

 Very Important Paper



# The Role of Spike Protein Mutations in the Infectious Power of SARS-COV-2 Variants: A Molecular Interaction Perspective

Santiago A. Gómez,<sup>[a]</sup> Natalia Rojas-Valencia,<sup>[a, b]</sup> Sara Gómez,<sup>[c]</sup> Chiara Cappelli,<sup>\*[c]</sup> and Albeiro Restrepo<sup>\*[a]</sup>

Specific S477N, N501Y, K417N, K417T, E484K mutations in the receptor binding domain (RBD) of the spike protein in the wild type SARS-COV-2 virus have resulted, among others, in the following variants: B.1.160 (20A or EU2, first reported in continental Europe), B.1.1.7 ( $\alpha$  or 20I501Y.V1, first reported in the United Kingdom), B.1.351 ( $\beta$  or 20H/501Y.V2, first reported in South Africa), B.1.1.28.1 ( $\gamma$  or P.1 or 20J/501Y.V3, first reported in Brazil), and B.1.1.28.2 ( $\zeta$ , or P.2 or 20B/S484K, also first reported in Brazil). From the analysis of a set of bonding descriptors firmly rooted in the formalism of quantum mechanics, including Natural Bond Orbitals (NBO), Quantum Theory of

Atoms In Molecules (QTAIM) and highly correlated energies within the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)), and from a set of computed electronic spectral patterns with environmental effects, we show that the new variants improve their ability to recognize available sites to either hydrogen bond or to form salt bridges with residues in the ACE2 receptor of the host cells. This results in significantly improved initial virus–cell molecular recognition and attachment at the microscopic level, which trigger the infectious cycle.

## 1. Introduction

Virus evolution is well understood in terms of the emergence of dominant populations with adapted characteristics resulting from evolutive pressures at the molecular scale.<sup>[1]</sup> The rate of mutation, estimated to be of the order of 1 in a million for each reproductive event,<sup>[2]</sup> is among the main obstacles to understand at the microscopic level the mechanism of the interactions between viruses and host cells and to produce efficient vaccines and treatments to fight viral infections. In the particular case of SARS-COV-2, cumulative mutations have originated a large number of variants, which are very difficult to track and characterize because of the fast changing reality.

As with any other disease-causing virus, the success of new variants is problematic, because of the potential to increase the rate of spread, to worsen the effects of the disease, to become more deadly, to evade detection protocols and to develop drug


resistance and even immunity against current vaccines.<sup>[3–9]</sup> A typical case is the influenza virus, for which an annual reformulation of the vaccine is needed to keep up with new variants.<sup>[10]</sup> A particular source of concern for SARS-COV-2 is that a few specific mutations have independently appeared in different places, thus suggesting that they are highly favorable for virus survival and transmission and that the virus is becoming stronger and better equipped to thrive in different local environments.<sup>[11,12]</sup>

It is well understood that the virus–cell molecular recognition and attachment is the initial event that triggers the viral cycle of infections.<sup>[13]</sup> For SARS-COV-2, this attachment involves molecular interactions between amino acids in the Receptor Binding Domain (RBD) of the spike protein and amino acids in the ACE2 receptor in host cells.<sup>[13]</sup> For the complexes reported between the wild type SARS-COV-2 and ACE2,<sup>[14]</sup> these contacts have been characterized as one strong salt bridge and a multitude of hydrogen bonds and weak non-covalent interactions.<sup>[15,16]</sup> Here, we analyze the structural, chemical, and spectral consequences of specific mutations on the RBD from a total of 5 newly reported variants,<sup>[17,18]</sup> namely, B.1.160, B.1.1.7 ( $\alpha$ ), B.1.351 ( $\beta$ ), P.1 ( $\gamma$ ), and P.2 ( $\zeta$ ), which are the subject of intensive research. In particular, we show that it is possible to rationalize the process of virus mutation and how this evolution increases the infectious power of the virus at the microscopic level, by exploiting a set of molecular descriptors related to the electronic rearrangements resulting from the RBD(S)–ACE2 interaction, and through the characterization of changes in absorption and electronic circular dichroism (ECD) spectral signatures at binding sites accounting for the chemical environment.

[a] S. A. Gómez, Dr. N. Rojas-Valencia, Prof. Dr. A. Restrepo  
Instituto de Química,  
Universidad de Antioquia UdeA  
Calle 70 No. 52–21, 050010, Medellín (Colombia)  
E-mail: albeiro.restrepo@udea.edu.co

[b] Dr. N. Rojas-Valencia  
Escuela de Ciencias y Humanidades  
Departamento de Ciencias Básicas,  
Universidad Eafit  
AA 3300 Medellín (Colombia)

[c] S. Gómez, Prof. Dr. C. Cappelli  
Scuola Normale Superiore,  
Classe di Scienze  
Piazza dei Cavalieri 7, 56126 Pisa (Italy)  
E-mail: chiara.cappelli@sns.it

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202100393>

Following the procedures introduced in a previous study of virus–cell intermolecular interactions for the wild-type SARS-CoV-2,<sup>[15]</sup> the strength of bonding interactions is described and quantified in terms of a set of QTAIM (first three items below) and NBO indexes rooted in quantum chemistry and firmly linked to the physicochemical nature of the interaction. In particular,<sup>[19–26]</sup> (i) Accumulation of electron densities at the bond critical points for intermolecular contacts. Larger densities indicate stronger interactions. (ii)  $|V(r_c)|/G(r_c)$ , i.e. the virial ratio evaluated at bond critical points. This ratio also increases for stronger interactions because larger ratios describe increasingly local dominance of the stabilizing potential energy. (iii)  $H(r_c)/\rho(r_c)$ , i.e. the bond degree parameter, which is complimentary to the virial ratio, because negative values indicate local dominance of the stabilizing potential energy while positive values indicate local dominance of the destabilizing kinetic energy, thus, larger positive values describe weaker interactions. Finally, (iv) orbital interaction energies in the donor→acceptor NBO picture and Wiberg bond indices, which are directly related to the strength of the interaction.

## 2. Discussion

Table 1 lists the primary sequences of amino acids in the RBD of the spike protein for all variants considered in this work. Table 2 lists NBO and QTAIM bonding descriptors for all variants, which are discussed next on an individual basis.

### 2.1. B.1.160

It is one of the two main variants first reported in Europe, also known as 20A/EU2, identified by Hodcroft et al.<sup>[28]</sup> Outside the RBD, this variant is defined by two mutations in the nucleocapsid and by four ORF1b mutations. One serine→asparagine mutation, S477N (Table 1, Figure 1), was found in the bonding motif of the RBD. For the calculation of descriptors and spectra in this mutation, the involved pairs were extracted from a snapshot of MD simulations performed by Singh and co-workers<sup>[29]</sup> (see section 4 for specific details).

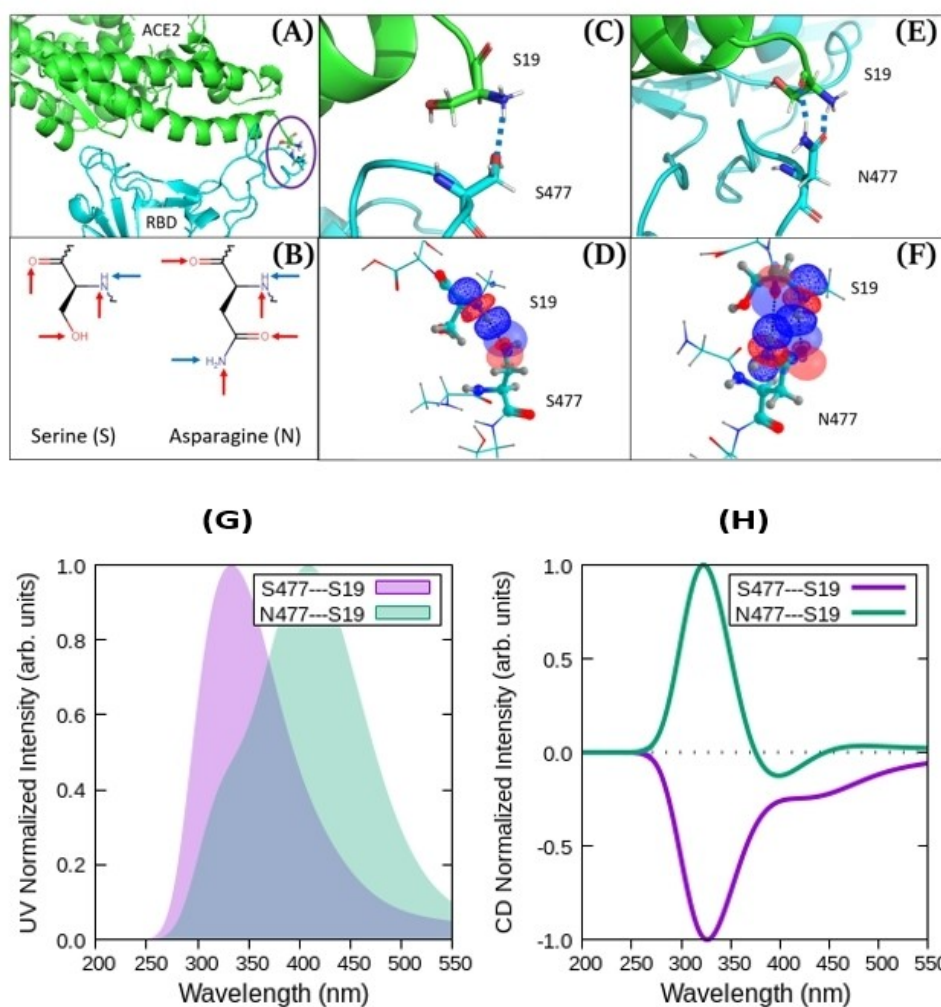
In the wild-type SARS-CoV-2, S477 had one HB active O–H group, however, previous studies<sup>[15]</sup> indicate that this amino acid does not take part in persistent bonding with the ACE2 receptor. Thus, if there were any S477→ACE2 hydrogen bonds, they were transient and did not contribute to the initial virus recognition and attachment.

**Table 1.** Mutations in the RBD of the spike protein of several variants of SARS-CoV-2.<sup>[17]</sup> The entire bonding motif is highlighted in purple. Mutations in the bonding motif are highlighted by colored boxes. Notice that for the P.1 variant, both K417N (blue box) and K417T (yellow box) are possible. We used the general notation for SARS-CoV-2 variants suggested by Rambaut and co-workers.<sup>[18]</sup>

SARS-CoV-2	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	370
B.1.160	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	
B.1.1.7	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	
B.1.351	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	
P.1	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	
P.2	N I T N L C P F G E	V F N A T R F A S V	Y A W N R K R I S N	C V A D Y S V L Y N	
SARS-CoV-2	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	410
B.1.160	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	
B.1.1.7	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	
B.1.351	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	
P.1	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	
P.2	S A S F S T F K C Y	G V S P T K L N D L	C F T N V Y A D S F	V I R G D E V R Q I	
SARS-CoV-2	A P G Q T G K I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	450
B.1.160	A P G Q T G K I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
B.1.1.7	A P G Q T G K I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
B.1.351	A P G Q T G N I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
P.1	A P G Q T G N I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
P.1*	A P G Q T G T I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
P.2	A P G Q T G K I A D	Y N Y K L P D D F T	G C V I A W N S N N	L D S K V G G N Y N	
SARS-CoV-2	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G S T P C	N G V E G F N C Y F	490
B.1.160	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G N T P C	N G V E G F N C Y F	
B.1.1.7	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G S T P C	N G V E G F N C Y F	
B.1.351	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G S T P C	N G V K G F N C Y F	
P.1	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G S T P C	N G V K G F N C Y F	
P.2	Y L Y R L F R K S N	L K P F E R D I S T	E I Y Q A G S T P C	N G V K G F N C Y F	
SARS-CoV-2	P L Q S Y G F Q P T	N G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	530
B.1.160	P L Q S Y G F Q P T	N G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	
B.1.1.7	P L Q S Y G F Q P T	Y G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	
B.1.351	P L Q S Y G F Q P T	Y G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	
P.1	P L Q S Y G F Q P T	Y G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	
P.2	P L Q S Y G F Q P T	N G V G Y Q P Y R V	V V L S F E L L H A	P A T V C G P K K S	

**Table 2.** Properties of virus cell contacts. Hydrogen bonds and salt bridges in the RBD(S) ACE2 complex of SARS-COV-2 variants. Multiple interactions for each pair of interacting fragments are possible, we only list the strongest contact in each case (see text for explicit details of the calculations). The arrows state the directionality of the donor→acceptor interaction in the corresponding hydrogen bond according to the classical electrostatic  $X^{\delta-}-H^{\delta+}-Y^{\delta-}$  description. IE: DLPNO-CCSD(T)/aug-cc-pVDZ interaction energies.  $E_{d-a}^{(2)}$  NBO orbital interaction energies (Figures 1–5). WBI: Wiberg bond indices.<sup>[27]</sup> The archetypal hydrogen bond in the water dimer is included for comparison purposes.

RBD(S)	ACE2	Variant	–IE kcal/mol	$E_{d-a}^{(2)}$ kcal/mol	NBO WBI	$10^2\rho(r_c)$ a.u.	QTAIM $\frac{ V(r_c) }{G(r_c)}$	$10^2 \frac{\chi(r_c)}{\rho(r_c)}$ a.u.	
S477	→	S19	Wild type	3.9	0.2	< 0.01	0.02	0.50	25.17
N477	↔	S19	B.1.160	13.1	7.2	0.04	2.88	0.94	4.63
N501	←	K353	Wild type	12.7	0.2	< 0.01	0.32	0.71	20.66
Y501	←	K353	B.1.1.7, B.1.351, P.1	29.8	10.77	0.04	2.71	0.89	9.72
K417	→	D30	Wild type	80.4	22.9	0.12	6.44	1.16	-13.75
N417	→	D30	B.1.351, P.1	4.6	< 0.05	< 0.01	0.04	0.35	42.21
417	→	30	P.1	1.4	< 0.05	< 0.01	0.01	0.41	30.88
E484	←	K31	Wild type	0.6	< 0.05	< 0.01	0.02	0.52	23.09
K484	→	E75	B.1.351, P.1, P.2	67.4	12.81	0.06	3.13	0.92	6.78
H–O–H	→	OH <sub>2</sub>		5.5	7.1	0.01	2.30	0.89	10.00



**Figure 1.** S477N mutation. A snapshot of the dynamics of the RBD/ACE2 complex<sup>[15]</sup> is shown in (A) and the possible acceptor (red arrows) and donor (blue arrows) sites for hydrogen bonding before and after the mutation are shown in panel (B). A close up of the S477(RBD spike) and S19(ACE2) positions with no persistent or transient interaction is shown in panel (C) and the associated donor acceptor orbital interactions are shown in (D). Hydrogen bonding in the mutated virus, in a snapshot taken from the MD simulations performed by Singh and co-workers<sup>[29]</sup> are shown in panel (E) and associated orbital interactions are shown in (F). The Computed UV and CD absorption spectra are provided in (G) and (H), respectively. Spectra of the interacting fragments involving the wild type are pictured in purple, spectra of the interacting fragments involving the mutated virus are depicted in green. Level: B3LYP/aug-cc-pVDZ/AMBER.

Computational<sup>[29]</sup> and experimental<sup>[30]</sup> evidence reveals an increase in the binding affinity towards ACE2 in the mutated virus in addition to an experimentally<sup>[31]</sup> measured reduction by one entire order of magnitude in the *in vitro* dissociation constant  $K_D$  for the attached complex. These observations, are rationalized using our calculations in several ways: There is a sensible increase in the DLPNO-CCSD(T)/aug-cc-pVDZ interaction energy between the extended pairs by one order of magnitude (−3.9 to −13.1 kcal/mol) due to the mutation. In addition, NBO and QTAIM descriptors in Table 2 uncover the molecular reasons behind the strengthening of the virus–cell interactions as a consequence of the mutation: a 30-fold increase in the orbital interaction energy, two orders of magnitude increase in the density of the bond critical point, a doubling of the virial ratio, and a remarkable decrease in the bond degree parameter. Furthermore, we see from Figure 1 that asparagine has seven possibilities of hydrogen bonding: two C=O groups and two N: groups that may act as HB acceptors and three protons that may be donated to HBs. Since it has been reported that in the mutated virus the asparagine 477 is in direct contact with S19 in ACE2,<sup>[31]</sup> the net result of this mutation is that the transient S477(S)–S19(ACE2)  $n_O \rightarrow \sigma_{N-H}^*$  HB (Figures 1(C), 1(D)) is replaced by two persistent primary  $n_O \rightarrow \sigma_{N-H}^*$  HBs, as shown in Figure 1(E), 1(F).

Moving to spectral signatures, we see that replacing the transient S477(S)–S19(ACE2) by the newly formed N477(S)–S19(ACE2) hydrogen bond leads to a noticeably ( $\Delta\lambda_{max} \approx 76$  nm) red-shifted maxima in the UV-absorption spectra. This means that the two electronic states involved in the transition get closer in energy, likely due to a larger stabilization of the excited state with respect to the ground state, when the mutation is present. More significant changes are predicted for ECD spectra. In fact, the presence of N477(S)–S19(ACE2) instead of S477(S)–S19(ACE2) interaction yields a change in sign of the spectral signal, whereas the absolute intensities of the peaks almost stay the same. This highlights differences in the mutual geometrical arrangement of transition (electric and magnetic) dipoles, from which the sign of the rotatory strength originates.<sup>[32–36]</sup>

### 2.2. B.1.1.7, $\alpha$

This variant was first reported in the United Kingdom and is also known as 20I or 501Y.V1. According to the Centers for Disease Control and Prevention (CDC),<sup>[12]</sup> this variant, which may be associated with an increased risk of death (as much as twice as deadly as the wild type according to Challen and co-workers<sup>[37]</sup>), is reported to have undergone at least six key mutations, of which only the asparagine→tyrosine N501Y (Figure 2) replacement falls within the binding motif of the RBD of the spike protein. In both the wild type and the mutated virus, G502 is hydrogen bonded to K353<sup>[15]</sup> via a  $n_O \rightarrow \sigma_{N-H}^*$  orbital interaction. An additional transient secondary N501(S)–K353 hydrogen bond described by a  $n_O \rightarrow \sigma_{N-H}^*$  orbital interaction is also present in the wild type. Computational evidence<sup>[38]</sup> indicates that the G502–K353  $n_O \rightarrow \sigma_{N-H}^*$  interaction

persists in the mutated virus and that the secondary N501(S)–K353 hydrogen bond is replaced by a primary Y501(S)–K353  $n_O \rightarrow \sigma_{N-H}^*$  hydrogen bond.

Our calculations indicate that in the Y501(S)–K353 contact, a collection of  $\pi_{aromatic} \rightarrow \sigma_{N-H}^*$  interactions, facilitated by the larger size of tyrosine and by the presence of the aromatic ring, emerge after amino acid replacement. This observation rationalizes a considerable increase in the complex binding energy reported elsewhere<sup>[38–40]</sup> and the  $\approx 17.1$  kcal/mol increase in the interaction energy of the extended pairs as a consequence of the mutation. Table 2 also reveals the molecular source of this mutation induced stabilization of the virus–cell complex: two factors work concomitantly to yield a larger stabilizing energy in the mutated virus, a synergistic action between an increase of the electron density at bond critical points by about an entire order of magnitude and an increase of  $\approx 25\%$  in the virial ratio on one hand, and a sensible reduction of the bond degree parameter in the direction of negative values, on the other. Thus, under a molecular perspective, the net result of the N501Y mutation is that the collection of new interactions enhances the ability of the mutated virus to attach to host cells.

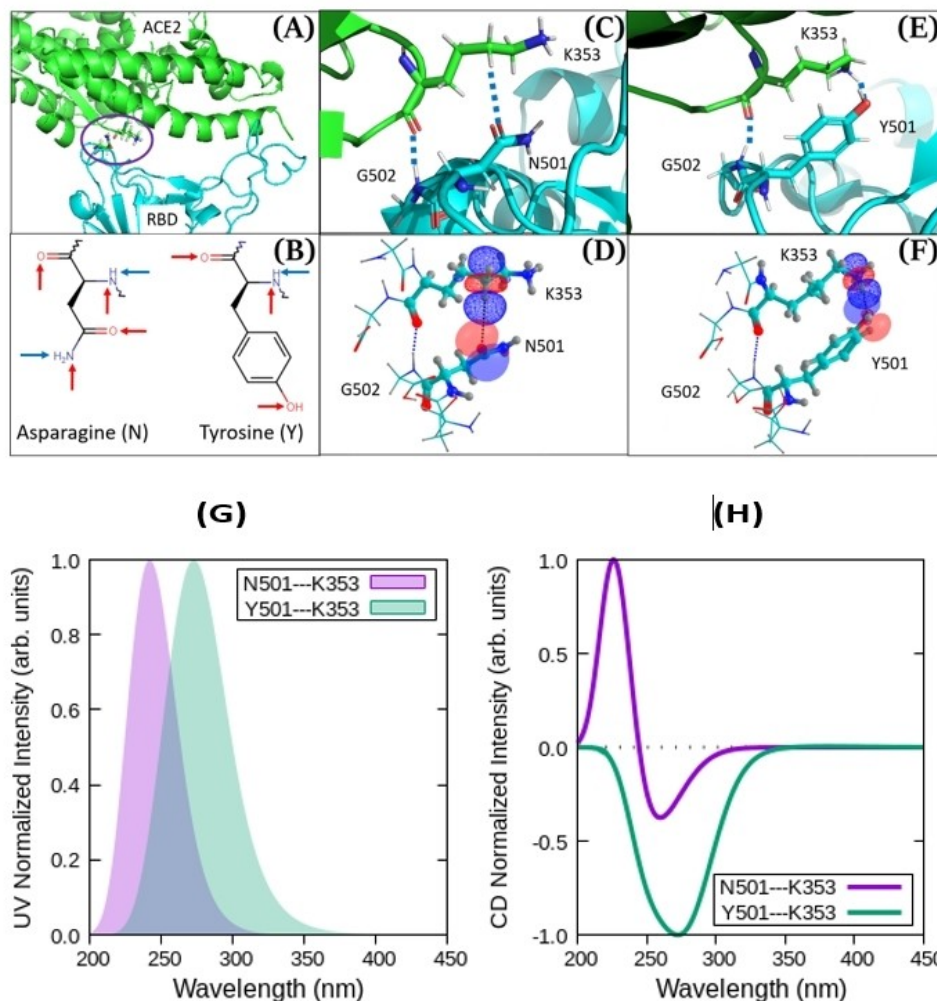
The N501Y mutation leads to a blue-shifted  $\Delta\lambda_{max} \approx 39$  nm due to the replacement of the N501(S)–K353 secondary HB by a similar Y501(S)–K353 HB and to the newly added aromatic ring which originates fresh  $\pi \rightarrow \pi^*$  transitions. Notice that these transitions are more energetic than similar transitions in typical aromatic compounds because this mutation also affects the occupied aromatic orbitals due to the collection of  $\pi_{aromatic} \rightarrow \sigma_{N-H}^*$  interactions. Moving to ECD spectra, the signal of the mutated virus is blue-shifted, as it is expected from UV signals. In addition, the entire ECD spectrum is now negative with a band appearing at about 272 nm, thus reflecting the newly established intermolecular interactions and their consequences on magnitude and orientation of electric and magnetic transition dipoles.<sup>[32–36]</sup>

### 2.3. B.1.351, $\beta$

First found in South Africa and reported by Tegally et al.,<sup>[41]</sup> this variant is also known as 20H or 501Y.V2. Outside of the RBD, this variant has one mutation in the spike protein, one mutation in the nucleocapsid, and one deletion in ORF1a. Within the RBD of the spike protein, asparagine→tyrosine N501Y, lysine→asparagine K417N, and glutamic acid→lysine E484K mutations are found, with both N501Y and E484K occurring in the binding motif. The effects of the N501Y mutation were analyzed in section 2.2. See section 4 for details of the preparation of the extended dimers.

In the wild-type SARS-COV-2, a very strong salt bridge was found to be responsible for the virus–cell contact,<sup>[14,15,39]</sup> dictated by a  $n_O \rightarrow \sigma_{N-H}^*$  orbital interaction between a  $CO_2^-$  group in D30(ACE2) and a  $-NH_3^+$  group in K417(S). Asparagine has an uncharged  $-NH_2$  group and is considerably smaller than the  $-NH_3^+$  charged lysine (Figure 3), then, the K417N replacement eliminates the positive charge and takes the interacting groups apart, effectively replacing the salt bridge with a very





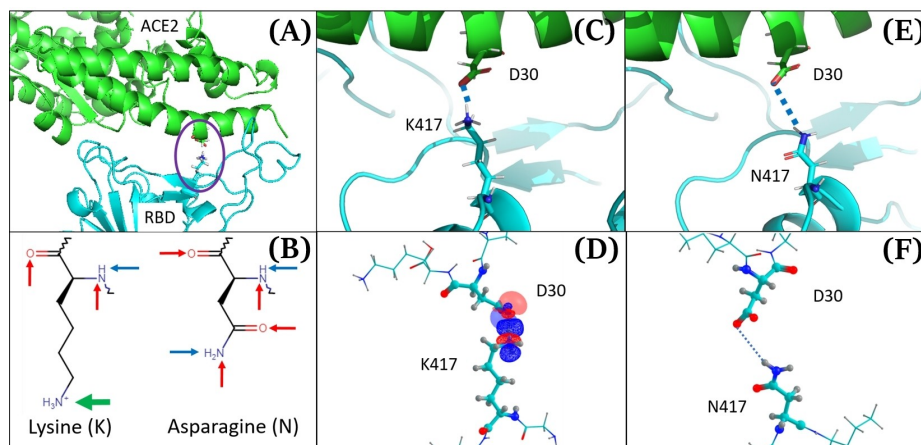
**Figure 2.** N501Y mutation. A snapshot of the dynamics of the RBD/ACE2 complex<sup>[15]</sup> is shown in panel (A) and the possible acceptor (red arrows) and donor (blue arrows) sites for hydrogen bonding before and after mutation are shown in panel (B). A close up of the G502, N501 (RBD spike) and K353(ACE2) positions is shown in panel (C) and the corresponding orbitals interactions for the transient N501(S) K353(ACE2) hydrogen bond are shown in panel (D). Y501 (S) K353 interactions in the mutated virus, taken from the crystal structures reported by Zahradnik and co-workers,<sup>[31]</sup> are shown in panel (E) and associated orbital interactions (only one  $\pi_{aromatic} \rightarrow \sigma_{N-H}^*$  is shown) in panel (F). The G502(S) K353(ACE2) interaction is preserved after the mutation. The computed UV and CD absorption spectra are provided in (G) and (H), respectively. Spectra of the interacting fragments involving the wild type are pictured in purple, spectra of the interacting fragments involving the mutated virus are depicted in green. Level: B3LYP/aug-cc-pVDZ/AMBER.

weak, long distance interaction according to the bond descriptors in Table 2.

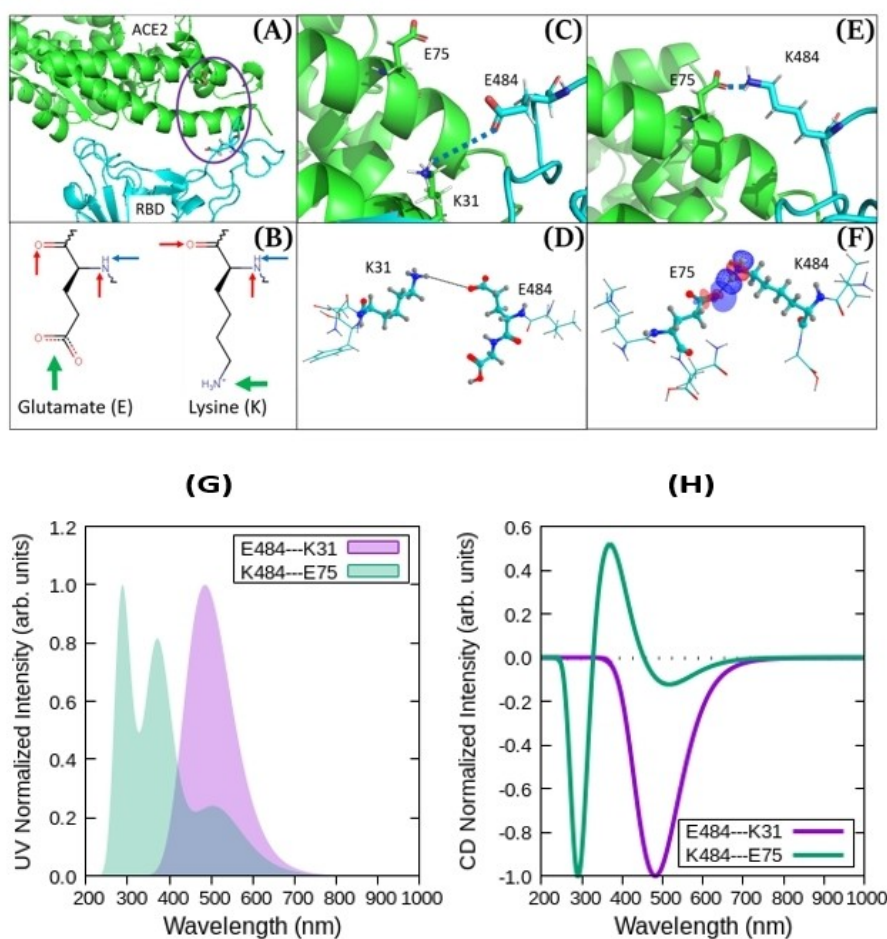
This observation rationalizes the reduction in interaction energy when replacing the K417(S) D30(ACE2) contact by N417(S) D30(ACE2) after this particular mutation, as reported in previous works.<sup>[39]</sup> Notwithstanding this particular reduction in interaction energy, eliminating the K417(S) D30(ACE2) salt bridge confers structural freedom to the RBD (again, the smaller, neutral asparagine is critical to this point) so that the other virus cell contacts become stronger.<sup>[42]</sup> Indeed, facilitated by this structural freedom, the E484K mutation (Figure 4) eliminates a relatively weak E484(s) K31(ACE2) interaction<sup>[39]</sup> and originates one K484(S) E75(ACE2) salt bridge,<sup>[42]</sup> which according to Table 2, leads to the strongest contact among all mutations studied here (even stronger than the hydrogen bond in the reference water dimer), described by a  $n_O \rightarrow \sigma_{N-H}^*$  orbital

interaction from the  $-\text{CO}_2^-$  group in E75 and a  $-\text{NH}_3^+$  group in K484. Thus, the ability for virus attachment is substantially enhanced. In addition to the new salt bridge in the bonding motif, the chemical environment for the newly added lysine is negatively charged in the virus cell complex, thus further favoring this mutation.<sup>[43]</sup>

In summary, for the B.1.351 variant, despite the specific bonding descriptors listed in Table 2 for the K484 E75 replacement in the mutated virus being weaker than for the K417 D30 interaction in the wild type, the additional N501Y mutation and the conformational freedom gained by the elimination of the K417 D30 salt bridge, result in a larger virus cell affinity. There is a subtle point hidden within this analysis: it has been argued that the evolution pressure in the molecular scale favors mutations that improve the ability of the virus to find available sites for hydrogen bonding.<sup>[15]</sup> In fact, the E484K mutation



**Figure 3.** K417N mutation. A snapshot of the dynamics of the RBD/ACE2 complex<sup>[15]</sup> is shown in panel (A) and the possible acceptor (red arrows) and donor (blue arrows) sites for hydrogen bonding before and after the mutation are shown in panel (B). A close up of the K417(RBD spike) and D30(ACE2) positions leading to a strong salt bridge is shown in panel (C) and the associated donor-acceptor orbital interactions are shown in panel (D). A very weak, long distance interaction (Table 2) and (E), (F) was found in the mutated virus, claimed to be unfavorable in other works.<sup>[39]</sup>



**Figure 4.** E484K mutation. A snapshot of the dynamics of the RBD/ACE2 complex<sup>[15]</sup> is shown in panel (A) and the possible acceptor (red arrows) and donor (blue arrows) sites for hydrogen bonding are shown in panel (B). A close up of the S477(RBD spike) and S19(ACE2) positions with no persistent or transient interaction is shown in panel (C) and the associated donor-acceptor orbital interactions are shown in panel (D). Hydrogen bonding in the mutated virus are shown in panel (E) and associated orbital interactions are shown in panel (F). The Computed UV and CD absorption spectra are provided in (G) and (H), respectively. Spectra of the interacting fragments involving the wild type are pictured in purple, spectra of the interacting fragments involving the mutated virus are depicted in green. Level: B3LYP/aug-cc-pVDZ/AMBER.

allows the virus to find a much more favorable attachment site at E75, a helix position in ACE2 far from the initial K31 helical position. Calculated UV-Vis and CD spectra with environmental effects for the E484K and K417N mutations are shown in Figures 4 and 5 respectively.

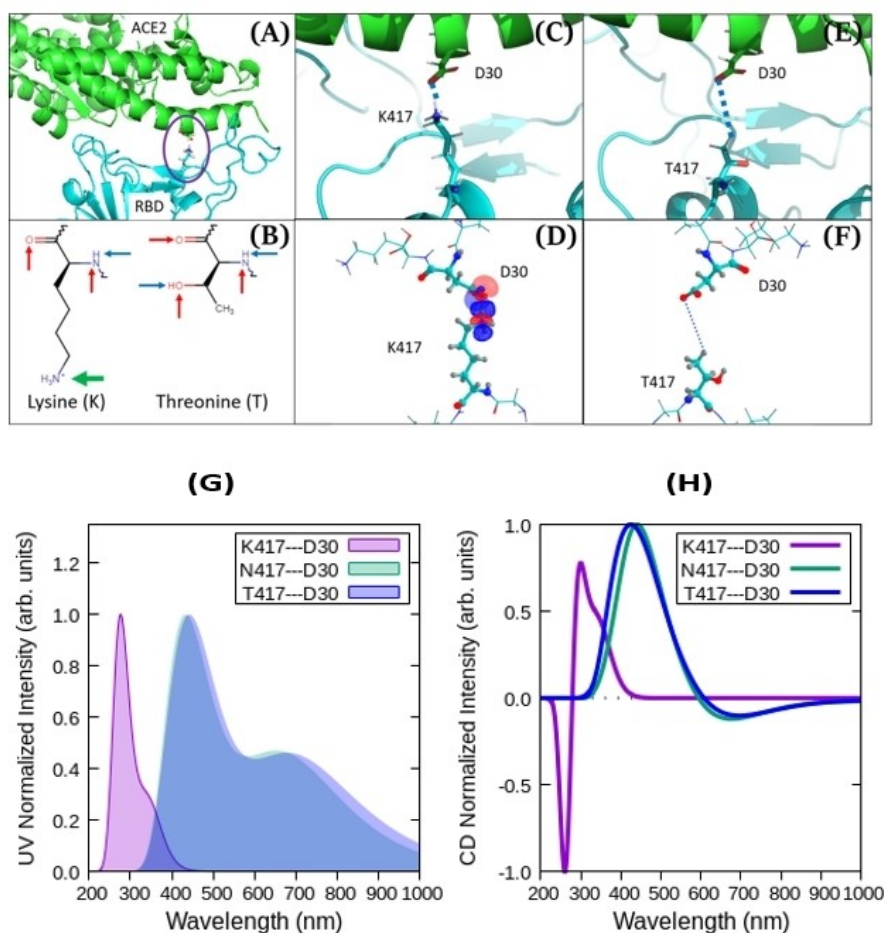
#### 2.4. P.1, $\gamma$

This variant is also referred to as B.1.1.28.1, 20J or 501Y.V3 and was first found in Brazil. It has initially been characterized by Faria et al.<sup>[44]</sup> as a likely case of convergent evolution with the B.1.351 variant.<sup>[44,45]</sup> Outside the RBD, P.1 has undergone one mutation in the nucleocapsid as well as the same deletion in ORF1a as in B.1.351. Within the RBD, the B.1.351 and P.1 variants have both N501Y and E484K mutations in common, the only difference in this domain is that P.1 may have either the same K417N replacement as in the B.1.351 variant or a K417T mutation. Since threonine and asparagine have neutral groups as opposed to the charged  $-\text{NH}_3^+$  in lysine, and since both are

quite smaller than lysine (Figures 3, 5), both K417N and K417T mutations have identical consequences in the ability of both variants to bind to ACE2. In fact, as analyzed above, the dominant contribution to the binding energies in both cases arises from the newly formed K484(S)--E75(ACE2) salt bridge. Fittingly, not only are B.1.351 and P.1 among the most contagious variants, but have been proven to have similar infecting capacity.<sup>[31]</sup> All spectral changes in this variant have already been dissected above, except for the K417T mutation, which, as expected, has the same effects as the K417N mutation, as shown in Figure 5.

#### 2.5. P.2, $\zeta$

Another variant first found in Brazil, also known as B.1.1.28.2, 20B or S.484K, which was first characterized by Voloch et al.<sup>[46]</sup> Outside the RBD of the spike protein, this variant exhibits two mutations in the nucleocapsid and two mutations in the ORF1a, and more importantly, does not contain N501 or K417



**Figure 5.** K417T mutation. A snapshot of the dynamics of the RBD/ACE2 complex<sup>[15]</sup> is shown in panel (A) and the possible acceptor (red arrows) and donor (blue arrows) sites for hydrogen bonding are shown in panel (B). A close up of the S477(RBD spike) and S19(ACE2) positions with no persistent or transient interaction is shown in (C) and the associated donor→acceptor orbital interactions are shown in panel (D). Hydrogen bonding in the mutated virus are shown in panel (E) and associated orbital interactions are shown in (F). The Computed UV and CD absorption spectra are provided in (G) and (H), respectively. The interacting fragments involving the mutated virus (K417N(T) mutations) are depicted in blue and green. Spectra of the interacting fragments involving the wild type are also provided for comparison and pictured in purple. Level:B3LYP/aug-cc-pVDZ/AMBER.

mutations. In the bonding motif of the RBD, a glutamate→lysine E484K mutation defines this variant. In the wild-type virus, the E484(S) glutamate interacts via very weak, long distance contacts with a K31 lysine in ACE2 (see Figure 4). The average distance between E484(S)⋯K31(ACE2) residues is 5.53 Å according to previously reported MD simulations.<sup>[15]</sup> At shorter distances this would characterize a strong  $-\text{CO}_2^- \cdots {}^+\text{HN}-$  salt bridge. The substitution of glutamate by lysine at the 484 position in the spike protein leads to a substantial change, that is, to the formation of the above analyzed K484(S)⋯E75(ACE2) salt bridge.<sup>[42]</sup> This results in a considerable increase of the Spike⋯ACE2 affinity, to the point that this mutation alone, with no simultaneous mutations in the RBD, is enough to stabilize the mutated virus⋯cell complex by more than 10 kcal/mol, when compared to the wild type,<sup>[43]</sup> however, as shown by Khan et al.,<sup>[47]</sup> this affinity it is still lower than the P.1, B.1.351 cases, which contain the additional N501Y and K417N mutations.

### 3. Conclusions

To summarize, we offer a dissection of the explicit residue to residue intermolecular interactions resulting from the specific S477N, N501Y, K417N, K417T, E484K mutations in the receptor binding domain of the spike protein in the wild-type SARS-COV-2 virus that have led to the highly concerning B.1.160, B.1.1.7 ( $\alpha$ ), B.1.351 ( $\beta$ ), P.1 ( $\gamma$ ), and P.2 ( $\zeta$ ) variants. A generalized qualitative observation is that the interaction energies listed in Table 2 are in perfect agreement with the available basic reproduction numbers  $R_0$ , reported to be 2.5<sup>[48]</sup> for the wild type and 4.7, 4.8, and 4.9 for the  $\alpha$ ,  $\beta$ ,  $\gamma$  variants, respectively, in the sense that higher reproduction numbers are associated to stronger interactions. We condensate our finding for each individual mutation in the following list:

1. The E484K mutation leads to a very strong K484(S)⋯E75(ACE2) salt bridge described by an  $\text{NH}_3^+ \cdots {}_2\text{OC}$  interaction arising from a  $n_{\text{O}} \rightarrow \sigma_{\text{N-H}}^*$  orbital interaction. Bonding descriptors of the strength of the interaction change in going from the wild type to the mutated virus as follows: -IE: 0.6→67.4 kcal/mol,  $10^2\rho(r_c)$ : 0.02→3.13 a.u., virial ratio: 0.52→0.92, bond degree parameter: 23.09→6.78 a.u. All descriptors indicate a much stronger affinity between the mutated virus and the ACE2 receptor. This mutation is present in the most contagious B.1.351, P.1, P.2 variants.
2. The N501Y mutation leads to a Y484(S)⋯K353(ACE2) contact described by a collection of  $\pi_{\text{aromatic}} \rightarrow \sigma_{\text{N-H}}^*$  orbital interactions, second in strength only to the interactions arising because of the E484K replacement. Bonding descriptors of the strength of the interaction change in going from the wild type to the mutated virus as follows: -IE: 12.7→29.8 kcal/mol,  $10^2\rho(r_c)$ : 0.32→2.71 a.u., virial ratio: 0.71→0.89, bond degree parameter: 20.66→9.72 a.u. This mutation is present in the B.1.351, P.1, B.1.1.7 variants
3. The S477N mutation leads to two N477(S)⋯S19(ACE2) persistent primary  $n_{\text{O}} \rightarrow \sigma_{\text{N-H}}^*$  hydrogen bonds. Bonding descriptors of the strength of the interaction change in

- going from the wild type to the mutated virus as follows: -IE: 3.9→13.1 kcal/mol,  $10^2\rho(r_c)$ : 0.02→2.88 a.u., virial ratio: 0.50→0.94, bond degree parameter: 25.17→4.63 a.u. This mutation is present in the most contagious B.1.160 variant
4. The K417N, K417T mutations lead to long range weak N417(S)⋯D30(ACE2) and T417(S)⋯D30(ACE2) interactions. This mutation is present in the B.1.351 and P.1 variants UV-absorption and ECD spectral signatures clearly indicate that mutations result in relevant changes in the electronic structure, and particularly, if a given mutation replaces one weak interaction by another (i.e. two HBs), the absorption peaks are red shifted by  $\approx 76$  nm. When the mutation involves the addition of an aromatic ring,  $\lambda_{\text{max}}$  is blue shifted by  $\approx 39$  nm. Also, substantially different ECD spectral patterns are predicted.

Our results provide a molecular perspective of the virus⋯cell interaction problem, elucidating and quantifying the strengths of all residue-to-residue contacts, thus offering the basics to rationalize the observed enhanced ability of the variants to bind to host cells. We clearly establish a direct link relating the strength of intermolecular interactions between the amino acids in the RBD of the spike protein and in the ACE2 receptor, and the initial virus⋯cell molecular recognition and attachment, i.e. the events that trigger the infectious cycle. Our studies provide a method based solely in either crystallographic or MD structural data to rapidly assess, ahead of the corresponding experimental measurements of binding energies and dissociation constants, the comparative infectious power of new virus variants that have mutations in the RBD of the spike protein by obtaining bonding descriptors with the ACE2 receptor.

## Computational Methods

### Wild type

One frame was randomly chosen from the late stages of the Molecular Dynamics trajectories of the wild type RBD/ACE2 complex reported by Gómez et al.<sup>[15]</sup> The frame was chosen so that the distances for all intermolecular contacts leading to persistent hydrogen bonds and salt bridges correspond with the average separations during the Dynamics. From the chosen frame, we extracted all extended interacting pairs explicitly involved in the virus⋯cell contacts. By extended interacting pairs we mean that for the NBO and QTAIM calculations, in order to account for changes in the local chemical environment, we included in our calculations not only the residues in direct contact, but also their nearest neighbors as if in their terminal configurations for a total of three amino acids in the spike protein and three amino acids in ACE2 (see Figures 1–5 in the main text for the specific geometries). The corresponding Cartesian coordinates are provided in the SI.

### Mutations

The structures for the complexes resulting from the interactions between the mutated virus and the ACE2 receptor were obtained in the following manner:

1. **S477N**: The extended pairs were extracted from a snapshot of the MD simulations carried out by Singh and co-workers.<sup>[29]</sup> The



- frame was chosen so that the N477(S)---S19(ACE2) distance corresponds with the average separation during the Dynamics.
- N501Y, K417T, E484K:** Professor Dong-Qing Wei from the Shanghai Jiao Tong University provided us with the required snapshots from their MD trajectories published in reference [47].
  - K417N:** Since there are no crystallographic data nor MD trajectories available, we substituted the amino acids involved in the wild type and kept the fragments in the configurations they had in the interacting system (this is more accurate to understand the virus---cell bonding interactions than re-optimizing the isolated pairs).

### Interaction Energies

We proceeded to compute quantum interaction energies for every mutated virus---cell region via single point energy calculations under the DLPNO-CCSD(T)/aug-cc-pVDZ model chemistry<sup>[49–52]</sup> in the ORCA suite of programs, version 4.0.1.2.<sup>[53]</sup>

### Bonding Interactions

The intermolecular interactions responsible for the initial virus---cell recognition and attachment for all the title variants were dissected using the tools provided by the natural bond orbitals (NBO7.0<sup>[54–57]</sup>) and by the quantum theory of atoms in molecules (QTAIM<sup>[58–60]</sup>) as implemented in AIMAll<sup>[61]</sup>.

### Absorption Spectra

The effects of virus mutations in the virus---cell interactions were analyzed by means of UV and CD absorption spectra calculated on the highly idealized fragments discussed above, enlarging the interaction region by including everything within a sphere of 12 °Å centered on each of the two interacting amino acids. It has been shown that B3LYP suffices to accurately calculate absorption spectra of the RBD(S)---ACE2 interacting fragments in the specific case of SARS-COV-2 without the need to add dispersion corrections.<sup>[15]</sup> Accordingly, in this work, all TD-DFT calculations were carried out using the B3LYP/aug-cc-pVDZ model chemistry.<sup>[51,52,62,63]</sup> We use QM/MM Electrostatic Embedding<sup>[64]</sup> by including in the quantum region the amino acid in ACE2 in direct contact with the amino acids involved in the mutation, and the classical region defined by the neighboring residues in both fragments, which was modelled via the Amber force field.<sup>[65]</sup> The first 20 excited states were considered during the TD-DFT QM/MM calculations in each case. In order to account for the systematic error in our chosen functional, we shifted all spectra by  $-0.7$  eV, as is needed to match the experimental maximum from the computed spectra of tyrosine.<sup>[66,67]</sup> All Spectra were convoluted with Gaussian line shapes with full width half maximum (FWHM) of 0.6 eV. All QM/MM calculations were carried out with Gaussian16.<sup>[68]</sup>

### Supporting Information

Cartesian coordinates for the interacting dimers are provided.

### Acknowledgements

Partial funding for this project from H2020-MSCA-ITN-2017 European Training Network "Computational Spectroscopy In Natural

sciences and Engineering" (COSINE), grant number 765739 is acknowledged. Internal support from Universidad de Antioquia via "Estrategia para la sostenibilidad" is also acknowledged. The authors are in debt to the following Professors and their research groups who graciously provided their research data: Christian Gruber at Medical University of Vienna, MD trajectories for the S477N mutation. Gideon Schreiber at the Weizmann Institute of Science, Israel, crystallographic data for the N501Y variant. Dong-Qing Wei from the Shanghai Jiao Tong University, snapshots from MD trajectories of the N501Y, K417T, E484K mutations.

### Conflict of Interest

The authors declare no conflict of interest.

- [1] B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, et al., *Cell* **2020**, *182*, 812–827.
- [2] J. W. Drake, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4171–4175.
- [3] Z. Gu, Q. Gao, E. A. Faust, M. A. Wainberg, *J. Gen. Virol.* **1995**, *76*, 2601–2605.
- [4] L. Moutouh, J. Corbeil, D. D. Richman, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6106–6111.
- [5] J. T. Blackard, D. E. Cohen, K. H. Mayer, *Clin. Infect. Dis.* **2002**, *34*, 1108–1114.
- [6] D. Zella, M. Giovanetti, F. Benedetti, F. Unali, S. Spoto, M. Guarino, S. Angeletti, M. Ciccozzi, *J. Med. Virol.* **2021**, <https://doi.org/10.1002/jmv.27196>.
- [7] T. Phan, *Infect. Genet. Evol.* **2020**, *81*, 104260.
- [8] V. Vinson, *Science* **2020**, *367*, 1438–1439.
- [9] J. Huo, A. Le Bas, R. R. Ruza, H. M. Duyvesteyn, H. Mikolajek, T. Malinauskas, T. K. Tan, P. Rijal, M. Dumoux, P. N. Ward, et al. *Nature Struc. Molec. Biol.* **2020**, *27*, 846–854.
- [10] K. Houser, K. Subbarao, *Cell Host Microbe* **2015**, *17*, 295–300.
- [11] H. Gu, Q. Chen, G. Yang, L. He, H. Fan, Y.-Q. Deng, Y. Wang, Y. Teng, Z. Zhao, Y. Cui, et al. *Science* **2020**, *369*, 1603–1607.
- [12] CDC, Emerging SARS-CoV-2 variants. 2021; <https://stacks.cdc.gov/view/cdc/100655>, Center for Disease Control and Prevention, Webpage, Accessed April 2, 2021.
- [13] Y. Yuan, D. Cao, Y. Zhang, J. Ma, J. Qi, Q. Wang, G. Lu, Y. Wu, J. Yan, Y. Shi, et al., *Nat. Commun.* **2017**, *8*, 15092.
- [14] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K.-Y. Yuen, et al., *Cell* **2020**, *181*, 894–904.
- [15] S. A. Gomez, N. Rojas-Valencia, S. Gomez, F. Egidi, C. Cappelli, A. Restrepo, *ChemBioChem* **2021**, *22*, 724–732.
- [16] Y. Wang, M. Liu, J. Gao, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 13967–13974.
- [17] Covariants, Overview of SARS-CoV-2 variants and mutations, 2020–2021; <https://covariants.org>, Emma Hodcroft, Covariants, Webpage, <https://covariants.org>. Accessed Aug. 3, 2021.
- [18] A. Rambaut, E. C. Holmes, A. O'Toole, V. Hill, J. T. McCrone, C. Ruis, L. du Plessis, O. G. Pybus, *Nat. microbial.* **2020**, *5*, 1403–1407.
- [19] N. Rojas-Valencia, S. Gomez, S. Montillo, M. Manrique-Moreno, C. Cappelli, C. Hadad, A. Restrepo, *J. Phys. Chem. B* **2020**, *124*, 79–90.
- [20] N. Rojas-Valencia, S. Gomez, F. Nuñez-Zarur, C. Cappelli, C. Hadad, A. Restrepo, *J. Phys. Chem. B* **2021**, *36*, 10383–10391.
- [21] N. Rojas-Valencia, C. Ibarqu'en, A. Restrepo, *Chem. Phys. Lett.* **2015**, *635*, 301–305.
- [22] S. Gomez, H. Ramirez-Malule, W. Cardona-G, E. Osorio, A. Restrepo, *J. Phys. Chem. A* **2020**, *124*, 9413–9426.
- [23] S. Llano, S. Gomez, J. Londoño, A. Restrepo, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3752–3760.
- [24] J. C. Coa, W. Cardona-Galeano, *A. Phys. Chem. Chem. Phys.* **2018**, *20*, 20382–20390.
- [25] C. Giraldo, S. Gomez, F. Weinhold, A. Restrepo, *ChemPhysChem* **2016**, *17*, 2022–2034.
- [26] E. Rengifo, S. Gomez, J. C. Arce, F. Weinhold, A. Restrepo, *Comp. Theo. Chem.* **2018**, *1130*, 58–62.

- [27] K. Wiberg, *Tetrahedron* **1968**, *24*, 1083–1096.
- [28] E. B. Hodcroft, M. Zuber, S. Nadeau, K. H. Crawford, J. D. Bloom, D. Veessler, T. G. Vaughan, I. Comas, F. G. Candelas, T. Stadler, et al. *MedRxiv* **2020**, <https://doi.org/10.1101/2020.10.25.20219063>.
- [29] A. Singh, G. Steinkellner, K. K'ochl, K. Gruber, C. C. Gruber, *Sci. Rep.* **2021**, *11*, 4320.
- [30] J. Chen, R. Wang, M. Wang, G.-W. Wei, *J. Mol. Biol.* **2020**, *432*, 5212–5226.
- [31] J. Zahradnik, S. Marciano, M. Shemesh, E. Zoler, J. Chiaravalli, B. Meyer, O. Dym, N. Elad, G. Schreiber, *Nat. Microbiol.* **2021**, *6*, 1188–1198.
- [32] L. Rosenfeld, *Zeitschr. Phys.* **1929**, *52*, 161–174.
- [33] N. Berova, P. L. Polavarapu, K. Nakanishi, R. W. Woody, *Comprehensive Chiroptical Spectroscopy: Instrumentation, Methodologies, and Theoretical Simulations*, Wiley, **2011**, Vol. 1.
- [34] N. Berova, P. L. Polavarapu, K. Nakanishi, R. W. Woody, *Comprehensive Chiroptical Spectroscopy: Applications in Stereochemical Analysis of Synthetic Compounds, Natural Products, and Biomolecules*, Wiley, **2012**, Vol. 2.
- [35] M. Scott, D. R. Rehn, S. Coriani, P. Norman, A. Dreuw, *J. Chem. Phys.* **2021**, *154*, 064107.
- [36] L. Barron, *Molecular Light Scattering and Optical Activity*; Cambridge University Press, 2004.
- [37] R. Challen, E. Brooks-Pollock, J. M. Read, L. Dyson, K. Tsaneva-Atanasova, L. Danon, *BMJ* **2021**, *372*, n579.
- [38] N. A. Murugan, P. S. Javali, C. J. Pandian, M. A. Ali, V. Srivastava, J. Jeyaraman, *bioRxiv* **2021**, <https://doi.org/10.1101/2021.01.25.428190>.
- [39] B. O. Villoutreix, V. Calvez, A.-G. Marcelin, A.-M. Khatib, *Int. J. Mol. Sci.* **2021**, *22*, 1695.
- [40] C. Watanabe, Y. Okiyama, S. Tanaka, K. Fukuzawa, T. Honma, *Chem. Sci.* **2021**, *12*, 4722–4739.
- [41] H. Tegally, E. Wilkinson, R. J. Lessells, J. Giandhari, S. Pillay, N. Msomi, K. Mlisana, J. N. Bhiman, A. von Gottberg, S. Walaza, et al., *Nat. Med.* **2021**, *27*, 440–446.
- [42] G. Nelson, O. Buzko, P. R. Spilman, K. Niazi, S. Rabizadeh, P. R. Soon-Shiong, *bioRxiv* **2021**, <https://doi.org/10.1101/2021.01.13.426558>.
- [43] W. B. Wang, Y. Liang, Y. Q. Jin, J. Zhang, J. G. Su, Q. M. Li, *bioRxiv* **2021**, DOI: <https://doi.org/10.1101/2021.02.17.431566>.
- [44] N. R. Faria, I. M. Claro, D. Candido, L. Moyses Franco, P. S. Andrade, T. M. Coletti, C. A. Silva, F. C. Sales, E. R. Manuli, R. S. Aguiar et al., *Virological* **2021**, Available at: <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manauas-preliminary-findings/586>, Accessed Mar. 25, 2021.
- [45] F. Naveca, V. Nascimento, V. Souza, A. Corado, F. Nascimento, G. Silva, A. Costa, D. Duarte, K. Pessoa, L. Goncalves et al., *Virological* **2021**, Available at: <https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585>, Accessed Mar. 25, 2021.
- [46] C. M. Voloch, R. da Silva Francisco, L. G. de Almeida, C. C. Cardoso, O. J. Brustolini, A. L. Gerber, A. P. d. C. Guimarães, D. Mariani, R. M. da Costa, O. C. Ferreira, et al. *J. Virol.* **2021**, DOI: 10.1128/JVI.00119–21.
- [47] A. Khan, T. Zia, M. Suleman, T. Khan, S. S. Ali, A. A. Abbasi, A. Mohammad, D.-Q. Wei, *J. Cell. Physiol.* **2021**, *236*, 7045–7057.
- [48] G. Lippi, B. M. Henry, *Ann. Trans. Med.* **2021**, *9*.
- [49] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 034106.
- [50] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
- [51] R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- [52] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [53] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- [54] E. D. Glendening, J. K. Badenhop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou, C. R. Landis, F. Weinhold, NBO 7.0. 2018; Theoretical Chemistry Institute, University of Wisconsin, Madison, WI.
- [55] A. E. Reed, F. Weinhold, *J. Chem. Phys.* **1983**, *78*, 4066–4073.
- [56] F. Weinhold, C. R. Landis, *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH, Hoboken NJ, 319pp, **2012**.
- [57] E. D. Glendening, C. R. Landis, F. Weinhold, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 1–42.
- [58] R. Bader, *Atoms in Molecules: A Quantum Theory*; Oxford Univ. press Oxford, 1990.
- [59] E. Espinosa, I. Alkorta, J. Elguero, E. Molins, *J. Chem. Phys.* **2002**, *117*, 5529–5542.
- [60] S. J. Grabowski, *Chem. Rev.* **2011**, *111*, 2597–2625.
- [61] T. Keith, AIMALL (version 19.10.12). 2019; TK Gristmill Software, Overland Park KS, USA, [aim.tkgristmill.com](http://aim.tkgristmill.com).
- [62] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
- [63] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [64] S. Dapprich, I. Kom'aromi, K. Byun, K. Morokuma, M. J. Frisch, *J. Mol. Struct.: THEOCHEM* **1999**, *461–462*, 1–21.
- [65] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- [66] D. Creed, *Photochem. Photobiol.* **1984**, *39*, 563–575.
- [67] L. H. Fornander, B. Feng, T. Beke-Somfai, B. Nord'en, *J. Phys. Chem. B* **2014**, *118*, 9247–9257.
- [68] Gaussian 16 Revision B.01. 2016, M. J. Frisch et al., Gaussian Inc., Wallingford CT.

---

Manuscript received: August 3, 2021

Revised manuscript received: September 14, 2021

Accepted manuscript online: September 16, 2021

Version of record online: October 8, 2021