FRONT MATTER

Title

- Increased elastase sensitivity and decreased intramolecular interactions in the more transmissible SARS-CoV-2 variants' spike protein
- Analysis of the new UK and SA SARS-CoV-2 variants

Authors

S. Pokhrel[1]†, L. Lee[1]†, B.R. Kraemer[1]†, K. Samardzic[1], D. Mochly-Rosen[1*]

Affiliations

1. Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA
† These authors contributed equally to this work
* mochly@stanford.edu

Abstract

Two SARS-CoV-2 variants showing increased transmissibility relative to the Wuhan virus have recently been identified. Although neither variant causes more severe illness or increased risk of death, the faster spread of the virus is a major threat. Using computational tools, we found that the new SARS-CoV-2 variants may acquire an increased transmissibility by increasing the propensity of its spike protein to expose the receptor binding domain. This information leads to the identification of potential treatments to avert the imminent threat of these more transmittable SARS-CoV-2 variants.

Teaser

The more infective SARS-CoV-2 variants may expose its Achilles Heel – an opportunity to reduce their spreading.

Abbreviations

ACE2	Angiotensin converting enzyme 2
MOE	Molecular Operating Environment
PDB	Protein Data Bank
RBD	Receptor binding domain

Keywords

SARS-CoV-2, N501.V1, N501.V2, neutrophil elastase, conformation, neutralization, spike protein

MAIN TEXT
Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the novel coronavirus that has resulted in the coronavirus disease 2019 (COVID-19) pandemic, continues to mutate while spreading throughout the world. A critical protein on the surface of the virus is the spike protein, that mediates the entry of the virus into the host cells (*1*). When a mutation or set of mutations provide an advantage over the previous variants, the new variant becomes the dominant variant to spread. This is what occurred when the aspartate 614 in the spike protein of the initial (Wuhan) SARS-CoV-2 virus, mutated to a glycine (Asn614Gly); within a month, the new variant became the most dominant in Europe (*2*) and now globally (*3*). The Asn614Gly variant is associated with higher viral loads, but not with a more severe COVID-19 symptoms in infected individuals (*4*). Importantly, the Asn614Gly mutation does not affect the ability of antibodies to neutralize that variant (*5*). Nevertheless, as the virus continues to mutate at a very high rate, [for example, each amino acid in the 1273 amino acid of the viral spike protein has mutated on average 4 times since the protein was first sequenced, about a year ago (*6*)], the need to continue active surveillance for the emergence of new variants and examination of means to slow down the spread of the new variants is of utmost importance.

The SARS-CoV-2 United Kingdom variant, 501Y.V1, and the South African variant, 501Y.V2

A novel SARS-CoV-2 variant, known as B.1.1.7 or 20B/501Y.V1, emerged in the United Kingdom at the end of September 2020 and became the dominant variant within a month (*7*). This variant increases viral transmissibility between 40-70% (*8*). As of January 16 2021, 501Y.V1 has been reported in more than 50 countries (*9*). However, this new variant does not appear to increase the severity of COVID-19 (*10*). The variant has concomitant 3 deletions and 10 amino acid changes in the 1273 amino acid spike protein (Table 1a), compared with the initial SARS-CoV-2 index virus identified in Wuhan, China; only one (Asn501Tyr or N501Y) is in the angiotensin-converting enzyme 2 (ACE2) receptor-binding domain (RBD; amino acids 331-524) (*11*). Since this Asn501Tyr mutation in the spike's RBD was observed alone as early as April 2020 in Brazil and later in Australia without reports of increased transmissibility, (*12*), it is unlikely that this single substitution is sufficient to explain the new phenotype of B.1.1.7 variant.

De Olivera and collaborators identified another more transmittable and dominant variant, termed 501Y.V2 (aka 20H/501Y.V2 or B.1.351) South African variant (*13*). Identified first in the second week of October, this variant became dominant in South Africa within a month. Six fixed substitutions in all the South African variants were identified (Table 1b): Asn501Tyr [identical to the 501Y.V1 UK variant (*11*)] and Pro614Gly [identical to the European dominant variant, identified between March and April of 2020 (*2*)], the Asp80Ala, Lys417Asn, Glu484Lys and Ala701Val. It appears that the combination of these substitutions results in increased infectivity without increasing COVID-19 severity (*13*).

Here we focused on mutations in the SARS-CoV-2 viral surface spike protein to examine how the new variant became more infective. We used computational methods, including Molecular Operating Environment (MOE) analysis (*14*) and software to predict the outcome of substitutions on the protein structure, to examine the features acquired by the new variants that enable them to increase the rate of infection and spreading without increasing the severity of COVID-19, the pathology resulting from the infection. The goal of this study is to determine whether the acquisition of the more infective phenotype (resulting in a more efficient and rapid viral transmission) exposes new vulnerabilities to target these variants with current drugs.

Results

Potential acquired features of the SARS-CoV-2 501Y.V1 and SARS-CoV-2 501Y.V2 variants

The SARS-CoV-2 spike protein is a trimer protein on the surface of the virus. This protein mediates the binding to the human ACE2 receptor on the nasal mucosa (*15*). Each monomer in the trimeric spike protein is found in two confirmations: a closed conformation, in which the ACE2 RBD is folded in and is inaccessible, and the open conformation – induced by proteolysis of the spike protein, in which the RBD becomes exposed (*16*). When examining the location of the various mutations associated with increased transmissibility of 501Y.V1 and 501Y.V2, we found that these mutations are not located in a particular domain in the spike protein, nor are they all exposed in the close and or the open conformation (Fig. 1a-c).

As neither variant with increased transmissibility causes more severe COVID-19 disease (*3, 10*), and the mutations do not cluster to a particular site in the spike protein, we next examined the possibility that the acquired features in the new variants may lie in increased susceptibility of the spike protein to activation due to introduction of additional protease-activating sites. We hypothesized that the proteases responsible for this gain-of-function should be relatively unique to the initial site of SARS-CoV-2 entry, the nasal mucosa, as viral effects on internal organs do not seem to become more severe. At least seven ubiquitously expressed proteases have been suggested to act on the spike protein: TMPRSS2, furin, PC1, matriptase, trypsin, cathepsin L and/or cathepsin B (*17, 18*).

In contrast to the above proteases, the neutrophil elastase is enriched in the nose of humans compared to other tissues (*19*) and neutrophils, a source for this elastase, are recruited to the nose in SARS-CoV-2 infection (*20*). Indeed, the now worldwide dominant Asp614Gly variant, introduced a new elastase proteolysis site in the spike protein (*2, 21*). Using ExPASy (*22*), we identified Thr716Ile and Ser982Ala in 501Y.V1 as potential new elastase cleavage sites (Fig. 2a). However, whereas Thr716 is on the surface of the spike protein, Ser982 is barely exposed in the open (active) conformation and is buried in the closed conformation (Fig. 1a). We also identified new potential elastase cleavage sites in 501Y.V2: two of the substitutions are fixed in all variants of the new South African strain (Asp80Ala and Ala701Val) and the other two are not (Asp215Gly and Arg246Ile; Fig. 2b). Note that although proteolysis-induced activation was first mapped to the S1/S2 boundary in the spike protein (*23*), protease-induced cuts at other sites may increase the propensity of the spike protein to expose the RBD.

The new variants may decrease intramolecular interactions in the spike protein

We next examined a complementary possible mechanism for the new mutations to compete for the older viral variants, namely substitutions that may decrease intramolecular interactions in the spike protein, thus rendering it more likely to be in the open/active conformation. The open confirmation in this trimeric protein exposes the RBD in a monomer, thus enabling it to bind to ACE2 and enter the body. Indeed, previous work suggested that the Asp614Gly substitution is more prone to be in the open conformation (*24*).

Using MOE, we found that Asn501Tyr substitution in both 501Y.V1 in 501Y.V2 may increase the propensity of the spike protein to be in the 'open' or active conformation (Fig. 3a).

The backbones of Asn501 residues in the three monomers are about 14 Å from each other in the closed trimeric structure. Therefore, a substitution with a bulky tyrosine residue (~7 Å, each) may introduce clashes and cause a loosening of the contact sites between the monomers in the closed conformation (Fig. 3a). Note, however, MOE calculations suggest that 501Tyr substitution does not significantly destabilize the closed conformation (ΔStability 0.4369 kcal/mol) as compared with the index Wuhan 501N variant.

Another intramolecular interaction in 501Y.V1 that can affect transition from closed to open conformer is mediated by serine 982. Ser982 in one monomer forms a hydrogen bond with a Thr547 in an adjacent monomer in the closed conformation (Fig. 3b). The substitution serine to alanine in 501Y.V1 abolishes this interaction with Thr547 in the neighboring spike monomer, thus likely favoring the open conformation.

In addition to the Asn501Tyr, the South African variant, 501Y.V2, has another amino acid substitution that may favor open conformation of spike protein: Lys417Asn (Table 1b). In the closed conformation, Lys417 present in one monomer makes a proton-π interaction with Tyr369 present in the adjacent monomer (Fig. 3c). Asparagine substitution at this position may disrupt this interaction with Tyr369, potentially favoring the transition to open conformation.

Note also that Thr716Ile, Ser982Ala, and Asp1118His are also conserved in viruses isolated from bat, civet and pangolin, suggesting that these amino acid substitutions may be important for some function of the spike protein (Fig. 4). Similarly, Asp80Ala and Ala701Val amino acid that are mutated in 501Y.V2 are conserved in these other species (Fig. 4). Whether these mutations are associated with more efficient transmissibility in various species remain to be determined.

SARS-CoV-2 variants in the RBD

The Ala501Tyr mutation common to both 501Y.V1 and 501Y.V2 variants and Glu484Lys and Lys417Asn of the 501Y.V2 variant is among the amino acids shown to interact with ACE2 (25). It is therefore possible that these mutations increase the binding of the active conformer of the spike protein to ACE2 or to other associated receptors to mediate more effective transmissibility. Although the affinity calculations in MOE suggest that these mutations have minimal negative impact on the affinity of RBD with ACE2 (ΔAffinity; Table 1c), recent yeast display assay suggests that co-occurrence of these mutations increases the affinity of RBD to ACE2 by 2 and 10 folds for 501Y.V1 and 501Y.V2, respectively (26). However, we suggest that mutations in RBD may also affect the closed conformation's stability, thus increasing the probability of binding to ACE2; the monomers wrap around one another through RBD in the closed trimeric structure and mutations in RBD that abolish these critical inter-monomeric interactions could favor the transition to open conformation. Therefore, the effect of these mutations, alone or in combination, on RBD-ACE2 affinity should be tested using intact viral particles.

Discussion

Mutations and the impact on SARS-CoV-2 neutralizing antibodies

The emergence of new virus variants triggers the concern that previous exposure to the virus, antibody therapeutics and current vaccines against SARS-CoV-2 may not be effective. This public health concern needs to be addressed as soon as possible. Many neutralizing antibodies were mapped to the RBD, and four major classes of neutralizing sites have been identified in this

domain (*27, 28*). As one of the mutations in SARS-CoV-2 501Y.V1 and three mutations in 501Y.V2 are in the RBD, these mutations may decrease the neutralizing activities of antibodies (*13*). However, neutralization studies refute this idea; both sera of convalescent patients and immunized subjects appear effective (*29-31*). Notably, mutation E484K in the 501Y.V2 lineage reduces serum binding and neutralization in circulating SARS-CoV-2 isolates (*32*). Furthermore, although point mutations in the spike protein may affect the therapeutic benefit of one or a combination of two monoclonal antibodies (*33*), the benefit of polyclonal antibodies, such as those generated in infected individuals or when using either passive or active vaccines is less likely to be negatively impacted.

The increased transmissibility of the new variants may expose an Achilles Heel

Our analysis indicates that many of these mutations may also increase the exposure of the RBD of the spike protein. We therefore suggest that neutralization by antibodies may increase – more antigenic determinants per each viral particle may be exposed for the antibodies to bind to. Indeed, sera from hamsters infected with Asp614Gly virus exhibit slightly higher neutralization titers against the new dominant Gly614 variant than against the index variant Asp614 (*13*). Furthermore, each viral particle has as many as 80 spike proteins on each virus (*34*), only some of which need to bind ACE2, to induce viral entry into cells. If more spike proteins are activated and expose the antigenic RBD domain, neutralizing antibodies have an increased probability to bind to viral particles and agglutinate several particles together. The resulting agglutinated particles are less likely to have a successful fusion with the cell membrane and infection may be aborted. Increased infection efficiency due to increased propensity of the spike protein activation may also increase the vulnerability of the usually hidden spike protein core to the neutralizing antibodies.

Nasal neutrophil elastase and viral infections – a therapeutic opportunity?

Proteomic analyses identified five proteins with increased expression in nasopharyngeal swab of subjects with high polymerase chain reaction titer for SARS-CoV-2 (*20*). Among these five proteins is neutrophil elastase, whose levels increased 3-fold in SARS-CoV-2 infected *vs.* noninfected individuals. Furthermore, there is a 10-fold increase in neutrophil number in swabs from infected individuals. Elastase was also found to be critical for accelerating viral entry as well as inducing hypertension, thrombosis and vasculitis (*20, 35*). As there are several approved elastase inhibitors, including Sivelestat (which is approved for acute respiratory syndrome), Alvestat (α1-antitrypsin) and Roseltide, the use of these inhibitors, perhaps as intranasal spray or drops, should be considered.

Limitations to our analysis

Our study uses computational tools to examine the potential impact of the mutations found in SARS-CoV-2 501Y.V1 and 501Y.V2 on the viral infectivity has several limitations. We focused only on two aspects – sensitivity to protease and increased open conformer of the spike protein. Other mechanisms may increase viral transmissibility and were not examined here. For example, if proteolytic events contribute to viral inactivation at the mucosa, a decline in proteolysis susceptibility, especially to proteases present in the nose, may also contribute to increased infectivity. We also did not examine the possibility that the increased transmissibility of the new variants may be due to better survival of the virus in respiratory droplets, the potential increased affinity of the viral spike protein for receptors other than ACE2, nor the impact of mutations in other viral genes. Finally, our analysis tested mainly the potential impact of one

mutation at a time. As our analysis is *in silico*, experiments to test the hypotheses raised by this study should be conducted next, to determine if these SARS-CoV-2 variants with more efficient and rapid transmission capabilities are also more sensitive to existing therapies and suggest new therapeutic targets to slow down and arrest the COVID-19 pandemic.

Materials and Methods

Proteolysis Site Prediction Using ExPASy

The amino acid sequence of spike protein from the reference EPI_ISL_402124 (Wuhan) sequence was uploaded to ExPASy (*22*) (https://web.expasy.org/peptide_cutter/) and neutrophil elastase was selected as the protease. We examined also potential proteolysis site for the seven other proteases suggested to cleave SARS-CoV-2 spike proteins (*17*). No changes in cleavage sites for both 501Y.V1 and 501Y.V2 were observed for furin, Cathepsin L/B, and PC1. However, a loss of a cleavage site was observed for matriptase in the 501Y.V1, due to alanine of serine 982 mutation. For TMPRSS2 and trypsin, which cleave at single arginine or lysine residues, a proteolysis site was gained at position 484, and a loss of proteolysis site was observed at both positions 246 and 417 in the 501Y.V2.

Protein Structures

Molecular Operating Environment (MOE) software (version 2019.01) (*14*) was used to generate representative structural models of the open and closed conformations of SARS-CoV-2 variant spike proteins. Protein Data Bank (PDB) ID: 7A98 (*25*) was used to model open conformation whereas PDB ID: 6ZB5(*36*) was used for closed conformation.

Residue Scan and Affinity Calculation

PDB ID: 7A98 was prepared using the QuickPrep functionality at the default settings, to optimize the H-bond network and perform energy minimization on the system in MOE. The ΔAffinity calculations were performed using 7A98.A (S-protein monomer) and 7A98.D (ACE-2) chains with 7A98.D defined as ligand. Residues K417, E484 and N501 in S-protein (7A98.A) were selected and residue scan application was performed. The difference in affinity (ΔAffinity (kcal/mol)) between mutated residue and the wild-type were calculated as per MOE's definition.

Residue Scan and Stability Calculation

QuickPrep functionality was used as above to prepare PDB ID: 6ZB5 in MOE. 6ZB5.A, 6ZB5.B and 6ZB5.C chains were used to perform the calculations. N501 residue in S-protein in all three chains were selected and residue scan application was performed for the observed variants S, Y, T and R that are tolerated at this position with site-limit set to 3. The change in stability (ΔStability; kcal/mol) between the variants and the reference wild-type sequence were calculated as per MOE's definition.

Calculating Predicted Effect of Variants in PROVEAN and SIFT

The amino acid sequence of spike protein from the index EPI_ISL_402124 (Wuhan) sequence was uploaded to PROVEAN (*37*) (http://provean.jcvi.org/index.php) and SIFT (*38*) (https://sift.bii.a-star.edu.sg). Every variant observed in 501Y.V1 (UK variant) and 501Y.V2 (South Africa variant) was also uploaded to compare against the reference sequence. Each variant was either predicted to be 'deleterious' or 'neutral' in PROVEAN or 'deleterious' or 'benign' in SIFT.

Multiple Sequence Alignment

MEGA X software (*39*) and the ClustalW alignment function was used to align the amino acid sequences of the human index, bat, civet, pangolin, 501Y.V1 and 501Y.V2 SARS-CoV-2 spike proteins.

References

1. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural Basis for the Recognition of SARS-CoV-2 by Full-Length Human ACE2. *Science* **367**, 1444 (2020).
2. S. K. Mishra, T. Tripathi, One Year Update on the COVID-19 Pandemic: Where Are We Now? *Acta Trop* **214**, 105778 (2021).
3. World Health Organization, in *Disease Outbreak News*. (WHO, 2020).
4. E. Volz, V. Hill, J. T. McCrone, A. Price, D. Jorgensen, Á. O'Toole, J. Southgate, R. Johnson, B. Jackson, F. F. Nascimento, S. M. Rey, S. M. Nicholls, R. M. Colquhoun, A. da Silva Filipe, J. Shepherd, D. J. Pascall, R. Shah, N. Jesudason, K. Li, R. Jarrett, N. Pacchiarini, M. Bull, L. Geidelberg, I. Siveroni, C. Koshy, E. Wise, N. Cortes, J. Lynch, S. Kidd, M. Mori, D. J. Fairley, T. Curran, J. P. McKenna, H. Adams, C. Fraser, T. Golubchik, D. Bonsall, C. Moore, S. L. Caddy, F. A. Khokhar, M. Wantoch, N. Reynolds, B. Warne, J. Maksimovic, K. Spellman, K. McCluggage, M. John, R. Beer, S. Afifi, S. Morgan, A. Marchbank, A. Price, C. Kitchen, H. Gulliver, I. Merrick, J. Southgate, M. Guest, R. Munn, T. Workman, T. R. Connor, W. Fuller, C. Bresner, L. B. Snell, T. Charalampous, G. Nebbia, R. Batra, J. Edgeworth, S. C. Robson, A. Beckett, K. F. Loveson, D. M. Aanensen, A. P. Underwood, C. A. Yeats, K. Abudahab, B. E. W. Taylor, M. Menegazzo, G. Clark, W. Smith, M. Khakh, V. M. Fleming, M. M. Lister, H. C. Howson-Wells, L. Berry, T. Boswell, A. Joseph, I. Willingham, P. Bird, T. Helmer, K. Fallon, C. Holmes, J. Tang, V. Raviprakash, S. Campbell, N. Sheriff, M. W. Loose, N. Holmes, C. Moore, M. Carlile, V. Wright, F. Sang, J. Debebe, F. Coll, A. W. Signell, G. Betancor, H. D. Wilson, T. Feltwell, C. J. Houldcroft, S. Eldirdiri, A. Kenyon, T. Davis, O. Pybus, L. du Plessis, A. Zarebski, J. Raghwani, M. Kraemer, S. Francois, S. Attwood, T. Vasylyeva, M. E. Torok, W. L. Hamilton, I. G. Goodfellow, G. Hall, A. S. Jahun, Y. Chaudhry, M. Hosmillo, M. L. Pinckert, I. Georgana, A. Yakovleva, L. W. Meredith, S. Moses, H. Lowe, F. Ryan, C. L. Fisher, A. R. Awan, J. Boyes, J. Breuer, K. A. Harris, J. R. Brown, D. Shah, L. Atkinson, J. C. D. Lee, A. Alcolea-Medina, N. Moore, N. Cortes, R. Williams, M. R. Chapman, L. J. Levett, J. Heaney, D. L. Smith, M. Bashton, G. R. Young, J. Allan, J. Loh, P. A. Randell, A. Cox, P. Madona, A. Holmes, F. Bolt, J. Price, S. Mookerjee, A. Rowan, G. P. Taylor, M. Ragonnet-Cronin, F. F. Nascimento, D. Jorgensen, I. Siveroni, R. Johnson, O. Boyd, L. Geidelberg, E. M. Volz, K. Brunker, K. L. Smollett, N. J. Loman, J. Quick, C. McMurray, J. Stockton, S. Nicholls, W. Rowe, R. Poplawski, R. T. Martinez-Nunez, J. Mason, T. I. Robinson, E. O'Toole, J. Watts, C. Breen, A. Cowell, C. Ludden, G. Sluga, N. W. Machin, S. S. Y. Ahmad, R. P. George, F. Halstead, V. Sivaprakasam, E. C. Thomson, J. G. Shepherd, P. Asamaphan, M. O. Niebel, K. K. Li, R. N. Shah, N. G. Jesudason, Y. A. Parr, L. Tong, A. Broos, D. Mair, J. Nichols, S. N. Carmichael, K. Nomikou, E. Aranday-Cortes, N. Johnson, I. Starinskij, A. da Silva Filipe, D. L. Robertson, R. J. Orton, J. Hughes, S. Vattipally, J. B. Singer, A. D. Hale, L. R. Macfarlane-Smith, K. L. Harper, Y. Taha, B. A. I. Payne, S. Burton-Fanning, S. Waugh, J. Collins, G. Eltringham, K. E. Templeton, M. P. McHugh, R. Dewar, E. Wastenge, S. Dervisevic, R. Stanley, R. Prakash, C. Stuart, N. Elumogo, D. K. Sethi, E. J. Meader, L. J. Coupland, W. Potter, C. Graham, E. Barton, D. Padgett, G. Scott, E. Swindells, J. Greenaway, A. Nelson, W. C. Yew, P. C. Resende Silva, M. Andersson, R. Shaw, T. Peto, A. Justice, D. Eyre, D. Crooke, S. Hoosdally, T. J. Sloan, N. Duckworth, S.

Walsh, A. J. Chauhan, S. Glaysher, K. Bicknell, S. Wyllie, E. Butcher, S. Elliott, A. Lloyd, R. Impey, N. Levene, L. Monaghan, D. T. Bradley, E. Allara, C. Pearson, P. Muir, I. B. Vipond, R. Hopes, H. M. Pymont, S. Hutchings, M. D. Curran, S. Parmar, A. Lackenby, T. Mbisa, S. Platt, S. Miah, D. Bibby, C. Manso, J. Hubb, M. Chand, G. Dabrera, M. Ramsay, D. Bradshaw, A. Thornton, R. Myers, U. Schaefer, N. Groves, E. Gallagher, D. Lee, D. Williams, N. Ellaby, I. Harrison, H. Hartman, N. Manesis, V. Patel, C. Bishop, V. Chalker, H. Osman, A. Bosworth, E. Robinson, M. T. G. Holden, S. Shaaban, A. Birchley, A. Adams, A. Davies, A. Gaskin, A. Plimmer, B. Gatica-Wilcox, C. McKerr, C. Moore, C. Williams, D. Heyburn, E. De Lacy, E. Hilvers, F. Downing, G. Shankar, H. Jones, H. Asad, J. Coombes, J. Watkins, J. M. Evans, L. Fina, L. Gifford, L. Gilbert, L. Graham, M. Perry, M. Morgan, M. Bull, M. Cronin, N. Pacchiarini, N. Craine, R. Jones, R. Howe, S. Corden, S. Rey, S. Kumziene-Summerhayes, S. Taylor, S. Cottrell, S. Jones, S. Edwards, J. O'Grady, A. J. Page, J. Wain, M. A. Webber, A. E. Mather, D. J. Baker, S. Rudder, M. Yasir, N. M. Thomson, A. Aydin, A. P. Tedim, G. L. Kay, A. J. Trotter, R. A. J. Gilroy, N.-F. Alikhan, L. de Oliveira Martins, T. Le-Viet, L. Meadows, A. Kolyva, M. Diaz, A. Bell, A. V. Gutierrez, I. G. Charles, E. M. Adriaenssens, R. A. Kingsley, A. Casey, D. A. Simpson, Z. Molnar, T. Thompson, E. Acheson, J. A. H. Masoli, B. A. Knight, A. Hattersley, S. Ellard, C. Auckland, T. W. Mahungu, D. Irish-Tavares, T. Haque, Y. Bourgeois, G. P. Scarlett, D. G. Partridge, M. Raza, C. Evans, K. Johnson, S. Liggett, P. Baker, S. Essex, R. A. Lyons, L. G. Caller, S. Castellano, R. J. Williams, M. Kristiansen, S. Roy, C. A. Williams, P. L. Dyal, H. J. Tutill, Y. N. Panchbhaya, L. M. Forrest, P. Niola, J. Findlay, T. T. Brooks, A. Gavriil, L. Mestek-Boukhibar, S. Weeks, S. Pandey, L. Berry, K. Jones, A. Richter, A. Beggs, C. P. Smith, G. Bucca, A. R. Hesketh, E. M. Harrison, S. J. Peacock, S. Palmer, C. M. Churcher, K. L. Bellis, S. T. Girgis, P. Naydenova, B. Blane, S. Sridhar, C. Ruis, S. Forrest, C. Cormie, H. K. Gill, J. Dias, E. E. Higginson, M. Maes, J. Young, L. M. Kermack, N. F. Hadjirin, D. Aggarwal, L. Griffith, T. Swingler, R. K. Davidson, A. Rambaut, T. Williams, C. E. Balcazar, M. D. Gallagher, Á. O'Toole, S. Rooke, B. Jackson, R. Colquhoun, J. Ashworth, V. Hill, J. T. McCrone, E. Scher, X. Yu, K. A. Williamson, T. D. Stanton, S. L. Michell, C. M. Bewshea, B. Temperton, M. L. Michelsen, J. Warwick-Dugdale, R. Manley, A. Farbos, J. W. Harrison, C. M. Sambles, D. J. Studholme, A. R. Jeffries, A. C. Darby, J. A. Hiscox, S. Paterson, M. Iturriza-Gomara, K. A. Jackson, A. O. Lucaci, E. E. Vamos, M. Hughes, L. Rainbow, R. Eccles, C. Nelson, M. Whitehead, L. Turtle, S. T. Haldenby, R. Gregory, M. Gemmell, D. Kwiatkowski, T. I. de Silva, N. Smith, A. Angyal, B. B. Lindsey, D. C. Groves, L. R. Green, D. Wang, T. M. Freeman, M. D. Parker, A. J. Keeley, P. J. Parsons, R. M. Tucker, R. Brown, M. Wyles, C. Constantinidou, M. Unnikrishnan, S. Ott, J. K. J. Cheng, H. E. Bridgewater, L. R. Frost, G. Taylor-Joyce, R. Stark, L. Baxter, M. T. Alam, P. E. Brown, P. C. McClure, J. G. Chappell, T. Tsoleridis, J. Ball, D. Gramatopoulos, D. Buck, J. A. Todd, A. Green, A. Trebes, G. MacIntyre-Cockett, M. de Cesare, C. Langford, A. Alderton, R. Amato, S. Goncalves, D. K. Jackson, I. Johnston, J. Sillitoe, S. Palmer, M. Lawniczak, M. Berriman, J. Danesh, R. Livett, L. Shirley, B. Farr, M. Quail, S. Thurston, N. Park, E. Betteridge, D. Weldon, S. Goodwin, R. Nelson, C. Beaver, L. Letchford, D. A. Jackson, L. Foulser, L. McMinn, L. Prestwood, S. Kay, L. Kane, M. J. Dorman, I. Martincorena, C. Puethe, J.-P. Keatley, G. Tonkin-Hill, C. Smith, D. Jamrozy, M. A. Beale, M. Patel, C. Ariani, M. Spencer-Chapman, E. Drury, S. Lo, S. Rajatileka, C. Scott, K. James, S. K. Buddenborg, D. J. Berger, G. Patel, M. V. Garcia-Casado, T. Dibling, S. McGuigan, H. A. Rogers, A. D. Hunter, E. Souster, A. S. Neaverson, I. Goodfellow, N. J. Loman, O. G. Pybus, D. L. Robertson, E. C. Thomson, A. Rambaut, T. R. Connor, Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell* **184**, 64-75.e11 (2021).

5. D. Weissman, M.-G. Alameh, T. de Silva, P. Collini, H. Hornsby, R. Brown, C. C. LaBranche, R. J. Edwards, L. Sutherland, S. Santra, K. Mansouri, S. Gobeil, C. McDanal, N. Pardi, N. Hengartner, P. J. C. Lin, Y. Tam, P. A. Shaw, M. G. Lewis, C. Boesler, U. Şahin, P. Acharya, B. F. Haynes, B. Korber, D. C. Montefiori, D614G Spike Mutation Increases SARS CoV-2 Susceptibility to Neutralization. *Cell Host Microbe* **29**, 23-31.e24 (2021).

6. S. Pokhrel, B. R. Kraemer, D. Mochly-Rosen, Natural Variants in SARS-CoV-2 S Protein Pinpoint Structural and Functional Hotspots; Implications for Prophylaxis Strategies. *bioRxiv*, 2021.2001.2004.425340 (2021).

7. E. Volz, S. Mishra, M. Chand, J. C. Barrett, R. Johnson, L. Geidelberg, W. R. Hinsley, D. J. Laydon, G. Dabrera, Á. O'Toole, R. Amato, M. Ragonnet-Cronin, I. Harrison, B. Jackson, C. V. Ariani, O. Boyd, N. J. Loman, J. T. McCrone, S. Gonçalves, D. Jorgensen, R. Myers, V. Hill, D. K. Jackson, K. Gaythorpe, N. Groves, J. Sillitoe, D. P. Kwiatkowski, S. Flaxman, O. Ratmann, S. Bhatt, S. Hopkins, A. Gandy, A. Rambaut, N. M. Ferguson, Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from Linking Epidemiological and Genetic Data. *medRxiv*, 2020.2012.2030.20249034 (2021).

8. World Health Organization, in *Disease Outbreak News*. (WHO, 2020).

9. M. Roberts, Covid: How Worrying Are the UK, South Africa, and Brazil Coronavirus Variants? *BBC News Online* https://www.bbc.com/news/health-55659820 (2021).

10. Centers for Disease Control and Prevention. (CDC, 2020).

11. A. Rambaut, N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D. L. Robertson, E. Volz, "Preliminary Genomic Characterisation of an Emergent SARS-CoV-2 Lineage in the UK Defined by a Novel Set of Spike Mutations" (2020; https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563).

12. J. W. Tang, P. A. Tambyah, D. S. C. Hui, Emergence of a New SARS-CoV-2 Variant in the UK. *J Infect*, (2020).

13. H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E. J. San, N. Msomi, K. Mlisana, A. von Gottberg, S. Walaza, M. Allam, A. Ismail, T. Mohale, A. J. Glass, S. Engelbrecht, G. Van Zyl, W. Preiser, F. Petruccione, A. Sigal, D. Hardie, G. Marais, M. Hsiao, S. Korsman, M.-A. Davies, L. Tyers, I. Mudau, D. York, C. Maslo, D. Goedhals, S. Abrahams, O. Laguda-Akingba, A. Alisoltani-Dehkordi, A. Godzik, C. K. Wibmer, B. T. Sewell, J. Lourenço, L. C. J. Alcantara, S. L. K. Pond, S. Weaver, D. Martin, R. J. Lessells, J. N. Bhiman, C. Williamson, T. de Oliveira, Emergence and Rapid Spread of a New Severe Acute Respiratory Syndrome-Related Coronavirus 2 (SARS-CoV-2) Lineage with Multiple Spike Mutations in South Africa. *medRxiv*, 2020.2012.2021.20248640 (2020).

14. Chemical Computing Group ULC. (Montreal, QC, Canada, 2021).

15. I. Hamming, W. Timens, M. L. Bulthuis, A. T. Lely, G. Navis, H. van Goor, Tissue Distribution of ACE2 Protein, the Functional Receptor for SARS Coronavirus. A First Step in Understanding SARS Pathogenesis. *J Pathol* **203**, 631-637 (2004).

16. X. Xiong, K. Qu, K. A. Ciazynska, M. Hosmillo, A. P. Carter, S. Ebrahimi, Z. Ke, S. H. W. Scheres, L. Bergamaschi, G. L. Grice, Y. Zhang, J. Bradley, P. A. Lyons, K. G. C. Smith, M. Toshner, A. Elmer, C. Ribeiro, J. Kourampa, S. Jose, J. Kennet, J. Rowlands, A. Meadows, C. O'Brien, R. Rastall, C. Crucusio, S. Hewitt, J. Price, J. Calder, L. Canna, A. Bucke, H. Tordesillas, J. Harris, V. Ruffolo, J. Domingo, B. Graves, H. Butcher, D. Caputo, E. Le Gresley, B. J. Dunmore, J. Martin, E. Legchenko, C. Treacy, C. Huang, J. Wood, R. Sutcliffe, J. Hodgson, J. Shih, S. Graf, Z. Tong, F. Mescia, T. Tilly, C. O'Donnell, K. Hunter, L. Pointon, N. Pond, M. Wylot, E. Jones, S. Fawke, B. Bullman, L. Bergamaschi, L. Turner, I. Jarvis, O. Omarjee, A. De Sa, J. Marsden, A. Betancourt, M.

Perera, M. Epping, N. Richoz, G. Bower, R. Sharma, F. Nice, O. Huhn, H. Stark, N. Walker, K. Stirrups, N. Ovington, E. Dewhust, E. Li, S. Papadia, J. A. Nathan, S. Baker, L. C. James, H. E. Baxendale, I. Goodfellow, R. Doffinger, J. A. G. Briggs, C.-N. C.-B. C. The, A Thermostable, Closed SARS-CoV-2 Spike Protein Trimer. *Nat Struct Mol Biol* **27**, 934-941 (2020).

17.   M. Seyran, K. Takayama, V. N. Uversky, K. Lundstrom, G. Palù, S. P. Sherchan, D. Attrish, N. Rezaei, A. A. A. Aljabali, S. Ghosh, D. Pizzol, G. Chauhan, P. Adadi, T. Mohamed Abd El-Aziz, A. G. Soares, R. Kandimalla, M. Tambuwala, S. S. Hassan, G. K. Azad, P. Pal Choudhury, W. Baetas-da-Cruz, Á. Serrano-Aroca, A. M. Brufsky, B. D. Uhal, The Structural Basis of Accelerated Host Cell Entry by SARS-CoV-2. *FEBS J*, (2020).

18.   P. J. Thul, L. Åkesson, M. Wiking, D. Mahdessian, A. Geladaki, H. Ait Blal, T. Alm, A. Asplund, L. Björk, L. M. Breckels, A. Bäckström, F. Danielsson, L. Fagerberg, J. Fall, L. Gatto, C. Gnann, S. Hober, M. Hjelmare, F. Johansson, S. Lee, C. Lindskog, J. Mulder, C. M. Mulvey, P. Nilsson, P. Oksvold, J. Rockberg, R. Schutten, J. M. Schwenk, Å. Sivertsson, E. Sjöstedt, M. Skogs, C. Stadler, D. P. Sullivan, H. Tegel, C. Winsnes, C. Zhang, M. Zwahlen, A. Mardinoglu, F. Pontén, K. von Feilitzen, K. S. Lilley, M. Uhlén, E. Lundberg, A Subcellular Map of the Human Proteome. *Science* **356**, (2017).

19.   P. V. Tomazic, R. Birner-Gruenberger, A. Leitner, B. Obrist, S. Spoerk, D. Lang-Loidolt, Nasal Mucus Proteomic Changes Reflect Altered Immune Responses and Epithelial Permeability in Patients with Allergic Rhinitis. *J Allergy Clin Immunol* **133**, 741-750 (2014).

20.   E. Akgun, M. B. Tuzuner, B. Sahin, M. Kilercik, C. Kulah, H. N. Cakiroglu, M. Serteser, I. Unsal, A. T. Baykal, Proteins Associated with Neutrophil Degranulation Are Upregulated in Nasopharyngeal Swabs from SARS-CoV-2 Patients. *PLOS ONE* **15**, (2020).

21.   R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W. J. Liu, D. Wang, W. Xu, E. C. Holmes, G. F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding. *The Lancet* **395**, 565-574 (2020).

22.   E. Gasteiger, C. Hoogland, A. Gattiker, M. R. Wilkins, R. D. Appel, A. Bairoch, "Protein Identification and Analysis Tools on the Expasy Server" in *The Proteomics Protocols Handbook* (Springer, 2005), pp. 571-607.

23.   M. Örd, I. Faustova, M. Loog, The Sequence at Spike S1/S2 Site Enables Cleavage by Furin and Phospho-Regulation in SARS-Cov2 but Not in SARS-Cov1 or MERS-CoV. *Sci Rep* **10**, 16944 (2020).

24.   J. A. Plante, Y. Liu, J. Liu, H. Xia, B. A. Johnson, K. G. Lokugamage, X. Zhang, A. E. Muruato, J. Zou, C. R. Fontes-Garfias, D. Mirchandani, D. Scharton, J. P. Bilello, Z. Ku, Z. An, B. Kalveram, A. N. Freiberg, V. D. Menachery, X. Xie, K. S. Plante, S. C. Weaver, P.-Y. Shi, Spike Mutation D614G Alters SARS-CoV-2 Fitness. *Nature*, (2020).

25.   D. J. Benton, A. G. Wrobel, P. Xu, C. Roustan, S. R. Martin, P. B. Rosenthal, J. J. Skehel, S. J. Gamblin, Receptor Binding and Priming of the Spike Protein of SARS-CoV-2 for Membrane Fusion. *Nature* **588**, 327-330 (2020).

26.   J. Zahradník, S. Marciano, M. Shemesh, E. Zoler, J. Chiaravalli, B. Meyer, O. Dym, N. Elad, G. Schreiber, SARS-CoV-2 RBD in Vitro Evolution Follows Contagious Mutation Spread, yet Generates an Able Infection Inhibitor. *bioRxiv*, 2021.2001.2006.425392 (2021).

27.   L. Liu, P. Wang, M. S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J. F. W. Chan, V. Sahi, A. Figueroa, X. V. Guo, G. Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.-Y. Yuen, P. D. Kwong, J. G. Sodroski, M. T. Yin, Z. Sheng, Y. Huang, L. Shapiro, D. D. Ho, Potent Neutralizing Antibodies against Multiple Epitopes on SARS-CoV-2 Spike. *Nature* **584**, 450-456 (2020).

28.   C. O. Barnes, C. A. Jette, M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, H. B. Gristick, A. G. Malyutin, N. G. Sharaf, K. E. Huey-Tubman, Y. E. Lee, D. F. Robbiani, M. C. Nussenzweig, A. P. West, P. J. Bjorkman, SARS-CoV-2 Neutralizing Antibody Structures Inform Therapeutic Strategies. *Nature* **588**, 682-687 (2020).

29.   X. Zhu, D. Mannar, S. S. Srivastava, A. M. Berezuk, J. P. Demers, J. W. Saville, K. Leopold, W. Li, D. S. Dimitrov, K. S. Tuttle, S. Zhou, S. Chittori, S. Subramaniam, Cryo-Em Structure of the N501Y SARS-CoV-2 Spike Protein in Complex with a Potent Neutralizing Antibody. *bioRxiv*, 2021.2001.2011.426269 (2021).

30.   X. Xie, J. Zou, C. R. Fontes-Garfias, H. Xia, K. A. Swanson, M. Cutler, D. Cooper, V. D. Menachery, S. Weaver, P. R. Dormitzer, P.-Y. Shi, Neutralization of N501Y Mutant SARS-CoV-2 by BNT162b2 Vaccine-Elicited Sera. *bioRxiv*, (2021).

31.   E. Callway, Could New Covid Variants Undermine Vaccines? Labs Scamble to Find Out. *Nature* **589**, (2021).

32.   A. J. Greaney, A. N. Loes, K. H. D. Crawford, T. N. Starr, K. D. Malone, H. Y. Chu, J. D. Bloom, Comprehensive Mapping of Mutations to the SARS-CoV-2 Receptor-Binding Domain That Affect Recognition by Polyclonal Human Serum Antibodies. *bioRxiv*, 2020.2012.2031.425021 (2021).

33.   Y. Weisblum, F. Schmidt, F. Zhang, J. DaSilva, D. Poston, J. C. Lorenzi, F. Muecksch, M. Rutkowska, H.-H. Hoffmann, E. Michailidis, C. Gaebler, M. Agudelo, A. Cho, Z. Wang, A. Gazumyan, M. Cipolla, L. Luchsinger, C. D. Hillyer, M. Caskey, D. F. Robbiani, C. M. Rice, M. C. Nussenzweig, T. Hatziioannou, P. D. Bieniasz, Escape from Neutralizing Antibodies by SARS-CoV-2 Spike Protein Variants. *eLife* **9**, e61312 (2020).

34.   Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C. J. Neufeldt, B. Cerikan, J. M. Lu, J. Peukes, X. Xiong, H.-G. Kräusslich, S. H. W. Scheres, R. Bartenschlager, J. A. G. Briggs, Structures and Distributions of SARS-CoV-2 Spike Proteins on Intact Virions. *Nature* **588**, 498-502 (2020).

35.   A. R. Thierry, B. Roch, Neutrophil Extracellular Traps and by-Products Play a Key Role in COVID-19: Pathogenesis, Risk Factors, and Therapy. *J Clin Med* **9**, 2942 (2020).

36.   C. Toelzer, K. Gupta, S. K. N. Yadav, U. Borucu, A. D. Davidson, M. Kavanagh Williamson, D. K. Shoemark, F. Garzoni, O. Staufer, R. Milligan, J. Capin, A. J. Mulholland, J. Spatz, D. Fitzgerald, I. Berger, C. Schaffitzel, Free Fatty Acid Binding Pocket in the Locked Structure of SARS-CoV-2 Spike Protein. *Science* **370**, 725 (2020).

37.   Y. Choi, A. P. Chan, Provean Web Server: A Tool to Predict the Functional Effect of Amino Acid Substitutions and Indels. *Bioinformatics* **31**, 2745-2747 (2015).

38.   R. Vaser, S. Adusumalli, S. N. Leng, M. Sikic, P. C. Ng, Sift Missense Predictions for Genomes. *Nat Protoc* **11**, 1-9 (2016).

39.   S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mega X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).

40.   P. C. Ng, S. Henikoff, Sift: Predicting Amino Acid Changes That Affect Protein Function. *Nucleic Acids Res* **31**, 3812-3814 (2003).

Acknowledgments

Author contributions: S.P, L.L and B.R.K provided data analysis, visualization and draft writing. K.S contributed to visualization and manuscript preparation. D.M-R conceived the project and supervised the analysis and writing.

Competing interests: The authors have no competing interests to declare.

Figures and Tables

Table 1: Predicted biological impact of mutations in the spike protein and ΔAffinity calculations. Predicted effects of amino acid substitutions common in **(A)** SARS-CoV-2 501Y.V1 and **(B)** SARS-CoV-2 501Y.V2 using Protein Variation Effect Analyzer (PROVEAN) (*37*) and SIFT (*40*). Invariant (fixed) mutations in all the 501Y.V2 isolates are shaded blue **(C)**: ΔAffinity calculations for RBD variants.

a.

SARS-CoV-2 501Y.V1

Variation	PROVEAN	SIFT
His69 Deletion	N/A	N/A
Val70 Deletion	N/A	N/A
Tyr 145 Deletion	N/A	N/A
Asn501Tyr*	Neutral	Tolerated
Ala570Asp	Neutral	Tolerated
Asp614Gly*	Neutral	Tolerated
Pro681His	Neutral	Tolerated
Thr716Ile	Deleterious	Deleterious
Ser982Ala	Neutral	Tolerated
Asp1118His	Neutral	Tolerated

* Mutation common to both variants.
N/A = not applicable as the software does not make predictions about deletions.

b.

SARS-CoV-2 501Y.V2

Variation	PROVEAN	SIFT
Leu18Phe	Neutral	Tolerated
Asp80Ala	Neutral	Tolerated
Asp215Gly	Neutral	Tolerated
Leu242 Deletion	N/A	N/A
Leu242-Leu244 Deletion	N/A	N/A
Arg246Ile	Neutral	Deleterious
Lys417Asn	Neutral	Tolerated
Glu484Lys	Neutral	Tolerated
Asn501Tyr*	Neutral	Tolerated
Asp614Gly*	Neutral	Tolerated
Ala701Val	Neutral	Tolerated

c.

Mutation	ΔAffinity (kcal/mol)
K417N	1.67
E484K	0.61
N501Y	2.29

**Fig 1**: **Mutant position in the open and closed conformation of the spike protein. (A)** The position of various mutations of the Y501.V1, **(B)** The N501Y and D614G mutations, which occur in both the Y501.V1 and Y501.V2 variants, and **(C)** Mutations unique for the 501Y.V2. Each mutation is shown in red in one of the monomers in the 3D structure of the spike protein trimer. Shown is the closed conformation (light gray) and the open (dark gray) conformation.
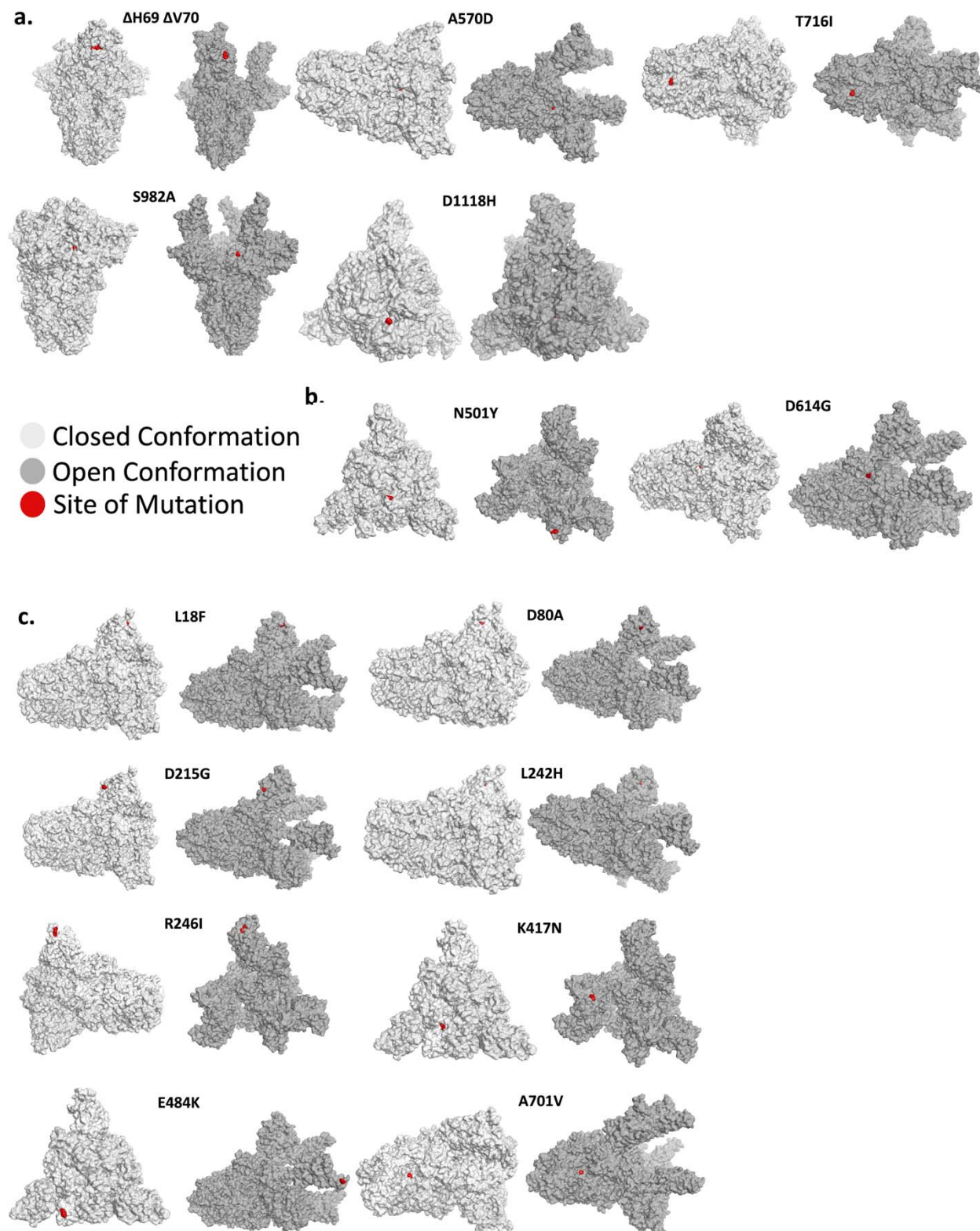
Fig. 2: Potential elastase sites in the spike protein of the new SARS-CoV-2 variants. Potential new elastase sites in **(A)** the more infective UK variant, 501Y.V1 and **(B)** the more infective South Africa variant, 501Y.V2. Images generated using BioRender. Blue boxes indicate fixed mutations and gray boxes – non-fixed in all the sequenced 501Y.V2 (*13*).
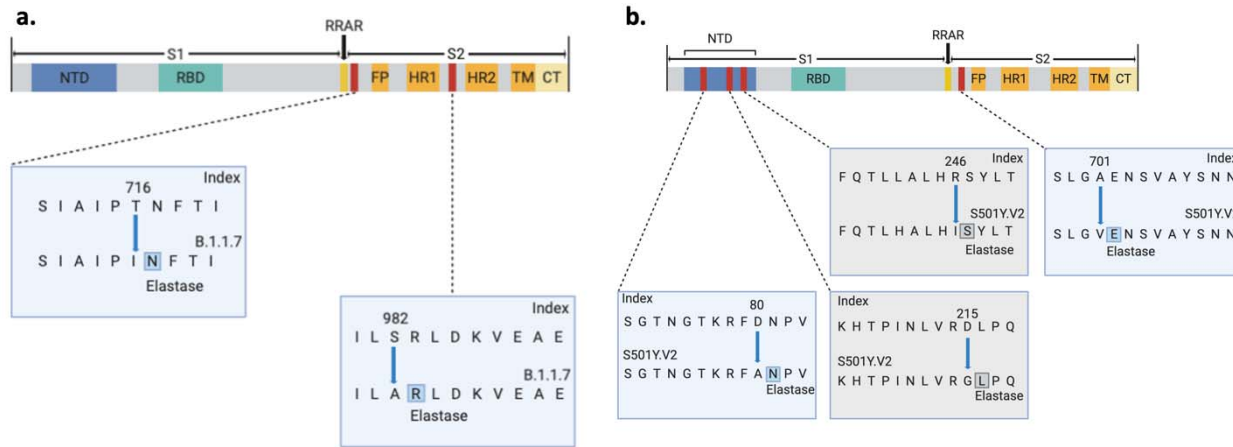
Fig. 3: Potential destabilizing effects of the point mutations on the closed conformation of the spike protein trimer. **(A)** Distance between the backbones of Asn501 in closed conformation is shown (left) and typical length of Tyr side chain is shown (right). Asn501Tyr mutation might introduce steric clashes and destabilize the closed conformation. **(B)** Ser982 (monomer 2)-Thr547 (monomer 1) interaction in closed (left) and open (right) conformation. Ser982Ala results in loss of this inter-monomeric interaction. **(C)** Lys417 (monomer 2)-Tyr369 (monomer 1) interaction in closed (left) and open (right) conformation. Lys417Asn mutation may reduce this inter-monomeric interaction.
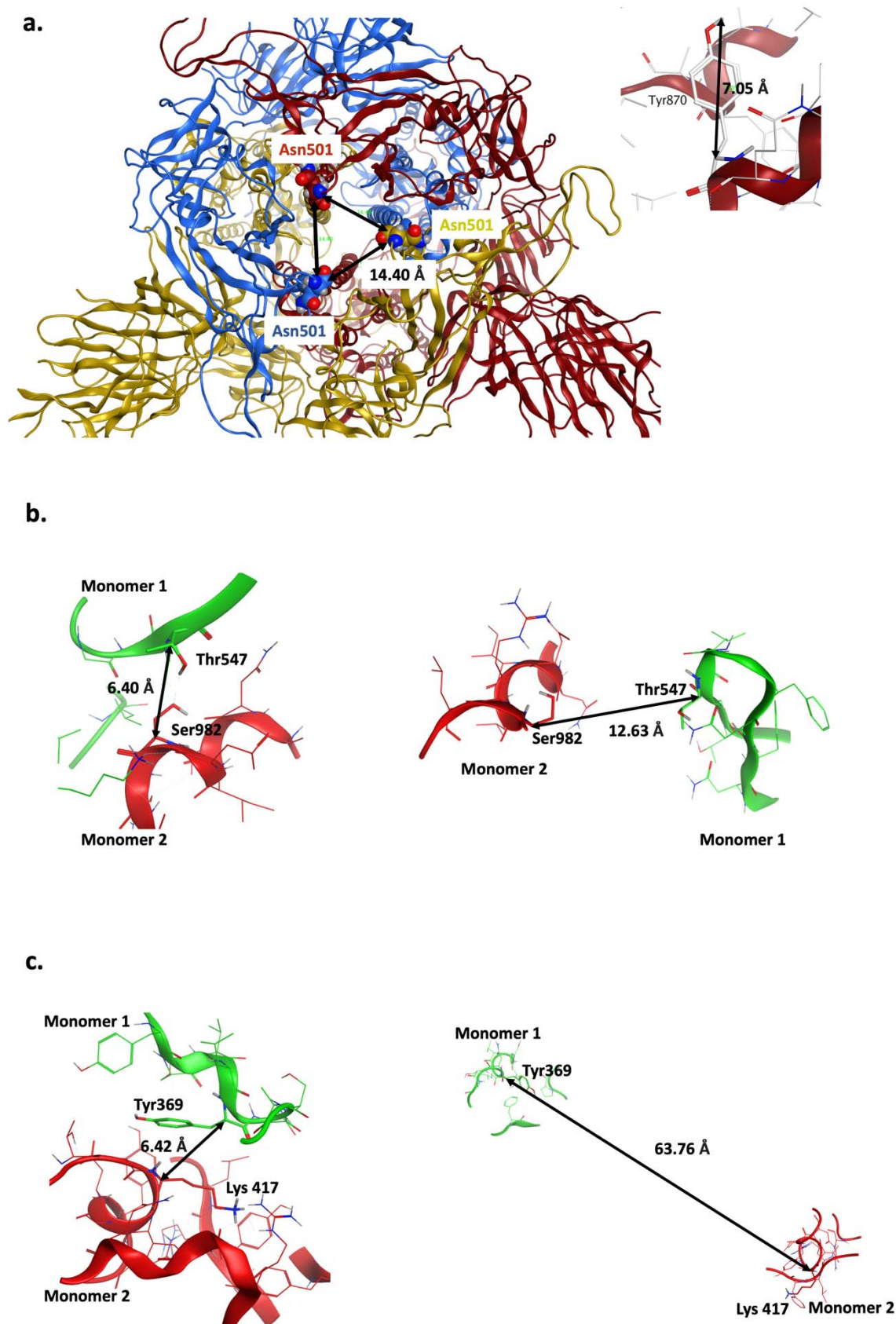
Fig. 4: Sequence alignment of the spike protein of the SARS-CoV-2 variants. (A) Conservation of mutations found in 501Y.V1 in comparison to the Wuhan index human, bat, civet and pangolin spike protein sequences. **(B)** Conservation of mutations found in 501Y.V2 in comparison to the index Wuhan human, bat, civet and pangolin spike protein sequences. Sequence alignments generated using MEGA X software.