	50.7	50.2	49.3	50.6	49.1	49.9	49.3

Table 2. T_m and ΔH values of RBD variants obtained from thermal denaturation curves using far-UV CD. Errors on ΔT_m were calculated using error propagation formulae.[69]

RBD variant	**Tm (°C)**	**ΔT_m (°C)**	**ΔH (kcal/mol)**
Wild-type	55.5 ± 1.0	-	-1.8 ± 0.4
K417N	55.4 ± 0.6	-0.1 ± 1.2	-2.6 ± 0.6
Y453F	57.9 ± 0.4	2.4 ± 1.1	-3.8 ± 0.6
S477N	56.6 ± 0.4	1.1 ± 1.1	-2.1 ± 0.2
E484K	54.7 ± 0.8	-0.8 ± 1.3	-1.8 ± 0.3
S494P	57.8 ± 1.0	2.3 ± 1.4	-1.5 ± 0.2
N501Y	57.1 ± 0.5	1.6 ± 1.1	-2.1 ± 0.3
K417T, E484K, N501Y	58.2 ± 0.7	2.7 ± 1.2	-1.5 ± 0.2

Table 3. Interaction parameters obtained from binding of RBD variants to ACE2 probed by ITC. Errors on ΔG and -TΔS were calculated using error propagation formulae.[69]

RBD variant	K_d (nM)	N	ΔH (kcal/mol)	ΔG (kcal/mol)	-TΔS (kcal/mol)
Wild-type	10.0 ± 3.1	1.0 ± 0.0	-11.8 ± 0.2	-10.7 ± 0.2	1.0 ± 0.3
K417N	94.1 ± 9.5	0.9 ± 0.0	-11.7 ± 0.1	-9.4 ± 0.1	2.3 ± 0.1
Y453F	11.4 ± 1.7	1.0 ± 0.0	-11.8 ± 0.1	-10.7 ± 0.1	1.2 ± 0.1
S477N	5.0 ± 1.4	1.0 ± 0.0	-16.4 ± 0.2	-11.1 ± 0.2	5.3 ± 0.3
E484K	51.5 ± 8.2	1.0 ± 0.0	-9.5 ± 0.2	-9.8 ± 0.1	-0.3 ± 0.2
S494P	13.8 ± 3.5	1.0 ± 0.0	-10.7 ± 0.2	-10.5 ± 0.1	0.2 ± 0.2
N501Y	3.0 ± 2.1	0.9 ± 0.0	-11.7 ± 0.2	-11.4 ± 0.4	0.2 ± 0.5
K417T, E484K, N501Y	1.6 ± 1.5	0.9 ± 0.0	-10.9 ± 0.2	-11.8 ± 0.5	-0.9 ± 0.6

Table 4. Interaction parameters obtained from binding of RBD variants to CC12.1 ScFv probed by ITC. Errors on ΔG and -TΔS were calculated using error propagation formulae.[69]

RBD variant	K_d (nM)	N	ΔH (kcal/mol)	ΔG (kcal/mol)	-TΔS (kcal/mol)
Wild-type	28.0 ± 8.8	0.9 ± 0.0	-4.8 ± 0.1	-10.1 ± 0.2	-5.3 ± 0.2
K417N	119 ± 50	0.8 ± 0.0	-1.8 ± 0.1	-9.3 ± 0.2	-7.5 ± 0.3
Y453F	827 ± 146	1.0 ± 0.0	-2.2 ± 0.1	-8.2 ± 0.1	-6.0 ± 0.1
S477N	25.6 ± 6.8	0.8 ± 0.0	-5.7 ± 0.1	-10.2 ± 0.2	-4.5 ± 0.2
E484K	1.7 ± 4.6	0.7 ± 0.0	-5.5 ± 0.1	-11.8 ± 1.6	-6.2 ± 1.2
S494P	23.0 ± 7.7	0.8 ± 0.0	-6.4 ± 0.1	-10.2 ± 0.2	-3.9 ± 0.2
N501Y	62.7 ± 21.9	0.9 ± 0.0	-6.2 ± 0.2	-9.7 ± 0.2	-3.4 ± 0.3
K417T, E484K, N501Y	433 ± 95	1.0 ± 0.0	-5.6 ± 0.2	-8.5 ± 0.3	-3.0 ± 0.2

Table 5. Comparison of protein expression, stability, binding to ACE2, and binding to CC12.1 of RBD variants. Red colored values indicate parameters favoring the natural selection of variants.

RBD variant	**Expression**	**T_m (°C)**	**K_d (nM) (ACE2 binding)**	**K_d (nM) (CC12.1 binding)**
Wild-type	-	55.5 ± 1.0	10.0 ± 3.1	28.0 ± 8.8
K417N	No change	55.4 ± 0.6	94.1 ± 9.5	119 ± 50.4
Y453F	Increase	57.9 ± 0.4	11.4 ± 1.7	827 ± 146
S477N	Increase	56.6 ± 0.4	5.0 ± 1.4	25.6 ± 6.8
E484K	Decrease	54.7 ± 0.8	51.5 ± 8.2	1.7 ± 4.6
S494P	Increase	57.8 ± 1.0	13.8 ± 3.5	23.0 ± 7.7
N501Y	No Change	57.1 ± 0.5	3.0 ± 2.1	62.7 ± 21.9
K417T, E484K, N501Y	Increase	58.2 ± 0.7	1.6 ± 1.5	433 ± 94.9

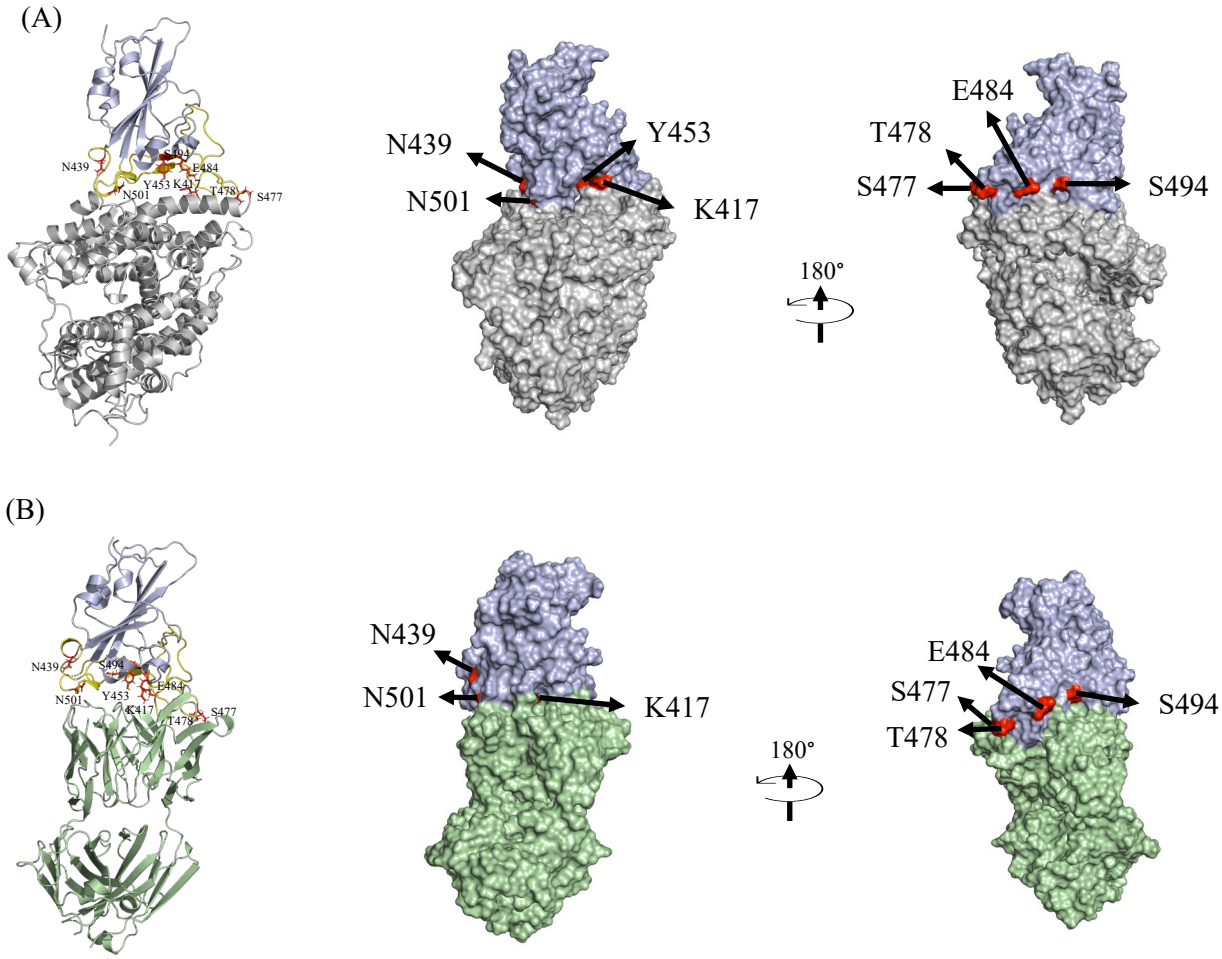

Figure 1. Structures of SARS-CoV-2 RBD (colored blue) interacting with (A) ACE2 (colored gray; PDB ID 6m0j) and (B) CC12.1 Fab (colored green; PDB ID 6xc2), showing the most frequently mutating residues in RBD- K417, N439, Y453, S477, T478, E484, S494 and N501 (colored red). RBM is shown in yellow color. The single mutants of RBD used in this study were K417N, N439K, Y453F, S477N, T478I, E484K, S494P and N501Y. A double mutant (E484K and N501Y) and triple mutants corresponding to South Africa variant (K417N, E484K and N501Y) and Brazil variant (K417T, E484K and N501Y) were also used. The position of Y453 is not visible in the surface view of RBD interacting with CC12.1 Fab as it is buried at the interface.

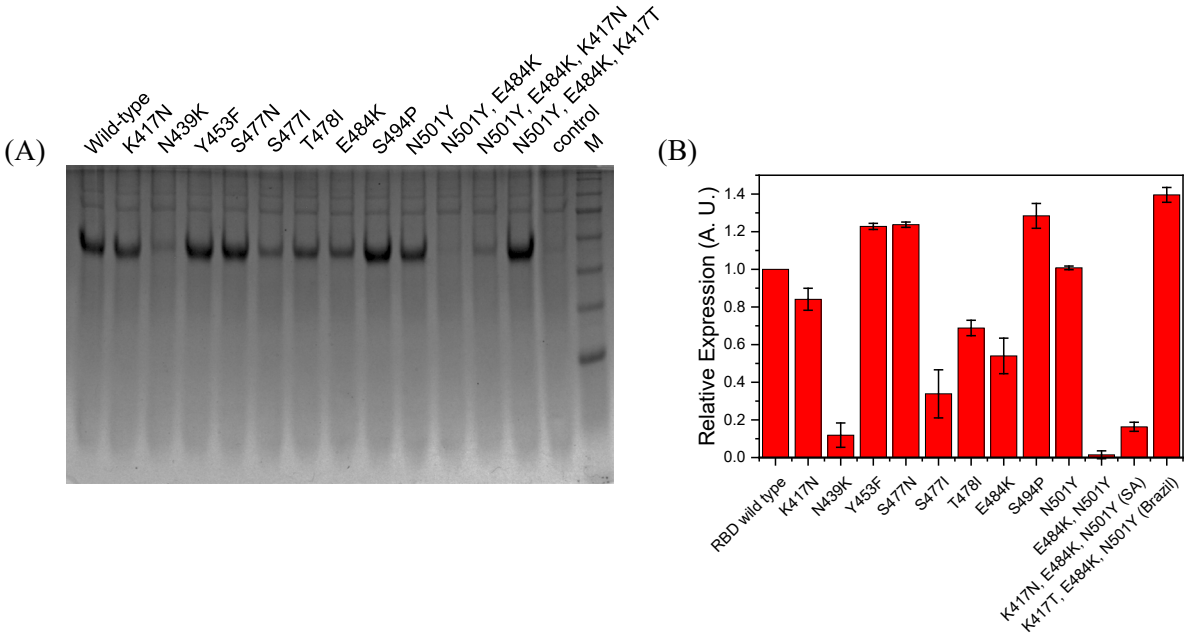

Figure 2. Comparison of relative expression of RBD and its mutants. (A) SDS-PAGE showing relative amounts of expressed RBD and its mutants. M represents molecular weight markers (From top to bottom: 180, 130, 100, 70, 55, 35 and 25 kDa) (B) Relative expression of mutants quantified from the band intensities in SDS-PAGE in panel A.

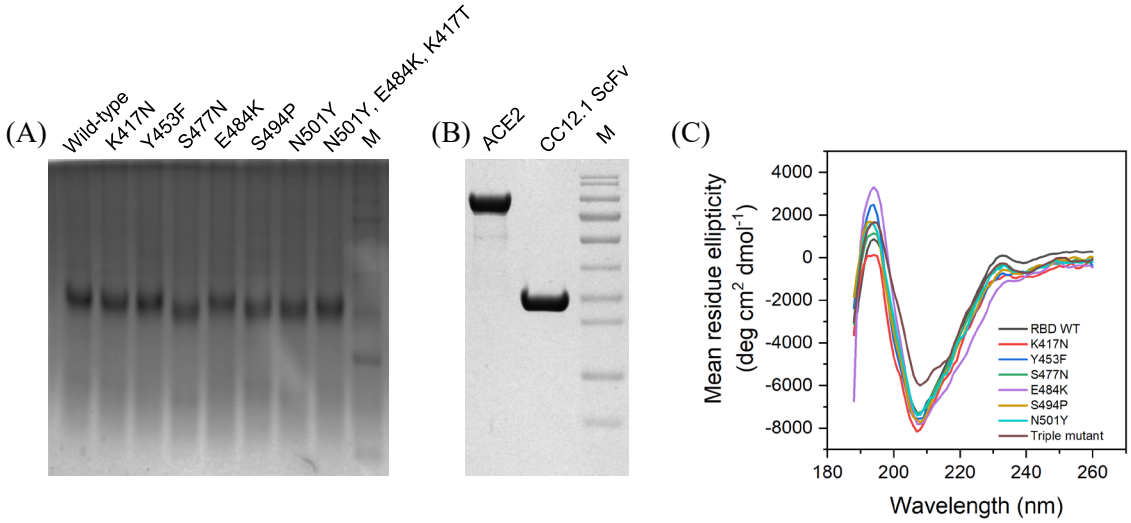

Figure 3. (A) Purified RBD and its mutants (B) ACE2 and CC12.1 ScFv. M represents molecular weight marker (From top to bottom: 180, 130, 100, 70, 55, 35, 25, 15 and 10 kDa). (C) Comparison of secondary structures of RBD and its mutants using far-UV CD spectroscopy.

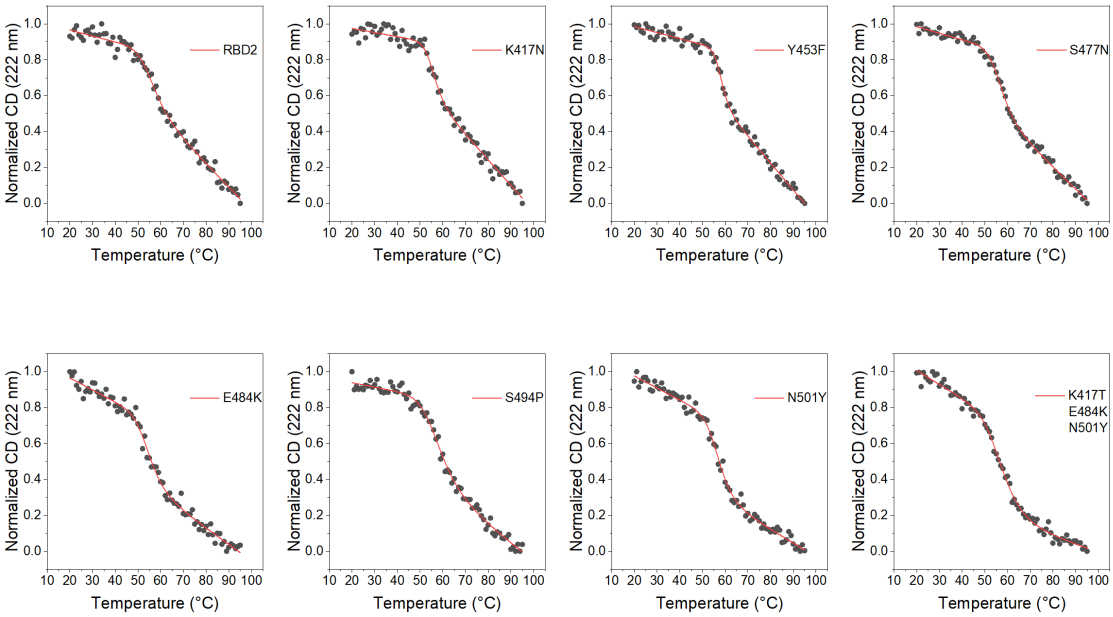

Figure 4. Thermal denaturation melts of RBD and its mutants obtained using far-UV CD spectroscopy. The solid lines show the fits to a 2-state unfolding equation (Eq. 1 in Methods section). Table 1 lists the T_m (midpoint meting temperature) and the ΔH (enthalpy change at Tm) values of RBD variants

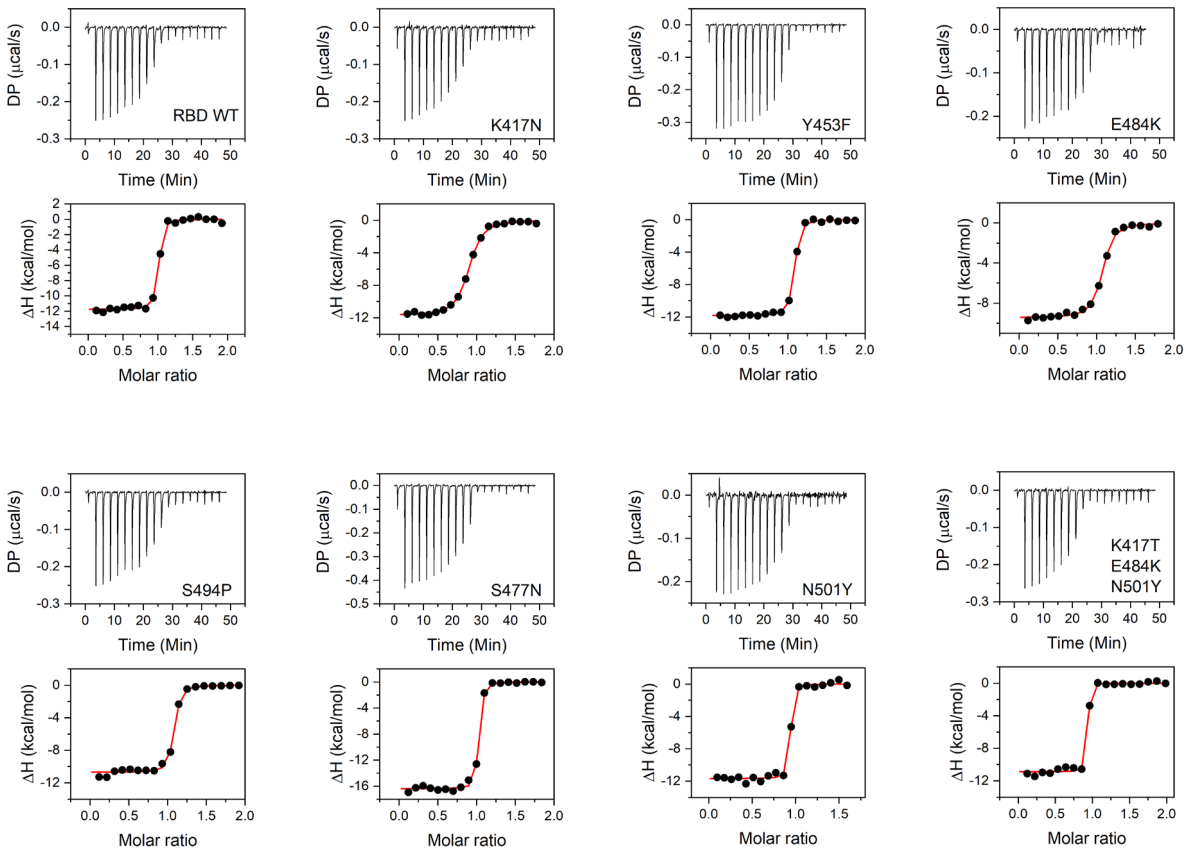

Figure 5. Binding of RBD and its variants to ACE2 studied using ITC. Table 2 lists the interaction parameters from the data fit.

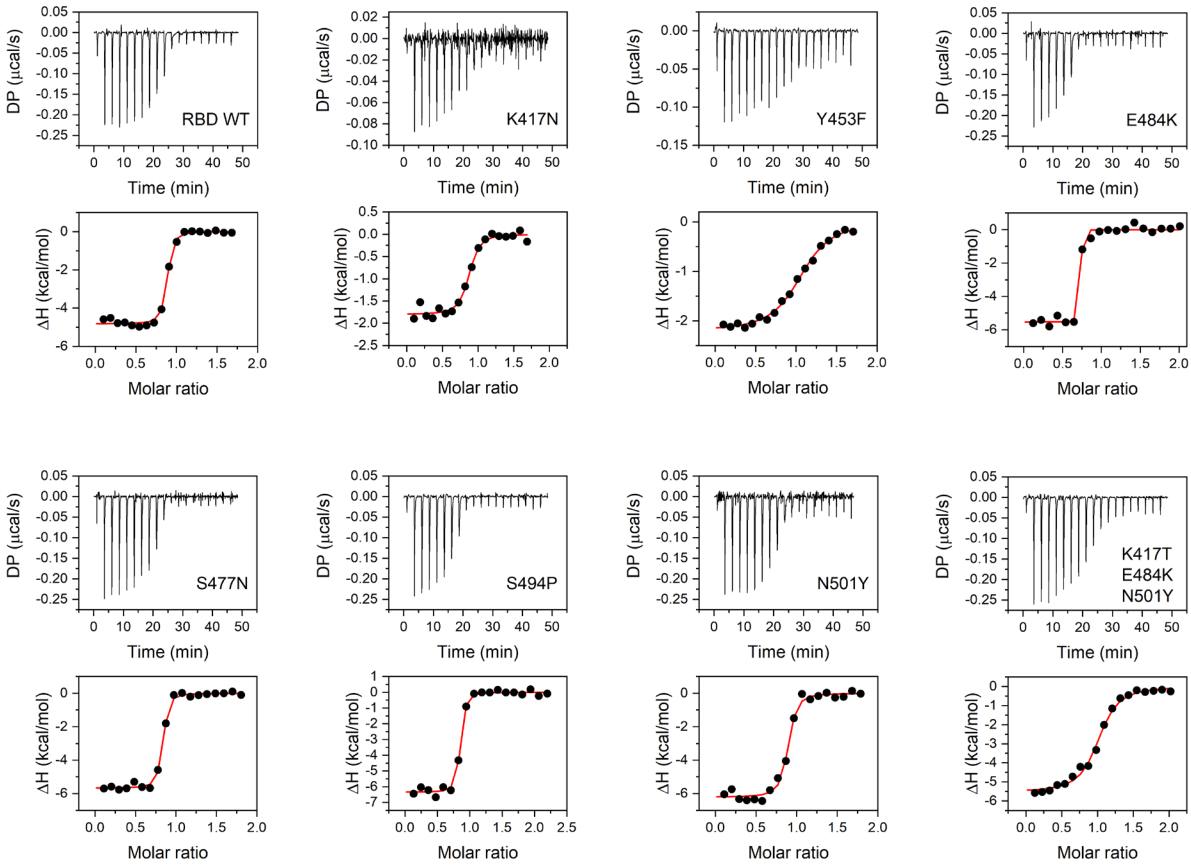

Figure 6. Binding of RBD and its variants to CC12.1 ScFv studied using ITC. Table 3 lists the interaction parameters from the data fit.

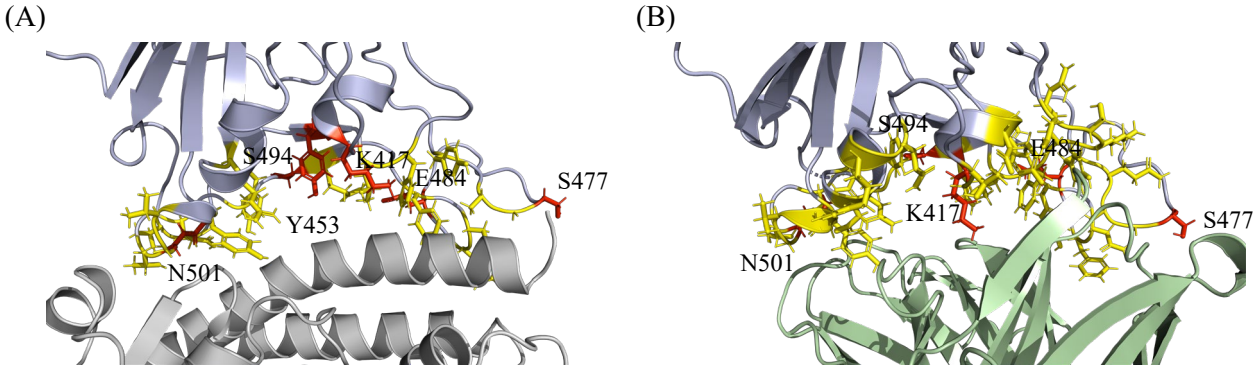

Figure 7. A model of SARS-CoV-2 RBD colored blue interacting with (A) ACE2 colored gray and (B) CC12.1 Fab colored green, showing the interface residues of RBD in yellow along with the side chains. The frequently mutating residues that are part of binding interface are colored red and are also shown with their side chains.