High-throughput molecular dynamics-based alchemical free energy calculations for predicting the binding free energy change associated with the common mutations in the spike receptor-binding domain of SARS-CoV-2

Rajendra Bhadane¹,²,*, Outi M. H. Salo-Ahen¹,²,*

¹Structural Bioinformatics Laboratory, Faculty of Science and Engineering, Biochemistry, Åbo Akademi University, FI-20520 Turku, Finland

²Pharmaceutical Sciences Laboratory, Faculty of Science and Engineering, Pharmacy, Åbo Akademi University, FI-20520 Turku, Finland

*Correspondence:

Rajendra Bhadane Tel + 358-2-2154006; Email: rajendra.bhadane@abo.fi

Outi M. H. Salo-Ahen Tel +358-2-215 4009; Email: outi.salo-ahen@abo.fi
Abstract:

The ongoing pandemic caused by SARS-CoV-2 has gone through various phases. From the initial outbreak the virus has mutated several times, with some lineages showing even stronger infectivity and faster spread than the original virus. Among all the variants, beta, gamma, delta and the latest (omicron) are currently classified as variants of concern (VOC) while the remaining are labelled either as variants of interest (VOI) or variants under monitoring (VUM). In this work, we have focused on the mutations observed in important variants, particularly at the receptor-binding domain (RBD) of the spike protein that is responsible for the interactions with the host ACE2 receptor and binding of antibodies. Studying these mutations is particularly important for understanding the viral infectivity, spread of the disease and for tracking the escape routes of this virus from antibodies. Molecular dynamics (MD) based alchemical free energy calculations have been shown to be very accurate in predicting the free energy change due to a mutation that could have a deleterious or a stabilising effect on the protein itself or its binding affinity to another protein. Here, we investigated the significance of six commonly observed spike RBD mutations on the stability of the spike protein binding to ACE2 by free energy calculations using high throughput MD simulations. For comparison, we also used other (rigorous and non-rigorous) binding free energy prediction methods and compared our results with the experimental data if available. The alchemical free energy-based method consistently predicted the free-energy changes with an accuracy close to ±1.0 kcal/mol when compared with the available experimental values. As per our simulation data the most significant mutations responsible for stabilising the spike RBD interactions with human ACE2 are N501Y and L452R.

Keywords: molecular dynamics simulation, SARS-CoV-2, alchemical free energy calculation, binding free energy, spike protein, receptor-binding domain, mutation
Introduction:

In the current pandemic caused by SARS-CoV-2, the study of viral mutations can give us understanding on the infectivity, pathogenesis, and drug resistance of the virus as well as insight into vaccination efficiency. Further, it may help tracking immune escape routes and consequent emergence of new diseases. During the second wave of COVID-19, SARS-CoV-2 mutations drew much more attention than those of any other virus studied till date. However, it is important to point out that viral mutation is not an unusual phenomenon. For any organism, the mutation rate is the change in genetic information that passes from one generation to the next. The viral mutation rate involves a process that starts from the viral particle’s attachment to the host, its entry, release of genetic material, replication of viral particles, assembly and release or escape from the host cell. Viruses show a higher mutation rate than other organisms due to their short generation time and large population size. The mutation rate of viruses also depends on the type of virus. RNA viruses show higher mutation rates than DNA viruses, e.g., HIV virus shows extremely high mutation rate due to its higher replication rate. When HIV and SARS-CoV-2 viruses are compared, despite both being RNA viruses, mutations of SARS-CoV-2 are slower than in HIV. There are also many other factors that could influence the mutation rate of a particular virus. It is also important to understand that not every mutation is beneficial to a virus as most mutations are deleterious, leading to the instability of the viral particles. The highly deleterious mutation gets rapidly purged while neutral or mildly deleterious stays. These mild mutations impact the virus phenotype and could be beneficial for pathogenicity, infectivity, transmissibility and/or antigenicity of virus particles.

Since the first outbreak of SARS-CoV-2 in Wuhan, China, the virus has mutated to various lineages. Among these so-called Pango (Phylogenetic Assignment of Named Global Outbreak) lineages some are classified as variants of concern (VOC), variants of interest (VOI) or variants under monitoring (VUM). According to the simplified nomenclature by the World Health Organisation (WHO), Pango lineages such as B.1.1.7, B.1.351, P.1 and B.1.617.2 are named with the letters of the Greek alphabet, alpha, beta, gamma and delta, respectively. Currently (February 2022), the B.1.1.529 lineage (omicron) and its descendants are the dominant variants worldwide. Among the known variants, beta, gamma, delta and the latest (omicron) are VOC as of February 2022 (https://www.ecdc.europa.eu/en/covid-19/variants-concern). Of special interest in the present work are the mutations occurring in the receptor-binding domain (RBD) of the SARS-CoV-2.
spike protein(8)(9) since that domain plays a major role in the entry of the virus into the host cell by binding to the human angiotensin-converting enzyme 2 (hACE2). For example, a sub lineage of the alpha variant that was first detected in the U.K in December 2020 carries E484K and N501Y mutations in the spike RBD and is 30-50% more infectious than the original virus(10).

The exponential growth in computational power has facilitated the utilisation of in silico methods. Computationally expensive free energy calculations that were previously thought to be impossible can now be performed in a few hours. For an ideal free-energy prediction method, the predictive accuracy should be of the same range as that obtained by experimental methods. In a so-called “alchemical” transformation, an amino acid is transformed from one state to another via a non-physical pathway. This transformation can be either reversible or irreversible and can be performed by an equilibrium or non-equilibrium method, respectively(11). The major advantage of this approach is the prediction accuracy that matches that reached in experiments(12).

In this work, we performed alchemical free energy calculations, utilising high throughput molecular dynamics (MD) simulations to predict the binding free energy change (with respect to binding partner ACE2) upon six spike RBD mutations commonly observed in current VOC and early VOI lineages. We performed altogether 24 equilibrium MD simulations and 1200 non-equilibrium MD simulations to ensure exhaustive conformational sampling of each of the mutants. We also carried out comparative studies with less rigorous computational methods and discuss the results in the light of available experimental data and/or other computational studies. The results of this comparative study may shed further understanding on the impact of these mutations on infectivity and spread of SARS-CoV-2. Most importantly, we show that the relatively fast free energy calculations may provide a reasonable alternative for expensive experimental mutation studies.

**Results:**

**Selection of mutations for the study:**

As per the Pango lineage definition, till date, there are more than 1200 lineages found for SARS-CoV-2(13). The B.1 lineages with the D614G mutation in the spike protein observed in early-2020 gave rise to more than 20% increase in transmissibility compared to the original
strain of SARS-CoV-2 Wuhan, China. Other variants also appeared, including many lineages with E484K as the common mutation that was found to reduce the binding affinity of antibodies(14–18). Figure 1 shows the rough timeline of VOC, VOI and VUM lineages with prominent mutations in the RBD region of the spike protein.

![Timeline of various SARS-CoV-2 mutations observed across the Pango lineages. The WHO proposed naming scheme is shown in brackets. The light green central arrow indicates the progress bar, and each circle represents a month, starting from November 2019 to January 2022. The circles with spike-like projections represent the month of the first appearance of new important lineages. Each chevron arrow is for individual lineages; the rectangle boxes contain the information about the key spike RBD mutations (*only in the B.1.1.7 sub lineage). Information on the WHO variants of concern are represented under the central progress bar whereas other variants of interest or under monitoring are above the bar.](image)

The epsilon variant (ε; B.1.427 and B.1.429) with a L452R mutation in the spike RBD was surfaced in California region in the USA in March 2020. The variant showed a 20% increase in transmission,(19) a decrease in susceptibility towards the monoclonal antibody combination containing bamlanivimab and etesevimab as well as reduced neutralization by convalescent and post-vaccination sera(19–21). The P2 variant (zeta, ζ) originated in Brazil during April 2020 and, like epsilon, also showed reduced neutralization by convalescent and post-vaccination sera and monoclonal antibody treatments. At about the same time, South Africa saw a big peak in COVID-19 infections, which was confirmed to be caused by the emergence of the beta variant that carries K417N, E484K and N501Y mutations in the spike RBD. Both K417N and N501Y mutations have been shown to increase the binding affinity of
SARS-CoV-2 RBD to hACE2 (22). It has been estimated that the beta variant had a 50% increase in transmission rate and a significantly reduced susceptibility towards bamlanivimab and etesevimab monoclonal antibody treatment (23–25). The gamma variant originating from Brazil has also an increased transmissibility and resembles the beta variant but has a K417T mutation instead of K417N. In spring 2021, India saw a big peak in the number of SARS-CoV-2 infections due to the delta variant (B.1.617.2) (26) that has a significantly increased infectivity and viral spread. This variant carries L452R and T478K mutations at RBD. The other variants carry also different combinations of the above-mentioned mutations. The lambda (λ) variant is classified under C.37 lineage and was reported the first time in Peru in December 2020. This variant with L452Q and F490S mutations at RBD has also improved transmissibility and ability to evade vaccine elicited antibodies (27,28), while the following variant (B.1.1.7, alpha) is known for about a 50% higher transmission rate than the original SARS-CoV-2 virus (8). Due to the transmission advantage of the alpha variant and its sub lineage (B1.1.7 + E484K) over B.1 and B.1.1 lineages, it become the dominant variant in the UK in the end of 2020, spreading to the other parts of the world. The kappa variant (B.1.617.1) first detected in India, carries L452R and E484Q mutations and has been shown to have an increased transmutability (29). The mu variant (B.1.621) from Columbia carries an additional R346K mutation not observed in the other variants, and it has shown significant resistance to antibodies elicited by both vaccination and natural SARS-CoV-2 infection (30). The iota and eta variants with E484K mutation at RBD were detected in November-December 2020 in the USA and Nigeria, respectively. In January 2021, the theta variant was reported in the Philippines. It carries E484K and N501Y mutations. Even the latest VOC, omicron, carries some of the same mutations: K417N, T478K and N501Y (+ E484A instead of E484K/Q). The six mutations selected from among the key RBD mutations for this study are R346K, L452R, T478K, E484Q, E484K and N501Y. They include both mutants that have been experimentally shown to affect the spike binding affinity to hACE2 as well as mutants that have either not been studied with respect to the spike-hACE2 binding affinity (to our knowledge) or have been shown to affect for example the immunity (antibody binding) rather than hACE2 binding (Table 1).
Table 1 Selected mutations for the study

| Mutation | Identified in variants | Increases hACE2 binding | Evading antibodies | Predicted affinity (kcal/mol) | Experimental binding affinity (K_D in nM) |
|----------|------------------------|-------------------------|--------------------|-------------------------------|-----------------------------------------|
| R346K    | µ                      | Yes(31)                 | 0.66              |                               | 34.4±0.06                               |
| L452R    | ε, β, α, κ, δ, γ, μ, θ | Yes(32–34)              | 0.395             | 1.20±0.06                     |
| T478K    | δ                      | Yes(32,34)              | -                 | -                            | 12.9±0.4                               |
| E484Q    | κ                      | Yes                     | 13±0.4            | 51.5±8.2                     |
| E484K    | ζ, β, α, γ, τ, μ, θ, 0 | Yes                     | -                 | -                            | 52.6±2.4                               |
| N501Y    | β, α, γ, μ, θ, O       | Yes                     | -                 | -                            | 89.4                                   |

N.B. Experimental binding affinity (Kd in nM) of WT spike glycoprotein to hACE2: 2.05±0.26(38) 17±0.6(41), 10±3.1(43), 62.6±7.7(44), 80(37), 75.1(47), 16(48) nM. *Binding affinity towards antibody; †Binding affinity towards hACE2; - data not available

Location of the selected mutations at the receptor-binding domain of the spike glycoprotein:

The envelope of SARS-CoV-2 has homotrimeric spike glycoproteins that consist of a S1, S2 and a short cytoplasmic domain. To interact with the human angiotensin-converting enzyme 2 (hACE2) the virus utilises the amino acids 331-524 of the S1 domain, that is, the receptor-binding domain (RBD)(49). To structurally analyse the selected mutations, we searched for the available 3D structures of the spike RBD bound with the hACE2 receptor. To date, there are hundreds of SARS-CoV-2 spike protein structures available in the Protein Data Bank (PDB), of which tens of structures are bound with hACE2. This includes structures determined both using X-ray crystallography and cryo-EM. The resolution of these structures varies between 2-5 Å. We set the resolution criteria of 2.0-3.0 Å. As we focused on the RBD of the spike protein, we ignored the structures of full-length spike glycoproteins. We chose the structure deposited by Wang et al. (PDB ID: 6LZG, resolution 2.5 Å)(49) for this study. The characteristic feature of the spike RBD core region (residues 333-437, 507-527) is a twisted, five-stranded β-sheet (formed by antiparallel β1-4 and β7 strands) with short connecting helices (α1-α3) and loops. Between β4 and β7 strands there is an external subdomain (residues 438-506) that is formed by α4 and α5 helices and a flexible loop that connects two short β strands (β5 and β6).
subdomain is also called as the receptor-binding motif (RBM) region since it makes direct
contacts with the hACE2 receptor. Except for the R346K mutation that is in the core, right after
the α1 helix, all other selected mutations are in the RBM region. L452R is at the β5 strand,
T478K and E484K/E484Q mutations are located between the β5 and β6 strands while N501Y
precedes the α5 helix(50). Of note, only residues at positions 501 and 484 have direct
interactions with hACE2. Figure 2 shows the location of the mutations selected to be studied
here.

Figure 2 Location of the studied mutations in the spike receptor-binding domain (RBD) of
SARS-CoV-2 (PDB ID: 6LZG). Right: Surface representation of the spike RBD in complex with
hACE2 (colour scheme: spike RBD in white opaque surface, hACE2 in blue surface, locations
of mutations in dark purple colour). Left: Transparent surface and cartoon representation of
the spike RBD (colour scheme: the core of spike RBD in transparent white surface, the
receptor-binding motif (RBM) domain in blue surface, the core beta sheets are shown in pink
cartoon, and the key amino acids in elemental ball and stick; oxygen atoms in red, carbon
atoms in grey, nitrogen atoms in blue).

Prediction of spike RBD-hACE2 interaction stability by non-rigorous methods:
Currently, there are various tools available for predicting the functional and/or structural impact
of mutations on a protein or protein-protein complexes. In this work, we have utilised
Schrödinger’s Prime/MM-GBSA-based residue scanning (Prime, Schrödinger, LLC, New York, NY, 2021)(51,52) as well as CUPSAT(53–55) and FoldX(56–59) tools for predicting stabilising or destabilising effect of the selected mutations on spike RBD-hACE2 complex. In addition, we calculated the effect of the mutations on the stability of the spike RBD alone. The major advantage of these calculations is that they are fast and computationally non-costly. CUPSAT and FoldX tools are also available for free. The downside of these predictions is that they are less accurate and, thus more qualitative than quantitative.

The predicted stabilising and destabilising effects of the selected spike RBD mutations on the RBD binding interaction with hACE2 and the stability of RBD are summarised in Table 2. A $\Delta \Delta G$ value $\geq |1.0|$ kcal/mol was counted as a significant change (“hotspot”) for CUPSAT/FoldX predictions and $\geq |3.0|$ kcal/mol for Prime predictions(60). Since all the mutations are found in the viable virus variants, it can be assumed that they should not destabilize the actual spike RDB structure drastically.

**Table 2 Predicted effect of mutations on the stability of the hACE2-spike RBD complex and the RBD alone**

| Spike RBD Mutation | CUPSATa | Prime Residue Scanning | FoldX |
|--------------------|---------|------------------------|-------|
|                    | Predicted $\Delta G$ (kcal/mol) complex / RBD only | Effect on the overall stability (complex / RBD only)b | $\Delta$Affinity$^c$ / $\Delta$Stability$^d$ (kcal/mol)$^e$ | Effect on the affinity / RBD stability | Predicted $\Delta G$ (kcal/mol) complex / RBD only | Effect on the overall stability (complex / RBD only) |
| R346K              | -0.27 / -0.4 | DS / DS | -0.03/7.41 | S / DS | -0.0003/-0.3517 | S / S |
| L452R              | 0.3 / 1.06 | S / S | -1.42/-2.14 | S / S | -0.3907/0.9966 | S / DS |
| T478K              | 0.25 / -0.06 | S / DS | -3.40/1.80 | S / DS | -0.0049/-0.22 | S / S |
| E484Q              | -1.51 / 1.15 | DS / S | 2.70/-1.38 | DS / S | -0.0564/0.2904 | S / DS |
| E484K              | -0.59 / 0.41 | S / S | 1.09/7.73 | DS / DS | -0.069/0.7285 | S / DS |
| N501Y              | -1.2 / 0.36 | DS / S | -6.15/-1.11 | S / S | 4.5373/0.0072 | DS / DS |

aThermal method; negative prediction value by CUPSAT indicates destabilisation, whereas the opposite is true for the Prime and FoldX predictions. bStabilising (S) / Destabilising (DS); cSide-chain prediction with backbone sampling using a cut-off 6 Å; dSide-chain prediction with backbone minimization using a cut-off 0 Å; eThe Prime/MM-GBSA energies can be related to experimental energies using a scale factor of 3.

The prediction tools consistently suggest that mutations L452R and T478K are stabilising the binding interaction. Predictions for the other mutations are inconsistent with these tools. According to the CUPSAT prediction for the spike RBD alone, none of the mutations affects the overall stability of the spike RBD protein significantly (only R346K and T478K are predicted to have a slight destabilising effect while the others are stabilising the protein).
The Prime/MM-GBSA-based residue scanning results for some mutants (especially N501Y) were observed to significantly depend on the used refinement cut-off distance for the side-chain predictions. The side-chain prediction with backbone sampling or simple backbone minimization (see Methods) was performed using different cut-off values for including the nearby residues in the refinement. Table 2 shows the affinity change ($\Delta$Affinity) predicted by Prime using the backbone sampling with cut-off of 6 Å. This cut-off gave the most reasonable $\Delta$Affinity value for the N501Y mutant as that mutation is known to increase the spike binding affinity to hACE2 (see the results with other cut-off values in Supporting Information, Table S1). On the other hand, the change in spike RBD stability was predicted using minimization of the mutated residue as the refinement method (cut-off 0 Å) to avoid possible errors caused by allowing the protein structure to move more than necessary. According to these results, all but the E484 mutants are predicted to increase the spike RBD binding affinity for hACE2 (T478K and N501Y showing the largest effect on the affinity), while E346K and E484K are predicted to have the greatest effect on the spike RBD stability among the mutants. Of note, the E484K mutant has been experimentally shown to exhibit a slightly lesser thermal stability than the WT whereas N501Y mutant is as stable as the WT(43). Almost all FoldX predictions for the mutants were less than 1.0 kcal/mol in quantity, suggesting an insignificant effect on the hACE2 affinity or spike RBD stability. However, inconsistent with experimental evidence, N501Y was predicted to reduce the spike RBD binding affinity to hACE2 by both FoldX and CUPSAT, and even by Prime with insufficient backbone sampling (Supporting Information, Table S1).

**Conventional MD simulation studies:**

The FoldX optimised structures of the wildtype (WT) and mutant spike RBD bound to hACE2 were used as input for conventional MD simulations. Root mean square deviation (RMSD) and root mean square fluctuation (RMSF) plots are shown in Figures S1-S2 in Supporting Information. The RMSD of the E484K and R346K mutants converges early, followed by E478K and E484Q and remains stable throughout the MD simulations. More fluctuations in RMSD can be seen for the L452R and N501Y mutants. The fluctuations are prominent towards the C-terminal end of these mutants. The number of intermolecular hydrogen bonds, Coulombic and van der Waals (vdW) interaction energies and the sum of Coulombic and vdW interaction energies for WT and spike RBD mutants with respect to hACE2 were extracted from each 100-ns MD simulation trajectory and averaged (see Figure 3).
It was observed that although the N501Y mutation reduces the average number of hydrogen bonding interactions during the simulation, the phenyl ring of the mutant Tyr501 stays very close (within 4.2-6.5 Å) to the phenyl ring of Tyr41 of hACE2 and, thus, is engaged in aromatic stacking interactions throughout the 100-ns MD simulation (Figures 4a, b; Supporting Information, Figures S3 and S4). In addition, the hydroxyl oxygen of Tyr501 may form a hydrogen bond/a dipole-ion interaction with Lys353 or its phenyl ring engages Lys353 in pi-cation interactions (Figures 4a, b).

Further, the trajectory analysis shows that L452R mutation increases hydrogen bonding as well as Coulombic and vdW interactions. This is likely due to the introduction of the charged residue (Arg452), which causes a change in the local environment (Figures 4c-4e), thus favouring increased electrostatic attraction and/or new polar interactions between the proteins. Interestingly, Figure 4c shows how Arg452 of the spike RBD mutant interacts with the backbone carbonyl of another RBD residue, Tyr449, thus indirectly providing additional stability to the interaction between Tyr449 and Asp38 of hACE2.

E484K mutation shows a slight increase in H-bonding and vdW interaction energy but the total interaction energy decreased due to the decrease in the Coulombic interaction energy as compared with the WT. The opposite charge seems to be non-optimal for the interaction as the
salt bridge between the WT Glu484 and hACE2 Lys31 is lost upon the mutation. In sum, according to the conventional MD simulations, L452R stabilises the spike RBD-hACE2 interactions while all the other studied mutations seem to destabilise the interactions somewhat.

**Figure 4.** The change in local environment at spike RBD-hACE2 interface caused by N501Y and L452R mutations. **a.** The interactions of Tyr501 with Tyr41 and Lys353 of hACE2 **b.** MD trajectory snapshots taken from the last 50 ns of the 100-ns simulation (each 50th frame) showing the pi-pi stacking between Tyr501 and Tyr41 as well as the pi-cation and hydrogen bond interactions of Tyr501 with Lys353 of hACE2. **c.** A new hydrogen bonding interaction formed between the spike RBD Tyr449 backbone carbonyl and the spike RBD mutant residue Arg452. **d.** Electrostatic surface representation of the spike RBD highlighting the WT residue Leu452. **e.** Electrostatic surface representation of the spike RBD highlighting the mutant residue Arg452. Colour scheme: spike RBD and hACE2 in cyan and white colour cartoon respectively. WT spike RBD, mutated and hACE2 amino acid residues in pink, green and grey colour elemental sticks representation, respectively; hydrogen bond, pi-pi and pi-cation interactions are shown in yellow, cyan and green colour dashed lines, respectively.

**MD-based alchemical free energy calculations:**
We did not observe a consensus between the conventional MD simulation results and those from the non-rigorous methods except for L452R. In addition, limited sampling during the MD simulations might affect the ability to predict the stability of the interactions upon mutation. Hence to increase the conformational sampling of the mutants, we carried out high-throughput MD simulations, utilizing non-equilibrium alchemical free energy calculations to investigate the effects of these mutations on hACE2 binding. To perform these calculations, hybrid topology files containing information on both WT and the mutated residues were created.

Among the selected mutations, only R346K and N501Y are non-charge-changing mutations i.e., before and after mutation the change in the charge for these mutants remains equal to zero. This is in contrast to charge-changing mutations (L452R, T478K, E484Q and E484K) where after mutation the net charge change does not equal to zero. Estimating the free energy difference for such charge-changing mutations required an additional step of performing corrections in the charge and was achieved by a so-called analytical post-simulation charge correction method(61). Table 3 contains the binding free energy calculated before and after the charge corrections (see also Table S2 in Supporting Information for the uncorrected and corrected binding free energies for each charge-changing mutation in both bound and unbound state).

Table 3 Binding free energies obtained from alchemical free energy calculations for the selected SARS-CoV-2 spike RBD mutations.

| Spike RBD mutant | Bound ΔG_{uncorr} (kJ/mol) | Unbound ΔG_{uncorr} (kJ/mol) | Uncorrected ΔΔG (kJ/mol) | Bound ΔG_{corr} (kJ/mol) | Unbound ΔG_{corr} (kJ/mol) | Corrected ΔΔG (kJ/mol) |
|------------------|----------------------------|-----------------------------|--------------------------|--------------------------|-----------------------------|--------------------------|
| R346K*           | -                          | -                           | -                        | 662.8±0.24               | 663.54±0.28                 | -0.74±0.37               |
| L452R            | -633.70±0.34               | -635.29±0.32                | -1.59±0.47               | 633.99±0.34              | 635.17±0.32                 | -1.17±0.47               |
| T478K            | 101.73±0.42                | 101.06±0.38                 | 0.67±0.57                | 101.58±0.42              | 100.94±0.38                 | 0.65±0.57                |
| E484Q            | 29.57±1.12                 | 29.19±1.51                  | 0.38±1.88                | 29.42±1.12               | 29.07±1.51                  | 0.36±1.88                |
| E484K            | 272.80±0.55                | 271.93±0.34                 | 0.87±0.65                | 272.50±0.55              | 271.68±0.34                 | 0.82±0.65                |
| N501Y*           | -                          | -                           | -                        | 290.22±0.84              | 304.24±0.46                 | -14.02±0.96              |

Note: * Charge correction not required

According to the free energy calculations, N501Y is the most significant stabilising mutation affecting the binding of spike RBD to hACE2. This is followed by L452R that shows a mild to moderate stabilising effect. The R346K mutation also shows a weak stabilising effect. The remaining three T478K, E484Q and E484K mutations are predicted to be weakly or negligibly destabilising.
We then compared the predicted binding affinity of three mutations with experimentally reported binding affinities (Figure 5, Table 1, Supporting Information, Table S3).

![Figure 5 Predicted versus experimentally reported binding affinity of spike receptor-binding domain with hACE2 for L452R, E484K and N501Y mutations. The experimental values were taken from Table 1 and converted to ΔG. The ΔΔG was calculated by subtracting the binding free energy of the WT spike (ΔΔGWT) from that of the mutant spike (ΔΔGMUT). N.B. CUPSAT values (see Table 2) have been converted to correspond to the convention that negative values are stabilizing and positive values destabilising.](image)

The results show that the MD-based alchemical free energy calculations accurately and consistently predicted the affinity when compared with the other prediction methods. The Prime-predicted values also look relatively accurate in Figure 5, but we need to remember that its reliability for some residues may drastically depend on the choice of a suitable distance cut-off during backbone sampling, which here was guided by the available information from experimental studies.

All methods were successful in predicting the effect of the L452R mutation; experimental range of ΔΔG from -0.50 to -0.35 kcal/mol (Table S1); prediction range from -0.47 to -0.28 kcal/mol (Tables 2 and 3). Even the Prime results with all used refinement cut-offs consistently showed that the mutation is increasing the spike RBD–hACE2 affinity. For the E484K mutant, the corresponding ranges were: experimental ΔΔG from -0.16 to 0.98 kcal/mol and predicted ΔΔG from -0.069 to 0.59 kcal/mol, which suggests a successful prediction even for this mutant that does not have a significant effect on the spike RBD–hACE2 affinity. The Prime-based predictions showed again variability in the values depending on the refinement protocol, some values being negative and some negative (Table S1). However, the N501Y mutant was a more difficult task for the non-rigorous methods. The experimental ΔΔG ranges from -0.72 to -2.20 kcal/mol and the predicted ΔΔG from -3.35 to 4.54 kcal/mol, with Prime (cut-off 6 Å) and the alchemical method performing the best and FoldX the worst. (If the other Prime refinement...
protocols are analysed, by far the worst prediction (7.88 kcal/mol) is given by the backbone minimization with cut-off 0 Å, see Table S1).

Discussion:

The viral mutation is an ongoing process and can evolve in response to selective pressures. The SARS-CoV-2 mutation pattern on spike protein observed till date suggests that during the first phase of COVID-19 infections, mutations with increased infectivity overcame the early variant. This can be seen for example from the disappearance of B.1 and B.1.1 lineages and their replacement by more infectious B.1.1.7 lineage in the UK. Though both B.1 and B.1.1.7 carry the characteristic D614G mutation that is known to cause a 20% increase in transmission efficiency, the latter in addition to D614G has the N501Y mutation that ensures more efficient binding of spike RBD to hACE2(62).

Among the six mutations investigated in the current study, only N501Y and E484Q/K are in direct contact with hACE2. Three other mutations are part of the RBM region that forms the binding interface with hACE2. However, none of them has a direct contact with hACE2, while R346K neither has a direct contact nor is part of the RBM region. Consistent with our results, recently reported investigations have shown that L452R and N501Y stabilize the interaction between the spike protein and the hACE2 receptor, thus increasing infectivity(19,36)(63)(64). There are still no reports on the effect of the R346K mutation on the binding interaction with hACE2, but it has been shown that it helps in escaping class 3 antibodies(65,66). The E484K has been shown to have a negligible or small effect on the spike-hACE2 binding affinity (for example, WT Kd 21.3 nM and E484K Kd 19.7 nM(67) (see also Table 1). It has also been identified as an escape mutation that emerges during exposure to monoclonal antibodies (mAbs) C121 and C144(68,69). On other hand there are no other studies on the effect of T478K or E484Q alone on the spike binding affinity to hACE2. As double mutants with L452R they have been shown to increase the binding affinity somewhat(67,70). According to our study, the individual mutations seem not to have a big effect on the spike RBD binding affinity towards hACE2.

When all the three approaches were compared, we observed a consensus only for the L452R mutation. Though we did not observe improved H-bonding, vdW or Coulombic interactions for the N501Y mutant with respect to the WT during the conventional MD simulations, we saw that the phenyl ring of Tyr501 can efficiently form aromatic stacking interactions with Tyr41
and a hydrogen bond or a pi-cation interaction with Lys353 at the hACE2 binding interface. More extensive conformational sampling (i.e., longer MD simulations with several parallel runs) might have given better interaction energies for that mutant, which would also be consistent with the Prime residue scanning results that emphasize the need for sufficient backbone sampling in the environment of the mutation. Notably, the rigorous alchemical free energy calculation method consistently predicted the free energy changes of all mutations with an accuracy close to ±1.0 kcal/mol when compared with the available experimental values.

In sum, the results of this study are consistent with the experimental data on that both N501Y and L452R mutations stabilise the binding interactions between spike RBD and hACE2. On the other hand, neither E346K in the spike RBD core nor the mutations E484Q, E484K and T478K in the RBM offer significant advantage on binding to hACE2. Our study shows also that the MD-based alchemical free-energy calculations work accurately and consistently for different types of mutations (including the more challenging charge-changing mutations) in different structural environments.

Materials and methods:

Selection of mutants:

Based on the reports on increased infectivity of alpha and beta variants and the sudden spike in infections in India in April-May 2021, the key mutations in the spike RBD of these variants were selected for the study.

Protein Preparation:

The 3D structure of the spike receptor binding domain (RBD) that is complexed with the angiotensin-converting enzyme-2 (ACE2) (PDB ID: 6LZG(49)) was retrieved from the Protein Data Bank (PDB)(70). The selected structures were processed using the Protein Preparation Wizard of Maestro (Schrödinger Release 2021-2 New York, NY, USA). Briefly, the missing hydrogen atoms were added to the structure and the hydrogen bond network was optimized using PROPKA at pH 7.0. All water molecules beyond 3 Å were removed and a restrained minimization was conducted using the OPLS4(71) force field with a convergence criterion of 0.3-Å root-mean-square deviation (RMSD) for all the heavy atoms.
Non-rigorous methods:

FoldX. The 3D structure of the spike RBD in complex with hACE2 (PDB ID: 6LZG) was stored in the FoldX directory. Using the RepairPDB command, residues having bad torsion angles, steric clashes, or high total energy were identified and repaired. The BuildModel command was then used to create the mutant PDB files and the corresponding WT structures. Finally, using the AnalyzeComplex command the interaction energy between the spike RBD and hACE2 was computed for each mutant and the corresponding WT structure, after which the free energy difference between the WT and the mutants was calculated according to \[\Delta\Delta G = \Delta G_{\text{mut}} - \Delta G_{\text{WT}}\] (72).

CUPSAT. The CUPSAT (Cologne University Protein Stability Analysis Tool) prediction model uses atom and torsion angle potentials that are specific to the empirical environment of amino acids and were derived from over 4000 protein structures (53). The model’s accuracy to predict the difference in free energy of unfolding between WT and mutant proteins has been shown to be over 80%. The complex stability upon mutation was predicted by providing the PDB ID of the protein-protein complex (6ZLG) to the webserver (http://cupsat.tu-bs.de/) using the thermal experimental method and selecting one mutation at a time. The stability of the spike RBD upon mutation was calculated by providing only the B chain of the PDB structure for the server.

Prime Residue Scanning. To predict the effect of individual mutations and to generate input files for MD simulations, residue scanning function of the BioLuminate module in Schrödinger’s Maestro (Schrödinger Release 2021-2: BioLuminate, Schrödinger, LLC, New York, NY, 2021) was utilized (60). It employs the Prime/MM-GBSA (molecular mechanics generalized Born surface area) method (Prime, Schrödinger, LLC, New York, NY, 2021) (51, 52) to calculate both the stability and the affinity of the protein in a complex. Using the residue tab, the mutations at selected positions were defined. To see the effect of the cut-off value that
defines how far from the mutated residue also the neighboring residues are minimized/sampled, the mutated structure was refined using backbone minimization with cut-off 0 Å and backbone sampling with cut-offs of 0, 4.0, 6.0 and 9.0 Å. The backbone sampling allows adjustments in backbone between the Cα atoms of the adjacent residues, between Cα and Cβ atoms of the residue as well as side chain rotamer adjustments before the final minimization, whereas the minimization adjusts only the side chain after which the whole residue is minimized. The output provides the change in the binding affinity (ΔAffinity) of the specified binding partners (spike RBD binding to hACE2) and change in the stability of the protein (spike RBD) upon mutation (ΔStability) in kcal/mol along with the wildtype (WT) and mutated protein structures. Here, we calculated the stability of the spike RBD separately using just the RBD structure (not the complex) as the input. The output structures from the residue minimization protocol with cut-off 0 Å were used as input for performing the conventional MD simulations. The OPLS4 force field was applied for these calculations.

Conventional MD simulations:

The WT and mutated structures of spike RBD in complex with hACE2 were submitted to a 100-ns molecular dynamics (MD) simulation using Desmond (Schrödinger Release 2021-2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, USA, 2021)(73). The simulation systems were prepared using the System Builder tool of the Desmond module. The single point charge (SPC) water(74) was chosen as the explicit solvation model. Each system was neutralized using an appropriate number of Na⁺ or Cl⁻ counter ions. An orthorhombic simulation box, with Periodic Boundary Conditions (PBC) and a 10-Å buffer space between the solute and the box edge, was used for each system. The simulation systems were minimized and equilibrated before the production simulation in a stepwise manner. After the system relaxation, the production simulation was performed in the NPT ensemble for 100 ns, using a reversible reference system propagation algorithm (RESPA) integrator(73). The
temperature (300 K) was set using the Nosé–Hoover chain thermostat\cite{75–77}, with a relaxation time of 1.0 ps. The pressure was set at 1.01325 bar with the Martyna–Tobias–Klein barostat\cite{78}, using isotropic coupling and a relaxation time of 2.0 ps. Long-range interactions were handled using the U-series method\cite{79} and for short range interactions, a cut-off radius of 9.0 Å was used. The MD trajectories were analyzed using the Maestro built-in Simulation Interactions Diagram and simulation event analysis tool and Microsoft Excel360.

**MD-based alchemical free energy calculations:**

Alchemical relative free energy calculations exploit and avoid the need to simulate binding and unbinding events by making use of the fact that the free energy is a state function and can be explained as in Equation 1\cite{80,81}).

\[
\Delta\Delta G_{bind,AB} \approx - k_B T \left( \ln \frac{Z(RB)}{Z(RA)} - \ln \frac{Z(R + B)}{Z(R + A)} \right)
\]

\[
= - k_B T \left( \ln \frac{Z(RB)}{Z(RA)} - \ln \frac{Z(B)}{Z(A)} \right)
\]

\[
= \Delta G_{bound} - \Delta G_{unbound}
\]

(Eq. 1)

where, \(\Delta G_{bound}\) and \(\Delta G_{unbound}\) represent the free energy of transformation of state A (WT) to state B (mutant) and the reverse transformation from state B to A. The unbound state is independent on the presence of the receptor in the simulations. The thermodynamic cycle\cite{80} is illustrated in Figure 5.
Figure 6. Alchemical free energy cycle represented for the SARS-CoV-2 spike RBD (WT and mutant) in unbound and bound state with hACE2. The arrows indicate the free energy cycle. Leu-452 in WT and Arg-452 in the mutant structure are shown in pink ball and stick representation, spike RBD in grey and hACE2 in purple cartoon with a white transparent surface representation. (The free energy cycle is repeated in same way for all the other mutations.)

According to Equation 1 and Figure 3, the difference in the free energy for example for the L452R mutation can be calculated by performing two independent MD simulations and estimating the free energy change from Leu-452 to Arg-452 in the presence or absence of hACE2, i.e., (ΔG\text{bound}) and (ΔG\text{unbound}), respectively. In an alchemical simulation this change from state A to state B is performed by introducing an alchemical progress parameter, lambda (\lambda) that controls the potential energy function U(\vec{q};\lambda) (q is the x,y,z coordinates of the simulation system). This is achieved by generating and simulating a hybrid topology of amino
acids composed of the atoms that represent both A and B states. A subset of the energetic terms in $U(\vec{q}; \lambda)$ are modulated such that at $\lambda_A$ the atoms representing state A are activated and those representing state B are non-interacting “dummy atoms” and vice versa at $\lambda_B$. In this way the alchemical transformation is performed and the free energy of mutating A to B in any environment (e.g., bound or unbound state) is computed as per the following Equation 2(80,82).

$$\Delta G_{env} = -k_B T \ln \frac{Z(\lambda_B)}{Z(\lambda_A)} = -k_B T \ln \frac{\int \Gamma_{env}(U(\vec{q}; \lambda_B)) d\vec{q}}{\int \Gamma_{env}(U(\vec{q}; \lambda_A)) d\vec{q}}$$  \hspace{1cm} (Eq. 2)

To achieve a good phase-space overlap between the two thermodynamic states, multiple intermediate alchemical states are introduced at values $\lambda_A = \lambda_0$, ..., $\lambda_A = \lambda_B$ during the simulations. These intermediate states do not belong to either state A or B, but the mixture of the two or a hybrid. The derivatives of the Hamiltonian with respect to $\lambda$ are recorded at every step and free energies are calculated from the work distributions obtained from integration as per the Equation 3(83).

$$W = \int_{\lambda=0}^{\lambda=1} \frac{\partial H}{\partial \lambda} d\lambda$$  \hspace{1cm} (Eq. 3)

Soft-core potentials are used for both electrostatics and Lennard-Jones interactions as implemented in GROMACS. Finally, $\Delta \Delta G$ is estimated for each mutation by calculating the intersection of the forward and backward work distributions according to the Bennet’s Acceptance Ratio (BAR) method(84).

**Preparation of Hybrid Topology.** The PMX webserver(85,86) was used to generate the hybrid topology files for the selected amino acids. The pdb2gmx command was executed to reassign correct hydrogen atoms in the input coordinate files. The Amber99SB force field was used(12,87) and the output files containing the topology information were generated in a Gromacs readable form.

**Equilibrium and non-equilibrium MD simulations.** All MD simulations (equilibrium and non-equilibrium) were carried out on Puhti HPC cluster(88) with GROMACS-
The hybrid topologies generated by the PMX webserver were used as an input. For each mutation, we prepared an individual simulation box and performed forward and backward transition with two simulation systems representing the protein in state A and state B (λ= 0 and λ=1, Figure 3). Amber99SB force field and the Joung and Cheatham ion parameters were used. Each state was solvated using a dodecahedral water box with simple point charge 216 (SPC216) 3-point water model and neutralized with Na⁺ and Cl⁻ ions at a 0.15 M concentration. Each simulation system was then energy-minimized for 100000 steps using the steepest descent algorithm. This was followed by 1-ns NPT ensemble simulation with positional restraints, using stochastic leap-frog integrator and isotropic Berendsen pressure coupling. The final production MD simulations were performed using the stochastic leap-frog integrator for 50 ns. All bond lengths were constrained using the LINCS algorithm. The constant pressure and temperature condition was maintained using the Parrinello–Rahman pressure coupling at 1 atm and temperature coupling at 298.15 K. A 2-fs time step was used, and the trajectory was recorded at every 10 ps. To treat the long-range electrostatics, Particle Mesh Ewald (PME) method with a direct space cut-off of 1.1 nm and Fourier grid spacing of 0.12 nm were used. The relative strength of the Ewald-shifted direct potential was set to 10⁻⁵. Van der Waals interactions were smoothly switched off between 1.0 and 1.1 nm.

To perform the alchemical transformations after the equilibrium MD simulations, fast-growth non-equilibrium simulations were performed to estimate the binding free energy (ΔΔG) value. For this, each equilibrium simulation was used as input for a corresponding non-equilibrium run; the first 25 ns of the trajectory were omitted, and the last 25 ns of the simulation were used to generate 50 snapshots from every 100 ps. The non-equilibrium simulation was performed for a total of 5 ns with the lambda value changing continuously in forward and reverse direction from state 0 to 1 and vice-versa at each time step with a frequency.
of transition of $2 \times 10^{-7}$. The standard errors of the $\Delta G$ estimates were obtained by bootstrapping and reflect the variability in the datasets (trajectories) analyzed.

**Correction of artefacts in charge transformative mutations.** Among the six mutations studied here, four mutations are charge transformative, which means that when alchemical transformation takes place from state A to state B, the total charge present on the amino acid changes. As we use the hybrid topology approach for the alchemical transformation, during the neutralisation stage only one state (state A or B) is sufficiently neutralised, leaving the system with a net charge of the other state\(^{(82)}\). This gives rise to a technical error during the simulation when using the state-of-the-art PME method for handling the Coulombic interactions. Hence, for a non-neutral system, any extra charge is neutralized by the implicit introduction of a uniform background charge\(^{(101,102)}\). Due to the contribution of this background charge in the free energy difference, an additional correction scheme is needed to correct the finite charge effect\(^{(61,103)}\). Here, we used an analytical post-simulation correction scheme described by Rocklin et al.\(^{(61)}\) which involves an elimination of the following finite size errors; periodicity-induced net charge correction ($\Delta G_{\text{NET}}$), periodicity-induced undersolvation ($\Delta G_{\text{USV}}$), residual integrated potential effect ($\Delta G_{\text{RIP}}$) and correction for discrete solvent effect ($\Delta G_{\text{DSC}}$). The term $\Delta G_{\text{RIP}}$ corrects for the residual integrated potential (RIP) and is important if two species are involved polyatomic charge distribution. Hence, for single point charges, a correction scheme utilising the terms $\Delta G_{\text{NET}}$ and $\Delta G_{\text{USV}}$ is sufficient to correct the finite charge effect due to neutralisation. Recently, Chen et al.\(^{(104)}\) compared different approaches required for charge corrections. They combined the ($\Delta G_{\text{NET}}$) and ($\Delta G_{\text{USV}}$) terms as per following Equation 4\(^{(104)}\):

$$
\Delta G_{\text{Corrected}} = \Delta G_{\text{Uncorrected}} + \Delta G_{\text{NET-USV}} + \Delta G_{\text{RIP}} + \Delta G_{\text{DSC}}
$$

(Eq. 4)
where, $\Delta G_{\text{NET}_{-\text{USV}}}$ term describes both periodicity-induced net charge correction and undersolvation terms. $\Delta G_{\text{RIP}}$ accounts for the difference between the solute with charges distributed on many atoms and a point charge assumed in $\Delta G_{\text{NET}_{-\text{USV}}}$ term. $\Delta G_{\text{DSC}}$ accounts for the difference between explicit water molecules used in the simulation and implicit solvent models considered in the $\Delta G_{\text{NET}_{-\text{USV}}}$ term.

**Author Contributions:** **R.B:** Performed the MD simulations, the data analysis and wrote manuscript, **O.M.H.S.-A:** Supervised the work and critically reviewed the manuscript.

**Funding:** R.B. gratefully acknowledges the financial support from the Tor, Joe and Pentti Borg Memorial Fund.

**Acknowledgements:** The Sigrid Jusélius Foundation, Biocenter Finland Bioinformatics, CSC IT Center for Science, Tor, Joe and Pentti Borg Memorial Fund and Prof. Mark Johnson and Dr. Jukka Lehtonen are gratefully acknowledged for the excellent computational infrastructure at the Åbo Akademi University.

**Conflict of Interests:** The authors declare no competing interests.

**References:**

1. CL A, I M, RDP M, TM O, J P, KTL T, et al. A Guide to COVID-19: a global pandemic caused by the novel coronavirus SARS-CoV-2. The FEBS journal. 2020 Sep 1;287(17):3633–50.
2. Sanjuán R, Domingo-Calap P. Mechanisms of viral mutation. Vol. 73, Cellular and Molecular Life Sciences. Birkhauser Verlag AG; 2016. p. 4433–48.
3. Burrell CJ, Howard CR, Murphy FA. Pathogenesis of Virus Infections. In: Fenner and White’s Medical Virology. Elsevier; 2017. p. 77–104.
4. SDW F, BR M, SL KP. Neutral Theory and Rapidly Evolving Viral Pathogens. Molecular biology and evolution. 2018 Jun 1;35(6):1348–54.
5. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nature Reviews Microbiology 2021 19:7. 2021 Jun 1;19(7):409–24.

6. Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature Microbiology [Internet]. 2020 Nov 1 [cited 2021 Jun 21];5(11):1403–7. Available from: https://pubmed.ncbi.nlm.nih.gov/32669681/

7. Tracking SARS-CoV-2 variants [Internet]. [cited 2021 Jun 21]. Available from: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

8. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. 2021 Apr 9;372(6538):eabg3055.

9. Collier DA, De Marco A, Ferreira IATM, Meng B, Datir R, Walls AC, et al. SARS-CoV-2 B.1.1.7 sensitivity to mRNA vaccine-elicited, convalescent and monoclonal antibodies. medRxiv: the preprint server for health sciences. 2021 Feb 15;

10. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. 2021 Apr 9 [cited 2021 Jun 21];372(6538):eabg3055. Available from: https://pubmed.ncbi.nlm.nih.gov/33658326/

11. Lee TS, Allen BK, Giese TJ, Guo Z, Li P, Lin C, et al. Alchemical binding free energy calculations in AMBER20: Advances and best practices for drug discovery. Journal of Chemical Information and Modeling. 2020 Nov 23;60(11):5595–623.

12. Gapsys V, Michielssens S, Seeliger D, de Groot BL. Accurate and Rigorous Prediction of the Changes in Protein Free Energies in a Large-Scale Mutation Scan. Angewandte Chemie - International Edition. 2016 Jun 20;55(26):7364–8.

13. O’Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evolution [Internet]. 2022 Jan 15 [cited 2022 Feb 28];7(2). Available from: https://academic.oup.com/ve/article/7/2/veab064/6315289

14. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host and Microbe [Internet]. 2021 Mar 10 [cited 2021 Jun 21];29(3):463-476.e6. Available from: https://pubmed.ncbi.nlm.nih.gov/33592168/

15. Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. Cell Host and Microbe. 2021 Jan 13;29(1):44-57.e9.

16. Planas D, Bruel T, Grzelak L, Guivel-Benhassine F, Staropoli I, Porrot F, et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. Nature Medicine. 2021 May 1;27(5):917–24.

17. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell. 2020 Sep 3;182(5):1295-1310.e20.

25
18. Zhou D, Dejnirattisai W, Supasa P, Liu C, Mentzer AJ, Ginn HM, et al. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. Cell. 2021 Apr 29;184(9):2348-2361.e6.

19. Deng X, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, et al. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. Cell. 2021;

20. Federation Drug American (FDA). Fact Sheet for Health Care Providers Emergency Use Authorization of Bamlanivimab and Etesevimab. 2020.

21. SARS-CoV-2 Variant Classifications and Definitions [Internet]. [cited 2021 Jun 26]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Interest

22. Mohammad T, Choudhury A, Habib I, Asrani P, Mathur Y, Umair M, et al. Genomic Variations in the Structural Proteins of SARS-CoV-2 and Their Deleterious Impact on Pathogenesis: A Comparative Genomics Approach. Frontiers in Cellular and Infection Microbiology. 2021 Oct 13;11:951.

23. Jangra S, Ye C, Rathnasinghe R, Stadlbauer D, Alshammary H, Amoako AA, et al. SARS-CoV-2 spike E484K mutation reduces antibody neutralisation. The Lancet Microbe. Elsevier Ltd; 2021.

24. Garcia-Beltran WF, Lam EC, St. Denis K, Nitido AD, Garcia ZH, Hauser BM, et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. Cell. 2021 Apr 29;184(9):2372-2383.e9.

25. Estimates of severity and transmissibility of novel SARS-CoV-2 variant 501Y.V2 in South Africa | CMMID Repository [Internet]. [cited 2021 Jun 21]. Available from: https://cmmid.github.io/topics/covid19/sa-novel-variant.html

26. How Dangerous Is the Delta Variant (B.1.617.2)? [Internet]. [cited 2021 Aug 12]. Available from: https://asm.org/Articles/2021/July/How-Dangerous-is-the-Delta-Variant-B-1-617-2

27. Issa E, Merhi G, Panossian B, Salloum T, Tokajian S. SARS-CoV-2 and ORF3a: Nonsynonymous Mutations, Functional Domains, and Viral Pathogenesis. mSystems [Internet]. 2020 Jun 30 [cited 2022 Feb 28];5(3). Available from: https://pubmed.ncbi.nlm.nih.gov/32371472/

28. Moghaddar M, Radman R, Macreadie I. Severity, Pathogenicity and Transmissibility of Delta and Lambda Variants of SARS-CoV-2, Toxicity of Spike Protein and Possibilities for Future Prevention of COVID-19. Microorganisms 2021, Vol 9, Page 2167 [Internet]. 2021 Oct 18 [cited 2022 Feb 28];9(10):2167. Available from: https://www.mdpi.com/2076-2607/9/10/2167/htm

29. Saville JW, Mannar D, Zhu X, Srivastava SS, Berezuk AM, Demers J-P, et al. Structural and biochemical rationale for enhanced spike protein fitness in delta and kappa SARS-CoV-2 variants. Nature Communications 2022 13:1 [Internet]. 2022 Feb 8 [cited 2022 Feb 28];13(1):1–10. Available from: https://www.nature.com/articles/s41467-022-28324-6

30. Uriu K, Kimura I, Shirakawa K, Takaori-Kondo A, Nakada T, Kaneda A, et al. Neutralization of the SARS-CoV-2 Mu Variant by Convalescent and Vaccine Serum. New England Journal of Medicine [Internet]. 2021 Dec 16 [cited 2022 Feb 28];385(25):2397–9. Available from: https://www.nejm.org/doi/full/10.1056/NEJMct2114706
31. Fratev F. R346K Mutation in the Mu Variant of SARS-CoV-2 Alters the Interactions with Monoclonal Antibodies from Class 2: A Free Energy Perturbation Study. Journal of Chemical Information and Modeling [Internet]. 2022 Feb 14 [cited 2022 Feb 23];62(3):627–31. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jcim.1c01243

32. Gurung AB, Ali MA, Lee J, Farah MA, Al-Anazi KM, Al-Hemaid F, et al. Structural and functional insights into the major mutations of SARS-CoV-2 Spike RBD and its interaction with human ACE2 receptor. Journal of King Saud University - Science. 2022 Feb 1;34(2):101773.

33. Deng X, Garcia-Knight MA, Khalid MM, Servellita V, Wang C, Morris MK, et al. Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. Cell [Internet]. 2021 Jun 24 [cited 2021 Jun 21];184(13):3426-3437.e8. Available from: https://pubmed.ncbi.nlm.nih.gov/33991487/

34. Deshpande A, Harris BD, Martinez-Sobrido L, Kobie JJ, Walter MR. Epitope Classification and RBD Binding Properties of Neutralizing Antibodies Against SARS-CoV-2 Variants of Concern. Frontiers in Immunology [Internet]. 2021 Jun 4 [cited 2022 Feb 23];12. Available from: /pmc/articles/PMC8212047/

35. Ferreira IATM, Kemp SA, Datir R, Saito A, Meng B, Rakshit P, et al. SARS-CoV-2 B.1.617 Mutations L452R and E484Q Are Not Synergistic for Antibody Evasion. The Journal of Infectious Diseases [Internet]. 2021 Sep 15 [cited 2022 Feb 23];224(6):989. Available from: /pmc/articles/PMC8420622/

36. Teng S, Sobitan A, Rhoades R, Liu D, Tang Q. Systemic effects of missense mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity. Briefings in Bioinformatics [Internet]. 2021 Mar 22 [cited 2021 Aug 17];22(2):1239–53. Available from: https://academic.oup.com/bib/article/22/2/1239/5917277

37. Deshpande A, Harris BD, Martinez-Sobrido L, Kobie JJ, Walter MR. Epitope Classification and RBD Binding Properties of Neutralizing Antibodies Against SARS-CoV-2 Variants of Concern. Frontiers in Immunology [Internet]. 2021 Jun 4 [cited 2022 Feb 23];12. Available from: /pmc/articles/PMC8212047/

38. Motozono C, Toyoda M, Zahradnik J, Saito A, Nasser H, Tan TS, et al. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. Cell Host & Microbe. 2021 Jul 14;29(7):1124-1136.e11.

39. Zhang J, Xiao T, Cai Y, Lavine CL, Peng H, Zhu H, et al. Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. [cited 2022 Feb 24]; Available from: https://www.science.org

40. Ding C, He J, Zhang X, Jiang C, Sun Y, Zhang Y, et al. Crucial Mutations of Spike Protein on SARS-CoV-2 Evolved to Variant Strains Escaping Neutralization of Convalescent Plasmas and RBD-Specific Monoclonal Antibodies. Frontiers in Immunology. 2021 Aug 17;12:3231.

41. Laffeber C, de Koning K, Kanaar R, Lebbink JHG. Experimental Evidence for Enhanced Receptor Binding by Rapidly Spreading SARS-CoV-2 Variants. Journal of Molecular Biology. 2021 Jul 23;433(15):167058.

42. Bai C, Wang J, Chen G, Zhang H, An K, Xu P, et al. Predicting Mutational Effects on Receptor Binding of the Spike Protein of SARS-CoV-2 Variants. Journal of the American Chemical
43. Upadhayay V, Lucas A, Panja S, Miyauchi R, Mallela KMG. Receptor binding, immune escape, and protein stability direct the natural selection of SARS-CoV-2 variants. Journal of Biological Chemistry. 2021 Oct 1;297(4):101208.

44. Barton MI, Macgowan S, Kutuzov M, Dushek O, Barton GJ, Anton Van Der Merwe P. Effects of common mutations in the sars-cov-2 spike rbd and its ligand the human ace2 receptor on binding affinity and kinetics. eLife. 2021 Aug 1;10.

45. Luan B, Wang H, Huynh T. Enhanced binding of the N501Y-mutated SARS-CoV-2 spike protein to the human ACE2 receptor: insights from molecular dynamics simulations. FEBS Letters [Internet]. 2021 May 1 [cited 2022 Feb 26];595(10):1454–61. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/1873-3468.14076

46. Hou X, Zhang Z, Gao J, Wang Y. SARS-CoV-2 spike protein N501Y mutation causes differential species transmissibility and antibody sensitivity: a molecular dynamics and alchemical free energy study. Molecular Systems Design & Engineering [Internet]. 2021 Nov 1 [cited 2022 Feb 26];6(11):964–74. Available from: https://pubs.rsc.org/en/content/articlehtml/2021/me/d1me00086a

47. Supasa P, Zhou D, Dejnirattisai W, Liu C, Mentzer AJ, Ginn HM, et al. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. Cell. 2021 Apr 15;184(8):2201-2211.e7.

48. Williams AH, Zhan CG. Fast prediction of binding affinities of the sars-cov-2 spike protein mutant n501y (UK variant) with ace2 and miniprotein drug candidates. Journal of Physical Chemistry B [Internet]. 2021 May 6 [cited 2022 Feb 26];125(17):4330–6. Available from: https://pubs.acs.org/doi/full/10.1021/acs.jpcb.1c00869

49. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell. 2020 May 14;181(4):894-904.e9.

50. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020 May 14;581(7807):215–20.

51. Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, et al. A hierarchical approach to all-atom protein loop prediction. Proteins: Structure, Function, and Bioinformatics. 2004 May 1;55(2):351–67.

52. Jacobson MP, Friesner RA, Xiang Z, Honig B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. Journal of Molecular Biology. 2002 Jul 12;320(3):597–608.

53. V P, MM G, D S. CUPSAT: prediction of protein stability upon point mutations. Nucleic acids research [Internet]. 2006 Jul [cited 2021 Aug 14];34(Web Server issue). Available from: https://pubmed.ncbi.nlm.nih.gov/16845001/

54. V P, MM G, C H, D S. Structural analysis and prediction of protein mutant stability using distance and torsion potentials: role of secondary structure and solvent accessibility. Proteins. 2007 Jan 1;66(1):41–52.
55. V P, MM G, M A, D S. Computational modeling of protein mutant stability: analysis and optimization of statistical potentials and structural features reveal insights into prediction model development. BMC structural biology. 2007;7.

56. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. Nucleic Acids Research. 2005 Jul 1;33(suppl_2):W382–8.

57. Petukhov M, Cregut D, Soares CM, Serrano L. Local water bridges and protein conformational stability. Protein Science. 1999 Jan 1;8(10):1982–9.

58. Vijayakumar M, Wong KY, Schreiber G, Fersht AR, Szabo A, Zhou HX. Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar. Journal of Molecular Biology. 1998 May 22;278(5):1015–24.

59. Schymkowitz JWH, Rousseau F, Martins IC, Ferkinghoff-Borg J, Stricher F, Serrano L. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proceedings of the National Academy of Sciences. 2005 Jul 19;102(29):10147–52.

60. Beard H, Cholleti A, Pearlman D, Sherman W, Loving KA. Applying Physics-Based Scoring to Calculate Free Energies of Binding for Single Amino Acid Mutations in Protein-Protein Complexes. PLOS ONE [Internet]. 2013 Dec 10 [cited 2022 Mar 4];8(12):e82849. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0082849

61. Rocklin GJ, Mobley DL, Dill KA, Hünenberger PH. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. Journal of Chemical Physics [Internet]. 2013 Nov 14 [cited 2021 Jun 22];139(18). Available from: https://pubmed.ncbi.nlm.nih.gov/24320250/

62. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. 2021 Apr 9;372(6538):eabg3055.

63. B K, WM F, S G, H Y, J T, W A, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 2020 Aug 20;182(4):812-827.e19.

64. E V, V H, JT M, A P, D J, Á O, et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell. 2021 Jan 7;184(1):64-75.e11.

65. Greaney AJ, Starr TN, Barnes CO, Weisblum Y, Schmidt F, Caskey M, et al. Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. Nature Communications 2021 12:1. 2021 Jul 7;12(1):1–14.

66. Liu Z, Vanblargan LA, Bloyet L-M, Fremont DH, Diamond MS, Whelan SPJ. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. 2021;

67. Augusto G, Mohsen MO, Zinkhan S, Liu X, Vogel M, Bachmann MF. In vitro data suggest that Indian delta variant B.1.617 of SARS-CoV-2 escapes neutralization by both receptor affinity and immune evasion. Allergy [Internet]. 2022 Jan 1 [cited 2022 Mar 4];77(1):111–7. Available from: https://pubmed.ncbi.nlm.nih.gov/34453338/
68. Barnes CO, Jette CA, Abernathy ME, Dam K-MA, Esswein SR, Gristick HB, et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. Nature. 2020 Dec 24;588(7839):682.

69. Y W, F S, F Z, J D, D P, JC L, et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. eLife. 2020 Oct 1;9:1.

70. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Vol. 28, Nucleic Acids Research. Oxford University Press; 2000. p. 235–42.

71. Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, et al. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. Journal of Chemical Theory and Computation. 2021 Jun 7;acs.jctc.1c00302.

72. Blanco JD, Hernandez-Alias X, Cianferoni D, Serrano L. In silico mutagenesis of human ACE2 with S protein and translational efficiency explain SARS-CoV-2 infectivity in different species. PLOS Computational Biology [Internet]. 2020 Dec 7 [cited 2022 Feb 28];16(12):e1008450. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008450

73. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. In: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC’06. New York, New York, USA: ACM Press; 2006. p. 84.

74. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction Models for Water in Relation to Protein Hydration. In Springer, Dordrecht; 1981. p. 331–42.

75. Nosé S. A unified formulation of the constant temperature molecular dynamics methods. The Journal of Chemical Physics. 1984 Jul 31;81(1):511–9.

76. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Molecular Physics. 1984;52(2):255–68.

77. Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Physical Review A. 1985 Mar 1;31(3):1695–7.

78. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. The Journal of Chemical Physics. 1994 Sep;101(5):4177–89.

79. Predescu C, Lerer AK, Lippert RA, Towles B, Grossman JP, Dirks RM, et al. The u-series: A separable decomposition for electrostatics computation with improved accuracy. Journal of Chemical Physics. 2020 Feb 28;152(8):084113.

80. Mey ASJS, Allen BK, Bruce Macdonald HE, Chodera JD, Hahn DF, Kuhn M, et al. Best Practices for Alchemical Free Energy Calculations [Article v1.0]. Living Journal of Computational Molecular Science. 2020 Dec 16;2(1):18378.

81. Gapsys V, Michielssens S, Seeliger D, de Groot BL. Insights from the First Principles Based Large Scale Protein Thermostability Calculations. Biophysical Journal. 2016 Feb 16;110(3):368a.

82. Aldeghi M, de Groot BL, Gapsys V. Accurate Calculation of Free Energy Changes upon Amino Acid Mutation. In: Methods in Molecular Biology. Humana Press Inc.; 2019. p. 19–47.
83. Seeliger D, de Groot BL. Protein thermostability calculations using alchemical free energy simulations. Biophysical Journal. 2010;98(10):2309–16.

84. Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. Journal of Computational Physics. 1976 Oct 1;22(2):245–68.

85. Gapsys V, De Groot BL. pmx Webserver: A User Friendly Interface for Alchemy. Journal of Chemical Information and Modeling. 2017 Feb 27;57(2):109–14.

86. Gapsys V, Michielssens S, Seeliger D, De Groot BL. pmx: Automated protein structure and topology generation for alchemical perturbations. Journal of Computational Chemistry. 2015 Feb 15;36(5):348–54.

87. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Vol. 65, Proteins: Structure, Function and Genetics. Proteins; 2006. p. 712–25.

88. Puhti - Puhti - Services for Research - CSC Company Site [Internet]. [cited 2021 Jun 26]. Available from: https://research.csc.fi/-/puhti

89. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. Vol. 7, Journal of Molecular Modeling. Springer; 2001. p. 306–17.

90. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. Vol. 26, Journal of Computational Chemistry. J Comput Chem; 2005. p. 1701–18.

91. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of Chemical Theory and Computation. 2008 Mar;4(3):435–47.

92. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013 Apr;29(7):845–54.

93. Páll S, Abraham MJ, Kutzner C, Hess B, Lindahl E. Tackling exascale software challenges in molecular dynamics simulations with GROMACS. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag; 2015. p. 3–27.

94. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015 Sep 1;1:1–2:19–25.

95. Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. Journal of Physical Chemistry B. 2008 Jul 31;112(30):9020–41.

96. Petrova SS, Solov’Ev AD. The Origin of the Method of Steepest Descent. Historia Mathematica. 1997 Nov 1;24(4):361–75.

97. Berendsen HJC, Postma JPM, Gunsteren WF van, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics. 1998 Aug 31;81(8):3684.
98. Van Gunsteren WF, Berendsen HJC. A Leap-Frog Algorithm for Stochastic Dynamics. Molecular Simulation. 1988;1(3):173–85.

99. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A Linear Constraint Solver for Molecular Simulations. Vol. 18, J Comput Chem. John Wiley & Sons, Inc; 1997.

100. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics. 1981 Aug 14;52(12):7182–90.

101. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics. 1993 Aug 31;98(12):10089–92.

102. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics. 1995 Aug 31;103(19):8577–93.

103. Lin YL, Aleksandrov A, Simonson T, Roux B. An overview of electrostatic free energy computations for solutions and proteins. Journal of Chemical Theory and Computation. 2014 Jul 8;10(7):2690–709.

104. Chen W, Deng Y, Russell E, Wu Y, Abel R, Wang L. Accurate Calculation of Relative Binding Free Energies between Ligands with Different Net Charges. Journal of Chemical Theory and Computation. 2018 Dec 11;14(12):6346–58.
Figure 3
Figure 5
Figure 6