

# PCCCP

Physical Chemistry Chemical Physics

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: V. Binette, S. Côté, M. Hadded, P. T. Nguyen, S. Bélanger, R. Gaudreault, S. Bourgault, C. Ramassamy and N. Mousseau, *Phys. Chem. Chem. Phys.*, 2021, DOI: 10.1039/D1CP01790J.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Cite this: DOI: 00.0000/xxxxxxxxxx

# Corilagin and 1,3,6-Tri-O-galloyl- $\beta$ -D-glucose: Potential inhibitors of SARS-CoV-2 Variants<sup>†</sup>

Vincent Binette,<sup>a</sup> Sébastien Côté,<sup>a,b</sup> Mohamed Haddad,<sup>c</sup> Phuong Trang Nguyen,<sup>d</sup> Sébastien Bélanger,<sup>e</sup> Steve Bourgault,<sup>d</sup> Charles Ramassamy,<sup>c</sup> Roger Gaudreault,<sup>a,‡</sup> and Normand Mousseau<sup>a,‡</sup>Received Date  
Accepted Date

DOI: 00.0000/xxxxxxxxxx

The COVID-19 disease caused by the virus SARS-CoV-2, first detected in December 2019, is still emerging through virus mutations. Although almost under control in some countries due to effective vaccines that are mitigating the worldwide pandemic, the urgency to develop additional vaccines and therapeutic treatments is imperative. In this work, the natural polyphenols corilagin and 1,3,6-tri-O-galloyl- $\beta$ -D-glucose (TGG) are investigated to determine the structural basis of inhibitor interactions as potential candidates to inhibit SARS-CoV-2 viral entry into target cells. First, the therapeutic potential of the ligands are assessed on the ACE2/wild-type RBD. We first use molecular docking followed by molecular dynamics, to take into account the conformational flexibility that plays a significant role in ligand binding and that cannot be captured using only docking, and then analyze more precisely the affinity of these ligands using MMPBSA binding free energy. We show that both ligands bind to the ACE2/wild-type RBD interface with good affinities which might prevent the ACE2/RBD association. Second, we confirm the potency of these ligands to block the ACE2/RBD association using a combination of surface plasmon resonance and biochemical inhibition assays. These experiments confirm that TGG and, to a lesser extent, corilagin, inhibit the binding of RBD to ACE2. Both experiments and simulations show that the ligands interact preferentially with RBD, while weak binding is observed with ACE2, hence, avoiding potential physiological side-effects induced by the inhibition of ACE2. In addition to the wild-type RBD, we also study numerically three RBD mutations (E484K, N501Y and E484K/N501Y) found in the main SARS-CoV-2 variants of concerns. We find that corilagin could be as effective for RBD/E484K but less effective for the RBD/N501Y and RBD/E484K-N501Y mutants, while TGG strongly binds at relevant locations to all three mutants, demonstrating the significant interest of these molecules as potential inhibitors for variants of SARS-CoV-2.

## 1 Introduction

The COVID-19 disease, first detected in late December 2019 in Wuhan China, has quickly spread worldwide leading to 145 mil-

lion reported cases and 3.1 million deaths as of April 2021<sup>1</sup>. This disease is caused by the virus SARS-CoV-2 of the coronavirus family, which is characterized by a lipid envelope sealing a genome made of a single positive RNA strand. To replicate its genome, SARS-CoV-2 has to penetrate and hijack the translation center of a host cell.

To do so, SARS-CoV-2 uses its *Spike* protein to bind to the angiotensin-converting enzyme 2 (ACE2)<sup>2,3</sup>, a receptor found at the cell surface in a wide variety of human organs, such as the heart, the liver, the kidneys and alveoli<sup>4</sup> and that plays a decisive regulating function in the renin-angiotensin system (RAS)<sup>5</sup>. More specifically, the crystal structure of the *Spike* protein and ACE2 interface<sup>6</sup> shows that the interactions with ACE2 are mediated by the receptor binding motif (RBM) of the receptor binding domain (RBD) of the *Spike* protein.

<sup>a</sup> Département de physique, Université de Montréal, Case postale 6128, succursale Centre-ville, Montréal (QC), H3C 3J7 Canada

<sup>b</sup> Cégep de Saint-Jérôme, 455 Rue Fournier, Saint-Jérôme (QC), J7Z 4V2 Canada

<sup>c</sup> Centre Armand-Frappier Santé Biotechnologie, 531 boulevard des Prairies, Laval (QC), H7V 1B7 Canada

<sup>d</sup> Département de Chimie, Université du Québec à Montréal, 2101 Rue Jeanne-Mance, Montréal (QC), H2X 2J6 Canada

<sup>e</sup> Department of Physics, McGill University, 3600 University Street, Montreal (QC), H3A 2T8 Canada

<sup>†</sup> Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/cXCP00000x/

<sup>‡</sup> Corresponding authors

The SARS-CoV-2 virus is constantly changing due to evolutionary pressure and mutations in its crucial *Spike*-protein were observed all over the world. The B.1.1.7 variant, initially observed in the United Kingdom, includes a mutation at position 501 of the RBD where the asparagine is replaced by a tyrosine (N501Y)<sup>7</sup>. The B.1.351 variant, identified first in South Africa, counts multiple mutations on the RBD including K417N, E484K and N501Y<sup>8</sup>. The B.1.1.28 variant, originating in Brazil, and its descendent, the P.1 mutant, contains multiple mutations on the RBD, including K417T, E484K and N501Y<sup>9</sup>. Among these mutations, the E484K and N501Y mutations, present in the above variants, could be critical as they are located on the RBD of the *Spike* protein and they have been shown experimentally to confer enhanced affinity for ACE2<sup>10</sup>.

Many strategies to prevent virus-induced infection aim at using small molecules to mitigate one (or many) of the steps of the SARS-CoV-2 mechanism of action<sup>11</sup>. On the *Spike* protein, two main regions could potentially be targeted; (1) The RBM of the RBD that directly interacts with ACE2<sup>6</sup> and, although not located directly at the interface, (2) the furin cleavage sites of the *Spike* protein have been shown to significantly affect the binding affinity between the *Spike* protein and ACE2<sup>2,12,13</sup>. In this work, we only focused our attention on the former.

One of the promising approaches considered for reducing the transmission of SARS-CoV-2 is the use of polyphenols because they are natural compounds found in plants and their therapeutic potential is already well documented for different diseases, e.g., neurodegenerative<sup>14–16</sup>, cardiovascular<sup>17</sup>, antihypertensive<sup>18</sup>, cancers<sup>17,19</sup>, HIV<sup>20,21</sup>, and antiviral<sup>22,23</sup> including antiviral drug candidates for SARS-CoV-1 and COVID-19<sup>24,25</sup>. Since the beginning of the actual pandemic, the potential of polyphenols against SARS-CoV-2 has been widely investigated<sup>26,27</sup>. For instance, molecular docking of amentoflavone, a natural compound found in *Ginkgo Biloba*, on SARS-CoV-2 *Spike* protein showed a high binding affinity<sup>28</sup>. Other results from molecular docking and MD simulation on SARS-CoV-2 *Spike* protein identified fisetin, kaempferol and quercetin, all natural compounds found in many fruits and vegetables, as having a high binding affinity and a network of interactions that could disrupt the interaction with ACE2<sup>29</sup>.

Here, we focus on two naturally occurring polyphenols that are promising therapeutic compounds against SARS-CoV-2; corilagin (C<sub>27</sub>H<sub>22</sub>O<sub>18</sub>) and TGG (C<sub>27</sub>H<sub>24</sub>O<sub>18</sub>). Both molecules share a very similar structures; corilagin phenolic rings (R3-R6) are joined compared to TGG, making the former rigid and the latter flexible<sup>30,31</sup>. Both molecules have very low toxicity even at high dosages<sup>32</sup> as well as promising therapeutic properties<sup>33</sup>. For example, corilagin was described as having anti-hypertensive<sup>34</sup>, anti-inflammatory and antioxidant<sup>35</sup> properties. On the other hand, the less studied TGG is closely related to the tetra-TGG molecule, a promising therapeutic compound against SARS-CoV-1<sup>24</sup>.

In this work, we probe in more details the crucial interactions between corilagin/TGG and the *Spike* protein/ACE2 interface. First, we use MD simulations on the ligand-protein complexes to take into account the conformational flexibility that plays a sig-

nificant role in ligand binding and that cannot be captured using only docking<sup>36,37</sup>. Second, we analyze more precisely the affinity of these ligands using MMPBSA binding free energy<sup>38</sup>. These numerical predictions and methodology are validated using experimental tools; Surface Plasmon Resonance (SPR) as well as Binding Inhibitor Assay (ELISA). Finally, we assess numerically the impact of emerging SARS-CoV-2 mutations (E484K, N501Y and E484K/N501Y) of the variants of concerns on the binding affinity of corilagin and TGG to RBD.

## 2 Materials and Methods

We investigate the mechanisms and binding affinity of corilagin and TGG with ACE2 and RBD using a combination of simulations (molecular docking, molecular dynamics and MMPBSA free energy calculations) and experiments (surface plasmon resonance and binding inhibitor assay).

### 2.1 MD simulations

As a first step, we perform 500-ns MD simulation on the ACE2/RBD complex as well as on ACE2 and RBD alone to evaluate their stability and fluctuations. Each system was prepared as follow: (1) the system undergoes an energy minimization step in vacuum using sequentially the steepest descent (SD) and conjugate gradient (CG) algorithms; (2) it is then solvated with explicit water molecules (TIP3P) inserted to fill the dodecahedron box; (3) and ions are added until neutrality; (4) the solvent configurational energy is minimized using sequentially the SD and CG algorithms with all non-hydrogen atoms of the protein kept in place using harmonic restraints; (5) the whole system is equilibrated in the NVT ensemble at 300 K over 10 ns, while maintaining harmonic restraints on non-hydrogen atoms; (6) this is followed by a 10-ns NPT equilibration also with harmonic constraints on non-hydrogen atoms; and (7) a full molecular dynamics (MD) simulation in the NPT ensemble, without any restraint.

All simulations are run with GROMACS v2019.3<sup>39</sup>. The all-atom AMBER14sb forcefield<sup>40</sup> is used for the parameters of the protein. The temperature is kept at 300 K using the Nosé-Hoover thermostat<sup>41,42</sup> with a coupling constant of 0.1 ps. This temperature is the same as the one used for AMBER14sb's parametrization and testing, and it is in line with the temperature in the experiments of our study. Counter ions (Na<sup>+</sup> and Cl<sup>-</sup>) were added to obtain neutrality. The pressure is fixed at 1 atm using the Parrinello-Rahman barostat<sup>43</sup> with a coupling constant of 2 ps. We apply a cutoff of 1 nm for both the van der Waals and electrostatic interactions. Long-range electrostatic interactions are computed using Particle Mesh-Ewald<sup>44,45</sup>. Bond lengths are constrained using LINCS<sup>46</sup> and water geometry are constrained using SETTLE<sup>47</sup>.

The ACE2/RBD complex, ACE2 and RBD simulations are analyzed on the 250 to 500 ns interval (see next section). The ACE2/RBD complex simulation is used to quantify the contacts, H-bonds and salt-bridges between ACE2 and RBD in order to characterize the ability of the ligands to block those interactions. The ACE2 and RBD simulations are used to determine an ensemble of configurations representative of their flexibility in order to

take it into account while performing the docking of the ligands. Their main configurations are identified using Daura's clustering algorithm<sup>48</sup> with a cutoff of 0.15 nm on the backbone atoms. Clusters containing at least 5% of the total population, four clusters for the RBD and three clusters for ACE2, are considered for docking (more details in section 2.2).

We also perform 100-ns MD simulations for three RBD mutants (E484K, N501Y and E484K with N501Y) using the protocol described above. These simulations were started from the center of the biggest cluster of the RBD simulation. The three RBD mutants (E484K, N501Y and E484K with N501Y) are generated using PyMOL<sup>49</sup>.

### Identification of the interactions between ACE2 and RBD in the complex

The interactions between ACE2 and RBD in terms of contacts, H-bonds and salt-bridges are determined from a 500-ns MD simulation on the ACE2/RBD complex, starting from the experimental structure determined using X-rays crystallography (PDB:6M0J)<sup>6</sup> (SFig. S1A). The interface of the ACE2/RBD complex remains globally similar to the crystal structure, particularly on the 250-500 ns convergence interval, as shown by the backbone-RMSD of the interface ( $0.22 \pm 0.02$  nm) and the probability of the secondary structure motifs for both ACE2 and RBD (SFig. S1B-C). In terms of secondary structure, the propensity of  $\alpha$ -helices,  $\beta$ -sheets, turns and coils for ACE2 and RBD are essentially the same as in the crystal structure, except for a drop from 11% in the crystal to  $3 \pm 2\%$  in the RBD  $\alpha$ -helix propensity over the simulation.

During the simulation, most ACE2-RBD contacts are between the A1A2 segment (residues 19-83, helix-helix) or the HS segment (residues 322-362, helix-sheet-sheet) of ACE2 and the RBM segment (residues 438-506, mainly disordered with a small helix and small sheets) of RBD, as shown in SFig. S2A. Overall, most of these ACE2-RBD contacts are present in the crystal structure: experimental contacts are observed  $72 \pm 6\%$  of the time during the simulation, with this percentage significantly rising up to  $89 \pm 4\%$  when using a slightly less stringent distance threshold of 0.6 nm for the contact definition during the simulation (instead of 0.4 nm as in the experiment). More precisely for RBD, there are only four residues that interact with ACE2 in the crystal structure, but that interact with ACE2 less than 60% of the time during the simulation: Lys-417, Gly-446, Gly-447 and Glu-484. On the other hand, three more residues of RBD interact with ACE2 during the simulation: Phe-490, Pro-491 and Leu-492. In terms of H-bonds between ACE2 and RBD, all those observed in the crystal structure (Asp-30, Gln-42, Tyr-83 and Lys-353 on ACE2's side and Gly-446, Asn-487 and Gly-502 on RBD's side) are also present to varying degrees during the simulation (SFig. S2B). Moreover, other relevant H-bonds are observed during the simulation because it takes into account the flexibility of the complex coming from it being in a solvated environment at 300 K and 1 atm. In terms of salt-bridges, the crystal D30-K417 salt-bridge is the most populated in our simulation and two new salt-bridges (E223-K458 and E37-R403) are also observed (SFig. S2C).

A more detailed comparison of the H-bonds present in the crystal structure is also presented in Table S1. After addition

of the hydrogen atoms in the crystal structure and minimization, we found that five out of the thirteen H-bonds identified in Table 1 of *Lan et al.*<sup>6</sup> satisfy our distance and angle criteria for H-bonds identification: Lys-417(RBD)/Asp30(ACE2), Asn-487(RBD)/Gln-24(ACE2), Asn-487(RBD)/Tyr-83(ACE2), Tyr-489(RBD)/Tyr-83(ACE2) and Tyr-505(RBD)/Glu-37(ACE2). In our MD simulation, most of these H-bonds are unstable and only Lys-417(RBD)/Asp-30(ACE2) and TYR505(RBD)-GLU37(ACE2) are formed with over 25% occurrence rate. In addition to these H-bonds, the minimization of the crystal structure leads to the formation of 10 new H-bonds. In our MD simulation, only Tyr-449(RBD)/Asp-38(ACE2), Gln-493(RBD)/GLU-35(ACE2) and Gln-493(RBD)/Lys-31(ACE2) are stable with an occurrence rate of 40.64%, 36.95% and 34.82% respectively (Table S1).

### Assessing the structural flexibility of the starting ACE2 and RBD structures

The representative configurations of ACE2 in solution are identified from the 500-ns MD simulation on ACE2 alone, starting from its structure in the crystal complex (PDB:6M0J) (SFig. S3A). We establish that the simulation is converged after the first 250 ns by looking at the backbone RMSD and the secondary structure as a function of time (SFig. S3B-C). ACE2 keeps a structure similar to when it is in the complex as shown by the backbone RMSD on the whole ( $0.29 \pm 0.02$  nm) and on the segments A1A2 and HS at the interface with RBD ( $0.26 \pm 0.02$  nm). In terms of per residue secondary structure, slightly longer  $\alpha$ -helices are observed in the A1A2 segment during the simulation, while slightly longer  $\beta$ -sheets are observed in the HS segment compared to the crystal structure in complex form (SFig. S3D).

Similarly, the representative configurations of RBD in solution are identified from the 500-ns MD simulation on RBD alone, starting from its structure in the crystal complex (PDB:6M0J)<sup>6</sup> (SFig. S4A), with convergence also achieved after 250 ns (SFig. S4B-C). While RBD deviates more from the complex structure than ACE2, it stays relatively near from it as shown by the backbone RMSD:  $0.37 \pm 0.03$  nm on the whole RBD and  $0.40 \pm 0.05$  nm on the RBM segment at the interface with ACE2. In terms of per residue secondary structure, the helix and sheets fluctuate with other motifs, while some sheets are slightly longer during the simulation (SFig. S4D). In particular, residues 443-447, 474-489 and 500-505 of the RBM, which are interacting with ACE2 in the complex, show the highest degree of fluctuations.

## 2.2 Molecular docking

The docking of corilagin and TGG are performed using AutoDock VINA v1.1.2<sup>50</sup>. The protein flexibility is taken into account by performing the docking on the center of all clusters representing at least 5% of the sampled population. The ligand flexibility is considered by VINA's methodology. VINA's estimation of the "correctness" of the poses is done based on a simplified physics-based potential with empirically determined weights<sup>50</sup>. Every atomic pair is affected by a steric term, and, depending on the pair type, a hydrophobic term and a hydrogen bond term<sup>50</sup>. However, it is important to note that docking and scoring techniques are of

ten simplified for efficiency and the predicted binding affinities only weakly correlates with experimental predictions<sup>51</sup>. The region of interest during the docking involves residues found at the ACE2/RBD interface in the complex. On ACE2, the region of interest involves two long  $\alpha$ -helices between residues 19 and 83 (referred to as A1A2) and small  $\alpha$ -helix followed by a small  $\beta$ -sheet between residues 322 and 362 (referred to as HS). The docking is targeted on these regions using a box with  $x$ ,  $y$ ,  $z$  dimensions of 22.640, 52.072 and 14.633 Å respectively. On the RBD, the region of interest is composed of residues 438 to 506 (referred to as the RBM) and docking is focused on this region using a  $27.625 \times 43.358 \times 26.026$  Å box. Both molecular systems can be visualized in Figure S6. VINA's exhaustiveness parameter is set to 100. The conversion between PDB and PDBQT format is done using Open Babel v3.1.0<sup>52</sup>.

### 2.3 Protein-ligand simulations

We use the best prediction (highest binding affinity) generated by AutoDock VINA as the starting point for additional MD simulations for each combination of the five proteins – ACE2, RBD, RBD(E484K), RBD(N501Y) and RBD(E484K-N501Y) – and two ligands – corilagin and TGG. The same simulation protocol described in section 2.1 is used to launch a 100-ns simulation for each of the 10 systems.

The ligand parameters were determined using the generalized AMBER forcefield (GAFF)<sup>53</sup> with their partial charges determined using the RESP protocol<sup>54,55</sup> with ANTECHAMBER<sup>53,56</sup>. The electrostatic potential of each ligand has been computed using HF6-31G\*/HF6-31G\* with Gaussian16<sup>57</sup>. The initial conformations of corilagin and TGG used for those computations were taken from previously published work<sup>30</sup> where they were determined using PM3 semi-empirical Molecular Orbital Theory. The initial structures of the ligands are shown in Figure S5. All files are converted GROMACS compatible format with the help of ACPYPE<sup>58</sup>.

### 2.4 Analysis

The analysis of the MD simulations is done using a combination of GROMACS tools<sup>39</sup> and in-house scripts. Secondary structures (SS) are determined using DSSP<sup>59</sup>. Hydrogen bonds (H-bonds) are defined using a 0.35 nm donor-acceptor cutoff and a 30° hydrogen-donor-acceptor angle cutoff. Contacts are defined with a 0.40 nm cutoff, the same cutoff used in the analysis of the experimental structure<sup>6</sup>. Salt-bridges are defined using a 0.40 nm distance cutoff between the oppositely charged groups<sup>60</sup>. Molecular visualization is done using PyMOL<sup>49</sup> and ligand/protein interaction visualization using LigPlot+<sup>61,62</sup>. Daura's algorithm is used for clusterization<sup>48</sup>.

### 2.5 Binding free-energy

The MMPBSA method<sup>38</sup> is used to estimate the protein/ligand binding free-energy ( $\Delta G_{bind}$ ), defined by

$$\Delta G_{bind} = \langle G_{RL} - G_R - G_L \rangle_{RL}$$

where  $G_{RL}$ ,  $G_R$  and  $G_L$  are the free-energy of the receptor/ligand complex, of the receptor alone and of the ligand alone, respectively. We use a single trajectory MMPBSA computation: the conformation of the complex (RL), receptor (R) and ligand (L) are all taken from a unique MD trajectory. The bracket pair  $\langle \rangle$  represent an ensemble average over all receptor/ligand conformations. More specifically, the free-energy  $G$  is estimated according to

$$\Delta G = U + G_{solvation} - TS,$$

where  $U$  is the internal energy, computed using the AMBER14sb forcefield field,  $G_{solvation}$  is the solvation free-energy and is usually decomposed into a polar part, computed by solving the Poisson-Boltzmann equation and a non-polar part that depends on the solvent accessible surface area (SASA),  $T$  is the temperature and  $S$  is the entropy<sup>63</sup>. The MMPBSA method offers a relatively quick and easy way to estimate the binding free-energy. It does, however, makes a few crude approximation: the solvation is considered implicitly and thus possibly neglects crucial water molecules at the binding site. Moreover, the entropic part of the equation is often neglected (as in this study)<sup>63,64</sup>. In spite of these limitations, the MMPBSA method has proven to be useful for refining the results of docking predictions<sup>63</sup>.

MMPBSA computations are done with `g_mmpbsa` utility<sup>65</sup>, which uses APBS<sup>66</sup> for computing the polar part of the solvation free-energy. The dielectric constants of the solute and solvent are set to 2 and 80 respectively. The surface tension ( $\gamma$ ) is set to 0.0226778 kJ/(mol Å<sup>2</sup>) and the temperature at 300 K. The results are computed from the convergence interval of the ligand-protein MD simulations using 40 ps snapshots. A 500-steps bootstrap analysis is used to compute the average and standard deviation of the free energy.

### 2.6 Products

Corilagin ( $\beta$ -1-O-Galloyl-3,6-(R)-hexahydroxydiphenoyl-D-Glucose), with molecular formula  $C_{27}H_{22}O_{18}$  and molecular weight of 634.45 g mol<sup>-1</sup>, was obtained from Cayman Chemical (USA). The powder material, C.A.S. 23094-69-1, is natural in origin, with purity >98%. TGG (1,3,6-tri-O-galloyl- $\beta$ -D-glucose) with molecular formula  $C_{27}H_{24}O_{18}$  and molecular weight of 636.46 g mol<sup>-1</sup>, was obtained from MuseChem (USA). The powder material, C.A.S. 18483-17-5, is natural in origin, with purity 98.23%. Host Cell Receptor Binding Domain (RBD) (RayBiotech, cat number: 230-30162) was expressed at Arg319-Phe541 region in human embryonic kidney (HEK293) cells with a His-tag at C-terminal. The protein was supplied as a 0.2  $\mu$ m filtered solution in PBS (pH 7.4) with purity > 95%. Recombinant Human ACE2 Protein was purchased from Bioss Inc. (Cat number: BS-46110P). Recombinant Human ACE2 Biotinylated Protein was purchased from (R&D Systems).

### 2.7 Surface Plasmon Resonance (SPR)

SPR analyses are performed using a Biacore T200 instrument (GE Healthcare). S1-RBD and ACE2 recombinant proteins are respectively immobilized on a carboxymethylated dextran CM5 sensor chip (GE Healthcare) using an amine-coupling strategy. Briefly,

the sensor chip surface is activated with a 1:1 mixture of N-hydroxysuccinimide (NHS) and 3-(N,N-dimethylamino)-propyl-N-ethylcarbodiimide (EDC). Recombinant protein solutions (20  $\mu\text{g/ml}$ ) are injected at a flow rate of 10  $\mu\text{l/min}$  using HBS-N running buffer (10 mM HEPES, 150 mM NaCl, pH 7.4) to reach a level of immobilization of 200 RU. Surfaces (protein and reference) are blocked by the injection of an ethanolamine-HCl solution. Binding kinetics of TGG and corilagin over the immobilized recombinant proteins sensor chip are evaluated in HBS-N buffer with increasing polyphenol concentrations (1 to 100 nM) at a flow rate of 20  $\mu\text{l/min}$ . Association time is set at 180 sec and dissociation time is extended up to 1,200 seconds. The sensor chip surface is regenerated by injecting 15  $\mu\text{l}$  of a 10 mM glycine solution, pH 3. For ACE2/RBD interactions, the binding partner is injected over the counterpart-functionalized surface with concentrations from 1 to 100 nM and surface is regenerated with 15  $\mu\text{l}$  of a 50 mM NaOH solution. For inhibition assay, 50 nM RBD recombinant protein is pre-incubated for 30 min at room temperature (RT) with increasing polyphenol concentrations and the mixtures are subsequently injected over an ACE2 functionalized CM5 surface. Binding sensograms are obtained by subtracting the reference flow cell (without protein). Experiments are performed at least in duplicate and data analysis is performed using the BIA evaluation software package (GE Healthcare) and fit to a one-site (1:1 molecular ratio) Langmuir binding model.

## 2.8 SARS CoV-2 RBD Spike Protein and Human ACE2 Binding Inhibitor Assay

The capacity of TGG and corilagin to inhibit the binding between the RBD Spike protein and the human ACE2 was assessed at different concentrations from 0.1 to 10  $\mu\text{M}$ , by ELISA. For this, ELISA plates were coated with 0.5  $\mu\text{g/ml}$  of RBD Spike protein and kept overnight at 4°C. Plates were then rinsed three times with the washing buffer (0.05% Tween 20 in phosphate buffered saline (PBS)) and then blocked with the blocking buffer (1% bovine serum albumin (BSA) in PBS) by incubating for 1 hour at 37°C. After three washing, one hundred microliters of biotinylated human ACE2 protein, diluted at 0.5  $\mu\text{g/ml}$  in the blocking buffer, were added to each well and incubated at 37°C for 1 h. After washing with the same washing buffer, diluted peroxidase-conjugated streptavidin was added to each well and incubated at 37°C for 30 min. Following three washes, chromogenic substrate solution was added to each well and incubated at 37°C for 30 min followed by 50  $\mu\text{L}$  of the stop solution (2N  $\text{H}_2\text{SO}_4$ ). The absorbance was then read at 450 nm. To note, a concentration response curve for the human ACE2 protein (0.015 to 2  $\mu\text{g/ml}$ ) was established to confirm a concentration-dependent increase of the absorbance at 450 nm (Figure S7A). For the competition assay, different concentrations of TGG and corilagin were incubated with immobilized RBD Spike protein for 1 hour at 37°C before the addition of the human ACE2 protein.

## 2.9 TGG, corilagin and human ACE2 binding assay

To study the possible binding of TGG or corilagin and their mixture to the human ACE2, the ELISA ACE2 detection kit (R&D Sys-

tems) was used with some modifications. Plates were coated with 0.5  $\mu\text{g/ml}$  of human ACE2 antibody which can bind to the extracellular region of the ACE2 protein (AA 18-740) at room temperature during overnight. After washing, wells were blocked with the blocking buffer. For the competition assay, biotinylated human ACE2 was mixed with various amounts of TGG or corilagin or their mixture for 1 hour at 37°C. After incubation, the mixture of human ACE2 and polyphenols was added to the coated wells and incubated for 1 hour at 37°C. After washing, diluted peroxidase-conjugated streptavidin was added to each well and incubated at 37°C for 30 min. Chromogenic substrate was added to each well after washing and incubated at 37°C for 30 min. The absorbance was then read at 450 nm in a fluorescent microplate reader. To note, a concentration response curve for the human ACE2 protein (0.015 to 2  $\mu\text{g/ml}$ ) was established to confirm a concentration-dependent increase in absorbance at 450 nm (Figure S7B).

## 2.10 Statistical analysis for binding assays

Data were analyzed using the GraphPad Prism program. For the inhibitory effects of TGG, corilagin and their mixture on the SARS CoV-2 Spike protein RBD and human ACE2 interactions, statistical analyses were performed using One-way ANOVA analysis followed by the Dunnett's t-test. A p value less than 0.05 was considered statistically significant.

# 3 Results

## 3.1 The impact of Corilagin and TGG on the ACE/RBD Wild-Type interactions

### 3.1.1 Molecular Docking

In order to probe the possible interaction sites between the ligands (corilagin/TGG) and ACE2/wild-type RBD (WT-RBD), we first carry out molecular docking using AutoDock VINA<sup>50,61</sup>. To consider the protein flexibility, we use representative structures extracted from the MD simulations run independently on ACE2 and RBD as described in Sect. 2.1. The docking simulations result in a wide variety of predicted conformations characterized by varying binding affinities and positions at the interface. Conformations sampled on RBD (WT and mutants) are shown as two-dimensional occurrence map as a function of VINA's binding energy and fraction of contacts with interface residues on Figure S8 for corilagin and Figure S9 for TGG.

Docking on the RBD-WT for corilagin leads to docked positions with VINA docking energy ranging from  $-5.8$  to  $-8.1$  kcal/mol and a fraction of contacts with interface residues going from 0.20 to 0.45. For the RBD-WT and TGG, the VINA docking energy spectrum is narrow and lower, from  $-7.0$  to  $-8.8$  kcal/mol, with a broader fraction of contacts from 0.20 to 0.55 among the identified docked conformations. Interestingly, many docked TGG conformations are characterized by a low binding energy and a high fraction of contacts with interface residues.

### 3.1.2 Molecular Dynamics

Molecular docking by itself takes into account only limited protein and ligand conformational dynamics. Yet, molecular flexibil-

ity is critical for a reliable and predictable characterization<sup>36,37</sup>. Thus, we perform a 100-ns MD simulation on the best predicted ACE2/RBD-WT with corilagin/TGG complexes given by Autodock VINA in order to allow for local rearrangements both on the protein and ligand sides.

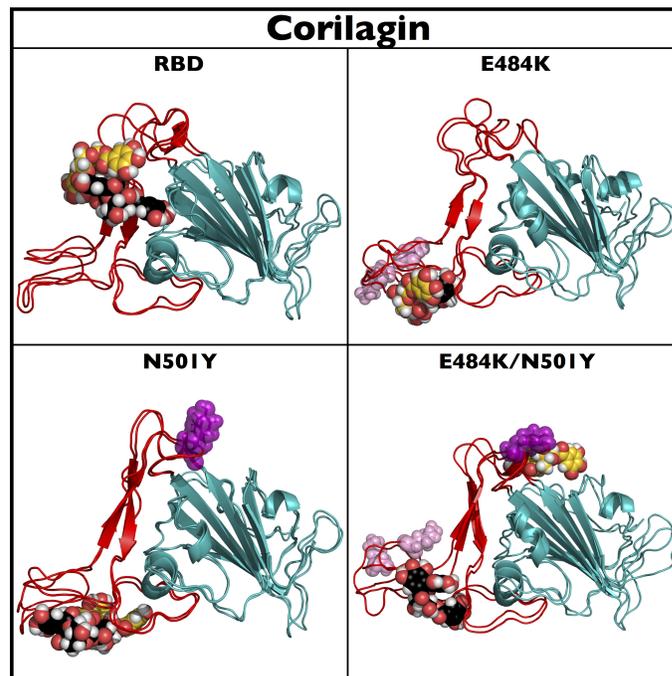
**ACE2.** ACE2's structure at the interface remains very stable when in contact with either corilagin or TGG (Figure S10). After docking, the A1A2 and HS backbone RMSD against the experimental structure is  $0.25 \pm 0.01$  and  $0.25 \pm 0.01$  nm for corilagin and TGG respectively as compared to  $0.26 \pm 0.02$  nm without ligand. Moreover, both the A1A2 and HS secondary structures remain unaffected by the presence of ligand: the  $\alpha$ -helix content is  $53 \pm 3\%$  (corilagin) and  $53 \pm 1\%$  (TGG) in the presence of the ligands as compared to  $55 \pm 1\%$  without the ligands, and the  $\beta$ -sheet content remains at  $15 \pm 2\%$  (corilagin) and  $15 \pm 2\%$  (TGG), while it is  $16 \pm 1\%$  without ligand. The initial docked positions of the corilagin, located in between the HS segment and the middle of the A1 helix, and TGG, located on the flexible loops of the HS segments, are also very stable and show little deformation during the MD.

**RBD.** The structure of the WT RBM segment is also only weakly affected by the ligands (Figures 1 and 2). Indeed, the average backbone RMSD measured with respect to the experimental structure computed is  $0.37 \pm 0.02$  nm,  $0.44 \pm 0.05$  nm for corilagin and TGG respectively as compared to  $0.40 \pm 0.05$  nm for the system without the ligands. The secondary structure is largely unaffected by the presence of the ligand. With a bound ligand, its  $\alpha$ -helix content is  $5 \pm 2\%$  (corilagin) and  $4 \pm 3\%$  (TGG), compared to  $6 \pm 3\%$  without it; the  $\beta$ -sheet content is  $22 \pm 6\%$  with corilagin, and  $16 \pm 4\%$  with TGG, as compared to  $16 \pm 4\%$  without any ligand. The solvent accessible surface area of the WT RBM is of  $49 \pm 2$  nm<sup>2</sup>. This relative stability on the protein side is reflected on the binding conformations of corilagin and TGG found with VINA, that both remain very stable over the MD simulation (Figures 1 and 2).

### 3.1.3 Interactions and Binding Energies

Molecular docking predictions use a simplified binding affinity score. In order to refine the estimation of the binding energies, we turn to the MMPBSA technique. We also characterize the interactions network between the proteins (ACE2/RBD-WT) and the ligands (Corilagin/TGG) and their evolution over the MD simulations.

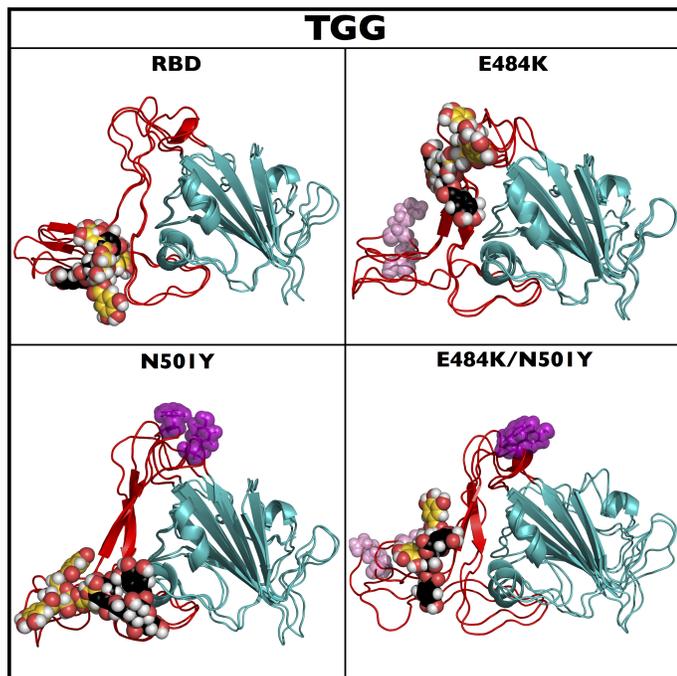
**ACE2.** The binding affinities of both ligands with ACE2 during the MD simulations are compared using the VINA score as well as the MMPBSA free energy as explained in Section 2 (Table 2). The average binding affinity of corilagin/ACE2 is  $-6.1 \pm 0.5$  kcal/mol (VINA score) and  $-0.1 \pm 0.2$  kcal/mol (MMPBSA). In spite of this negligible binding affinity, corilagin remains associated with ACE2 during the entire MD simulation, demonstrating at least the presence of a metastable state, once binding occurs. The binding affinity of TGG/ACE2 is more favorable with  $-6.0 \pm 0.4$  kcal/mol (VINA) and  $-14.4 \pm 0.2$  kcal/mol (MMPBSA). The LigPlot interaction maps between corilagin/TGG and ACE2 for the center of the biggest cluster (total population of 84% for corilagin and 98% for TGG) are shown on Figure S11. Corilagin is stabilized by



**Fig. 1** The docked position of corilagin on RBD and the three mutants. The RBM segment and the rest of the RBD are shown respectively in red and teal. Residues 484 and 501, both the location of tested mutation, are shown in pink and purple respectively. The ligand in black and gold is respectively the conformation after docking and the center of the biggest cluster sampled during the converged part of the MD simulation respectively.

nine H-bonds with multiple residues of the A1A2 segment (Asp30 twice, His34 and Glu37 twice) as well as with residues Arg393 (twice), Gln388 and Phe390. Four nonpolar contacts are formed with residues of the A1A2 segment (Asn33), the HS segment (Lys353 and Gly354) and the rest of the ACE2 (Pro389). For its part, TGG is forming four H-bonds with residues of the HS segment (Met323, Gln325 and Asp350 twice) and it is stabilized by a large number of nonpolar contacts, mainly with the HS segment (Asn322, Thr324, Gly326, Gly352, Gly354, Asp355 and Phe356) and the rest of ACE2 (Pro321, Met383, Ala386 and Arg393).

**RBD.** The binding affinities of both ligands with RBD during the MD simulations are compared using the VINA score as well as the MMPBSA free energy as explained in Section 2 and shown in Table 2. The corilagin/RBD binding affinity is  $-5.0 \pm 0.5$  kcal/mol (VINA) and  $-7.2 \pm 0.1$  kcal/mol (MMPBSA) and that for TGG/RBD is similar in terms of VINA ( $-6.0 \pm 0.4$  kcal/mol), but more favorable in terms of MMPBSA ( $-12.8 \pm 0.4$  kcal/mol). The LigPlot interaction maps between the corilagin/TGG and RBD for the center of the biggest cluster (total population of 96% for corilagin and 52% for TGG) are shown in Figures 3 and 4, respectively. Corilagin is stabilized by the formation of four H-bonds with Tyr449, Gln493, Ser494 and Gly496 as well as four nonpolar contacts with Tyr495, Gln498, Gly504 and Tyr505. On the other hand, TGG is stabilized by the formation of three H-bonds with Arg454, Glu471 and Pro491 and six nonpolar contacts with Leu455, Phe456, Arg457, Lys458, Thr470 and Leu492.



**Fig. 2** The docked position of TGG on RBD and the three mutants. The RBM segment and the rest of the RBD are shown respectively in red and teal. Residues 484 and 501, both the location of tested mutation, are shown in pink and purple respectively. The ligand in black and gold is respectively the conformation after docking and the center of the biggest cluster sampled during the converged part of the MD simulation respectively.

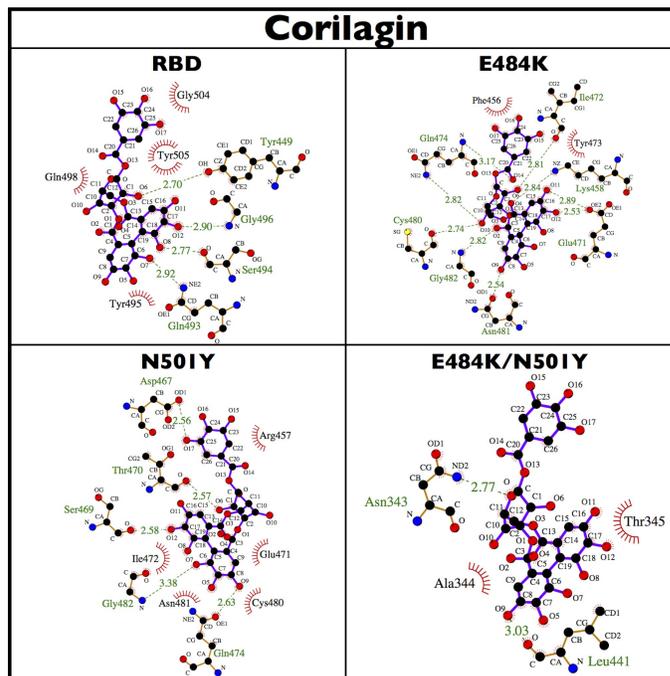
### 3.1.4 Corilagin/TGG ability to disrupt the ACE2/wild-type RBD interactions

Using a combination of molecular docking and molecular dynamics simulations, we find that both ligands, corilagin and TGG, are able to interact with ACE2 and the RBD-WT. We analyze below whether those interactions are compatible with a disruption of the ACE2/RBD-WT association.

**ACE2.** For ACE2/corilagin, although the binding energy is low ( $-0.1 \pm 0.2$  kcal/mol according to MMPBSA), the corilagin's localization on ACE2 is compatible with the disruption of many interface residues of the A1A2 segment (Asp30, His34, Glu37) and of the HS segment (Lys353, Gly354, Arg393).

For ACE2/TGG, TGG interacts only with the HS segment (Lys353, Gly354, Asp355, Arg393) and not with the A1A2 segment (SFig. S10) suggesting that it is predominantly interacting with a small portion of the ACE2's residues involved at the interface with RBD-WT. In this position, TGG might not be able to disrupt significantly the association of ACE2 with the RBD-WT.

**RBD-WT.** Our MMPBSA calculation shows that corilagin binds much more strongly to WT RBD than to ACE2. In its preferred binding site, corilagin interacts with five residues at the RBD/ACE2 interface – Tyr449, Gln493, Gly496, Gln498 and Tyr505 – and forms a H-bond with two of these – Tyr449 and Gly496 (Figure 1). Since these five residues form H-bonds with ACE2 in the complex, the presence of corilagin could interfere with the formation of these H-bonds and impair the complexa-



**Fig. 3** Corilagin interaction maps. The interaction maps of corilagin with wildtype RBD (top left), RBD/E484K (top right), RBD/N501Y (bottom left) and RBD/E484K-N501Y (bottom right) are shown for the center of the biggest cluster computed on the convergence interval using the protein backbone atoms and ligand non-hydrogen atoms. The nonpolar contacts, defined by a distance smaller than 0.40 nm, between the ligand and the protein are shown as red arcs. H-bonds and their donor/acceptor distance are shown in green. All figures were generated using LigPlot<sup>61,62</sup>.

tion.

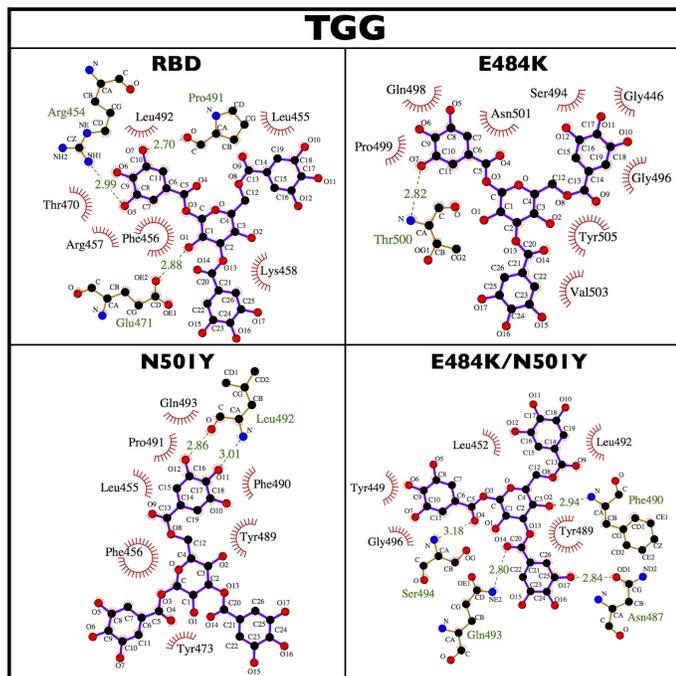
Our MMPBSA calculation also shows that the association between WT RBD and TGG is strong. TGG interacts with many residues involved at the ACE2-RBD interface in the complex such as nonpolar residues Leu455, Phe456, Phe490, Pro491 (with which it forms a H-bond) and Leu492 as well as polar residues: Lys458 and Gln493 (Table 1). Notably, these two polar residues are involved in a salt-bridge (Lys458) and a H-bond (Gln493) with ACE2.

In summary, both corilagin and TGG could have the potential to disrupt the interaction between ACE2 and the WT RBD. However, our numerical results shows that the disruption would be more important (better binding energy and better network of interactions) on the side of the WT RBD than on the side of ACE2, leaving the crucial physiological functions of ACE2 untouched.

### 3.1.5 Surface Plasmon Resonance

In order to validate our numerical observations, we turn to experiments for confirmation. First, we perform SPR measurements to determine the binding kinetics of TGG and corilagin to ACE2 and RBD.

The recombinant proteins ACE2 and RBD are respectively immobilized on carboxymethylated dextran sensor chips. The results (sensograms) show that both polyphenols bind avidly to the immobilized RBD (Figure 5A). Fitting the sensograms to a



**Fig. 4 TGG interaction maps.** The interaction maps of TGG with wildtype RBD (top left), RBD/E484K (top right), RBD/N501Y (bottom left) and RBD/E484K-N501Y (bottom right) for the center of the biggest cluster computed on the convergence interval using the protein backbone atoms and ligand non-hydrogen atoms. The nonpolar contacts, defined by a distance smaller than 0.40 nm, between the ligand and the protein are shown as red arcs. H-bonds and their donor/acceptor distance are shown in green. All figures were generated using LigPlot<sup>61,62</sup>.

one-site (1:1 molecular ratio) binding model leads to dissociation constant ( $K_D$ ) in the low nanomolar range, i.e., 1.8 nM for corilagin/RBD and 1.3 nM for TGG/RBD. In sharp contrast, no significant binding of neither corilagin, nor TGG, to ACE2 is observed by SPR over the range of 1 to 80 nM concentrations (Figure 5B). This observation indicates that the binding of polyphenols to RBD has clear specificity.

Next, we validate the interactions between ACE2 and RBD by means of two different experimental configurations; (i) binding of ACE2 to immobilized RBD and (ii) binding of RBD to immobilized ACE2. As expected, we observe strong interactions between these two proteins (Figure 5C) with  $K_D$  of 41 nM (ACE2 to immobilized RBD) and 63 nM (RBD to immobilized ACE2), in agreement with recent studies<sup>6,67,68</sup>.

Finally, the capacity of TGG and corilagin to inhibit the RBD/ACE2 interaction is evaluated by pre-incubating 50 nM RBD for 30 mins in presence, or absence, of increasing concentrations of polyphenols before injecting the mixtures onto an ACE2-functionalized sensor chip. Strikingly, 12.5 to 50 nM of TGG, or corilagin, fully inhibit the binding of RBD to surface-immobilized ACE2 (Figure 5D).

### 3.1.6 Binding assays

To complement the results obtained by MD and SPR, we also investigate the ability of TGG and corilagin to inhibit the interaction between the SARS-CoV-2 RBD protein and the human ACE2 us-

**Table 1 ACE2-RBD contacts blocked by the ligands.** RBD(RBM) residues blocked by corilagin (Cor.) and TGG. The RBM residues showed are those that are specifically involved in a contact pair with ACE2 that is formed with a probability of at least a 60% during the ACE2-RBD complex MD simulation. Nonpolar, polar, positively charged and negatively charged residues are shown respectively in gray, green, red and blue. The formation of a contact with the ligand is shown in gray. The presence of a H-bond or a salt-bridge is indicated by *HB* and *SB*, respectively. The star (\*) indicates that a H-bond is present in the experimental structure. The dagger (†) beside *SB* for Lys458 indicates that this residue was added to the table even if its contact probability with ACE2 is less than 60% (45%) because it forms a salt-bridge with E23 of ACE2.

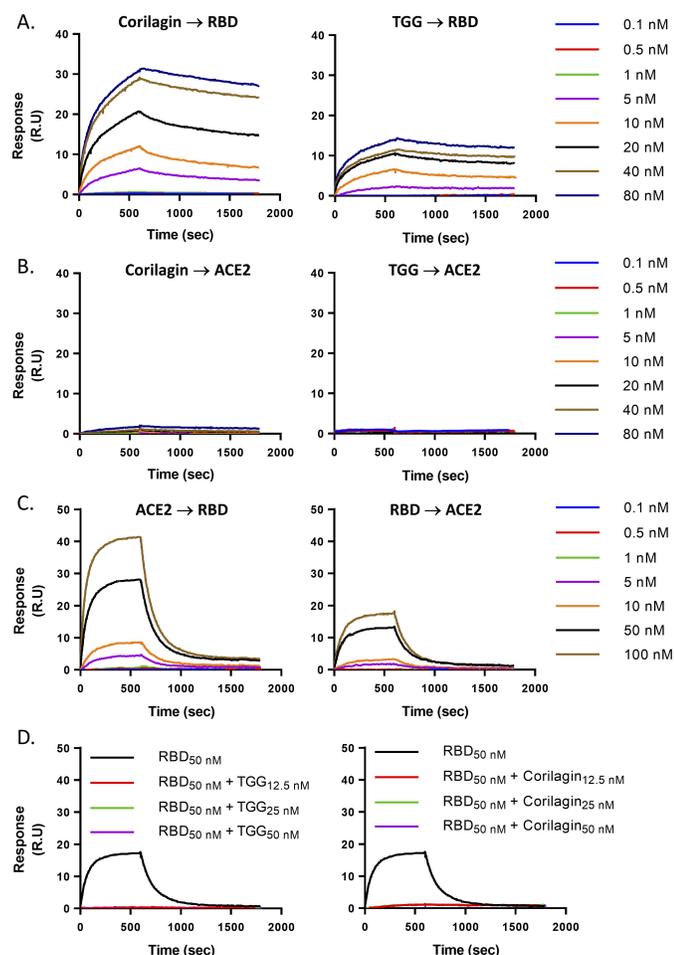
		RBM contacts blocked by the ligands							
		WT		E484K		N501Y		E484-N501Y	
In contact with A1A2/HS in the complex		Cor.	TGG	Cor.	TGG	Cor.	TGG	Cor.	TGG
Tyr-449	HB	x HB							x
Tyr-453	HB								
Leu-455			x				x		
Phe-456			x	x			x		
Lys-458	SB†/HB		x	x			x		
Tyr-473	HB			x			x		
Ala-475									
Gly-476	HB								
Phe-486									
Asn-487	HB*								
Tyr-489	HB						x		x
Phe-490	HB		x				x		x HB
Pro-491			x HB				x		
Leu-492			x				x HB		x
Gln-493	HB	x	x				x		x HB
Gly-496	HB	x HB			x				x
Gln-498	HB	x			x				
Pro-499					x				
Thr-500	HB				x HB				
Asn-501	HB				x HB				
Gly-502	HB*								
Tyr-505	HB	x			x				

ing binding assays. We find that the incubation of SARS-CoV-2 RBD Spike protein with TGG or corilagin results in a significant reduction of the interaction between RBD and ACE2, e.g., from 45% up to 75% for concentrations from 0.1 to 10  $\mu$ M (Figure 6A-B). Moreover, the mixture of TGG with corilagin, from 0.1 to 5  $\mu$ M, inhibits 50% of the interaction and does not potentiate the inhibitory effect of each compound (Figure 6C).

Next, we evaluate whether the inhibition of the RBD-ACE2 interaction by TGG and corilagin is associated with a preferential binding of these polyphenols to RBD in comparison to ACE2. For this, the ACE2 antibody is immobilized and the binding of ACE2 protein in absence or in presence of polyphenols is evaluated by ELISA. Strikingly, TGG and corilagin, used alone or in combination, do not reduce avidly the recognition of the ACE2 protein by the anti-ACE2 (Figure 7A-C). These results suggest that the inhibition of the binding of the WT RBD to ACE2 is mainly mediated by the binding of these polyphenols to the RBD protein, and less to the ACE2 protein.

### 3.2 The impact of Corilagin and TGG on mutant RBDS

Our results from both numerical techniques (molecular docking and molecular dynamics) and experimental techniques (SPR and ELISA assay), show that both corilagin and TGG could interfere

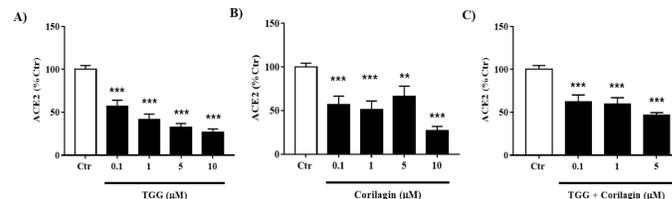


**Fig. 5** Characterization of molecular interactions by surface plasmon resonance. **A, B**) Binding kinetics of corilagin and TGG on immobilized **(A)** RBD and **(B)** ACE2. The recombinant proteins RBD and ACE2 are respectively immobilized on a CM5 sensor chip and increasing concentrations of polyphenols are injected to evaluate binding kinetics. **C**) Kinetics of ACE2 binding to immobilized RBD (left panel) and kinetics of RBD binding to immobilized ACE2 (right panel). **D**) Pre-incubation of RBD (50 nM) for 30 minutes with increasing concentrations of corilagin or TGG inhibit the binding of RBD to immobilized ACE2.

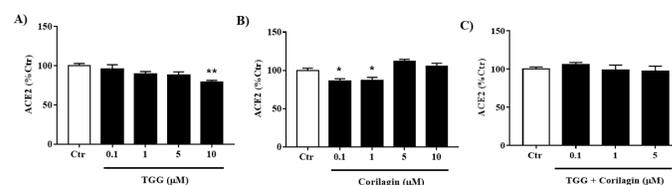
with the binding of the WT RBD to ACE2, by primarily interacting on the side of the RBD. However, due to evolutionary pressure, the SARS-CoV-2 virus is in constant mutations and led to the rise of three main variants of concerns with many critical mutations on the RBD; E484K (B.1.351, B.1.1.28) and N501Y (B.1.1.7, B.1.351, B.1.1.28). In order to probe the therapeutic potential of our two molecules, we also test their binding with three mutant RBDs, with mutation E484K, N501Y and E484K/N501Y, using the described numerical protocol we used on the WT RBD and that was validated using SPR and ELISA assay.

#### The impact of the mutations on RBD's stability

First, the impact of the mutants E484K, N501Y and E484K/N501Y on RBD are analyzed using 100-ns MD simulations starting from the center of the biggest cluster sampled during the MD on the WT. These systems converge quickly after about 40 ns as shown by the backbone RMSD and the secondary structure as a func-



**Fig. 6** Inhibitory effects of TGG, corilagin and their mixture on the interaction between SARS CoV-2 Spike protein and human ACE2. TGG **(A)** and corilagin **(B)** are tested at different concentrations (0.1, 1, 5 and 10  $\mu\text{M}$ ) and their mixture **(C)** (0.1, 1, 5  $\mu\text{M}$ ) to evaluate their ability to inhibit the binding of immobilized Spike protein (0.5  $\mu\text{g}/\text{ml}$ ) to human biotin labeled ACE2 (0.5  $\mu\text{g}/\text{ml}$ ), by using the ELISA assay. The absorbance values at 450 nm of human ACE2 (0.5  $\mu\text{g}/\text{ml}$ ) are set to 100%. Results are expressed as mean  $\pm$  standard error of the mean (SEM) of two (combined effect) or three independent assays. Statistical analysis is performed using the One-way ANOVA followed by the Dunnett's post hoc test with \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  compared to human ACE2 (0.5  $\mu\text{g}/\text{ml}$ ).



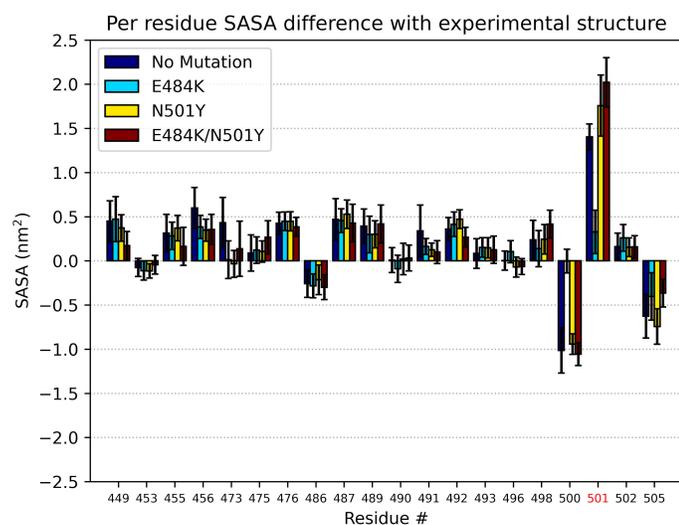
**Fig. 7** Inhibitory effects of TGG, corilagin and their mixture on the interaction between human ACE2 and ACE2 antibody (18-740 AA). TGG **(A)** and corilagin **(B)** are tested at different concentrations (0.1, 1, 5 and 10  $\mu\text{M}$ ) and their mixture **(C)** (0.1, 1, 5  $\mu\text{M}$ ) to study their ability to inhibit the binding of immobilized ACE2 antibody (0.5  $\mu\text{g}/\text{ml}$ ) to human biotin labeled ACE2 (0.5  $\mu\text{g}/\text{ml}$ ), by using the ELISA assay. The absorbance values at 450 nm of human ACE2 (0.5  $\mu\text{g}/\text{ml}$ ) were set to 100%. Results are expressed as mean  $\pm$  standard error of the mean (SEM) of two (combined effect) or three independent assays. Statistical analysis was performed using the One-way ANOVA followed by the Dunnett's post hoc test with \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  compared to human ACE2 (0.5  $\mu\text{g}/\text{ml}$ ).

tion of time (SFig. S12). The similarity of the RBM segment between the three mutants RBDs and the WT is characterized using three parameters: the backbone-RMSD compared to experimental structure, the secondary structure (SS) and the solvent accessible surface area (SASA).

In terms of RMSD, the RBM segment of the three mutant sequences closely resembles the WT experimental structure with  $0.34 \pm 0.03$  nm (E484K),  $0.41 \pm 0.04$  nm (N501Y) and  $0.32 \pm 0.05$  nm (E484K/N501Y). The secondary structure is also similar in terms of the  $\alpha$ -helical content with  $3 \pm 3\%$  (E484K),  $3 \pm 3\%$  (N501Y) and  $4 \pm 3\%$  (E484K/N501Y) as well as in terms of the  $\beta$ -sheet content with  $20 \pm 3\%$  (E484K),  $24 \pm 6\%$  (N501Y) and  $24 \pm 6\%$  (E484K/N501Y). In terms of SASA, the RBM of the three mutant systems is equally exposed to the solvent:  $47 \pm 1$  nm<sup>2</sup> (E484K),  $47 \pm 1$  nm<sup>2</sup> (N501Y) and  $49 \pm 1$  nm<sup>2</sup> (E484K/N501Y). Overall, these results indicate that the structure of the RBM segment of RBD is not much affected by the mutations and closely resembles our results on the WT RBD (presented in section 3.1.2) as well as with the crystal structure of RBD in complex with ACE2.

Beyond the overall structural stability, we also assess the poten-

tial impact of the mutations on RBD interactions with ACE2, since this association is most critical for SARS-CoV-2's cell recognition. To do so, we focus our attention on the residues of RBD that are in contact with ACE2 when they form a complex. We compute the difference between the average SASA of these residues for all simulations (WT, E484K, N501Y and E484K/N501Y) and the experimental structure of RBD alone (Figure 8). In agreement with what we observe in terms of RMSD and SS, these mutations have little impact on region at the interface, as most residues have similar accessibility. The only notable difference is observed for E484K for which the accessibility of Thr500 increases and the accessibility of Asn501 decreases significantly. For the two RBDs with the N501Y mutations, we do observe a slight increase in the accessibility for residue 501 compared to the WT; this is likely due to the size increase between TYR and ASN as we do not observe any significant structural change. Overall, the RBD residues involved at the interface with ACE2 stay accessible even in the presence of these mutations, hence, could still potentially interact with ACE2.



**Fig. 8 Mutations effect on the solvent accessibility of RBD alone.** RBD's per residue solvent accessible surface area (SASA) difference between the MD and the experimental structure. Only the residues of RBD interacting with ACE2 (contact probability greater than 60% during the MD simulation) in the complex structure are shown. The red residue number indicates the position of a mutation. The SASA of wildtype (blue), E484K (teal), N501Y (yellow) and E484K/N501Y (red) are compared. The error bars correspond to the standard deviation over the 250-500 ns interval.

### 3.3 Corilagin and TGG interactions with ACE2 and RBD

#### 3.3.1 Molecular Docking

Using the same molecular docking protocol used for the WT, the predicted interaction sites for corilagin and TGG with the three mutants were generated. Although, there are very few structural differences between the WT and the mutants, as shown in section 3.2, we clearly see changes in the generated ensembles of docked conformations.

For the RBD/E484K and RBD/E484K-N501Y mutant, the two-dimensional occurrence map is very similar to the one of the

WT for corilagin (Figure S8) and TGG (Figure S9; a very similar range of binding energies and fraction of contacts with interface residues are sampled, although, for both mutation, the distribution is slightly shifted toward higher binding energies. In contrast, with the N501Y mutation alone, the generated corilagin conformations have similar binding energy (from  $-5.4$  to  $-6.5$  kcal/mol), but are characterized by a wide variety of interface contact fraction (from 0.05 to 0.60). The same is observed for the docking results obtained using TGG (Figure S9).

#### 3.3.2 Molecular Dynamics

As with the WT, we follow the molecular docking with 100 ns of molecular dynamics simulation on the best predicted RBD/ligand complex for the three mutant RBD and the two ligands. Convergence of all systems is reached at 75 ns (except for RBD/E484K-N501Y with corilagin, for which, the convergence is reached at 90 ns) as monitored by the RMSD on the proteins' backbone atoms and the ligands' heavy atoms (SFig. S13).

**RBD/E484K.** The RMSD with respect to the experimental structure computed on the backbone atoms remains small:  $0.33 \pm 0.02$  nm and  $0.31 \pm 0.07$  nm for corilagin and TGG, respectively, compared with  $0.34 \pm 0.03$  nm without the ligands. The secondary structure remains also stable in terms of  $\alpha$ -helix content ( $1 \pm 2\%$ ,  $0 \pm 0\%$  and  $3 \pm 3\%$  for corilagin, TGG and without the ligands, respectively) and  $\beta$ -sheet content ( $19 \pm 2\%$ ,  $17 \pm 4$  and  $20 \pm 3\%$ , respectively).

On the ligand side, the position of corilagin remains very stable, while that of TGG moves slightly away from the  $\beta$ -sheet of RBM in order to interact with the flexible loop between GLN498 and GLY504 (Figures 1 and 2).

**RBD/N501Y.** The backbone RMSD measured against the experimental structure also remains small with  $0.36 \pm 0.02$  nm and  $0.39 \pm 0.03$  nm for corilagin and TGG, respectively, compared with  $0.41 \pm 0.04$  nm in the absence of a ligand. The secondary structure also stays very stable upon ligand addition: the  $\alpha$ -helix content is  $4 \pm 2\%$ ,  $0 \pm 0\%$  and  $3 \pm 3\%$  and the  $\beta$ -sheet content being  $25 \pm 5\%$ ,  $23 \pm 4\%$  and  $24 \pm 6\%$  for respectively corilagin, TGG and without the ligands.

Corilagin stays near its initial docking position during the MD, while a small movement is observed for TGG (Figures 1 and 2). From its initial position, between the loop associated with residues 405-424 and the 455-461 loop, TGG moves slightly to interact with the 455-458 loop and the 489-493 loop. The latter portion of the loop contains residues that interact directly with ACE2 in the complex.

**RBD/E484K-N501Y.** The backbone RMSD measured against the experimental structure remains also small for this system:  $0.32 \pm 0.05$  nm,  $0.33 \pm 0.05$  nm and  $0.32 \pm 0.05$  nm for corilagin, TGG and without the ligand, respectively. The secondary structure stays similar to WT: the  $\alpha$ -helix content is  $4 \pm 3\%$ ,  $5 \pm 2\%$  and  $4 \pm 3\%$  and the  $\beta$ -sheet content is  $24 \pm 6\%$ ,  $23 \pm 5\%$  and  $24 \pm 6\%$  for corilagin, TGG and without the ligands, respectively.

After around 80 ns of MD, starting from the VINA docked position, corilagin disassociates from the RBD and reassociates with it at around 90 ns, but at a completely new localization, in the region between the RBM and the rest of the RBD, near the mutation

N501Y (Figure 1). In this new position, the corilagin doesn't interact with RBM's residues involved in the association with ACE2. On the other hand, TGG is really stable at its initial docked position and stays near it for the entire MD calculation (Figure 2).

### 3.3.3 Interactions and Binding Energies

As for the WT RBD, the structure of the mutated RBD – E484K, N501Y and E484K/N501Y – are not affected significantly by corilagin or TGG over the 100-ns MD simulations and the VINA docking positions are generally fairly stable. Yet, small local rearrangements, which are made possible by the dynamical trajectories, contribute to a better exploration of their interactions, modifying the binding affinity and the ligand/RBD contact network as described in the following paragraphs.

**RBD/E484K.** The ligands' binding affinities with RBD/E484K during the MD simulations starting from an initial VINA docking are shown in Table 2. The computed corilagin/RBD binding affinity is  $-6.3 \pm 0.4$  kcal/mol (VINA) and  $-7.0 \pm 0.2$  kcal/mol (MMPBSA), while it is weaker for TGG/RBD when measured with VINA ( $-4.5 \pm 0.4$  kcal/mol), but stronger with MMPBSA ( $-10.9 \pm 0.4$  kcal/mol). The LigPlot interaction maps between corilagin/TGG and RBD/E484K for the center of the biggest cluster (total population of 92% for corilagin and 34% for TGG) are shown in Figures 3 and 4, respectively. Corilagin is stabilized by nine H-bonds involving seven RBM's residues (Lys458, Glu471 twice, Ile472, Gln474 twice, Cys480, Asn481, Gly482) and only two nonpolar contacts (Phe456 and Tyr473). For its part, TGG is only making one H-bond with Thr500, but has eight nonpolar contacts with Gly446, Ser494, Gly496, Gln498, Pro499, Asn501, Val503 and Tyr505.

**RBD/N501Y.** The binding affinities with RBD/N501Y are shown in Table 2. For corilagin/RBD, it is found to be  $-6.4 \pm 0.5$  kcal/mol with VINA and  $-11.3 \pm 0.1$  kcal/mol with MMPBSA, while the binding affinity for TGG/RBD is weaker with VINA ( $-3.8 \pm 0.5$  kcal/mol), but stronger with MMPBSA ( $-12.1 \pm 0.3$  kcal/mol). The LigPlot interaction maps between corilagin/TGG and RBD/N501Y for the center of the biggest cluster (total population of 99% for corilagin and 73% for TGG) are shown in Figures 3 and 4, respectively. Corilagin is forming five H-bonds with RBD (Asp467, Ser469, Thr470, Gln474 and Gly482) and is involved in five nonpolar contacts (Arg457, Gly471, Ile472, Cys480 and Asn481). TGG is again mainly stabilized by nonpolar contacts (Leu455, Phe456, Tyr473, Tyr489, Phe490, Pro491 and Gln493) and two H-bonds with Leu494 (mainchain).

**RBD/E484K-N501Y.** The binding affinities with RBD/E484K+N501Y are presented in Table 2. For corilagin/RBD, two binding sites are presented: the first site (from 20-70 ns), located at the interface, before the dissociation and the second site (90-100 ns), outside the interface, after the reassociation. For the first site, the binding affinities are  $-4.3 \pm 0.7$  kcal/mol (VINA) or  $-15.0 \pm 0.1$  kcal/mol (MMPBSA). In spite this high binding affinity computed with MMPBSA, corilagin still dissociates of the interface after around 80 ns. On the new binding sites, after reassociation, the binding energy is  $-4.1 \pm 1.0$  kcal/mol (VINA) or  $-1.4 \pm 0.3$  kcal/mol (MMPBSA). For TGG/RBD, we computed a binding affinity of:  $-4.9 \pm 0.4$

kcal/mol with VINA and  $-15.1 \pm 0.1$  kcal/mol with MMPBSA. Although the computed MMPBSA binding energy is similar to the one computed for corilagin on the first site, TGG stays strongly associated to the RBD for the entire simulated 100 ns. The LigPlot interaction map between corilagin/TGG and RBD/E484K-N501Y for the center of the biggest cluster (total population of 40% and 93% for corilagin and TGG, respectively) is shown in Figures 3 and 4, respectively. For corilagin, the interactions with the RBD are minimal, with the formation of only two H-bonds with Asn343 and Leu441 and two nonpolar contacts with Ala344 and Thr345. On the other hand, TGG is stabilized by the formation of four H-bonds with Asn487, Phe490, Gln493 and Ser494 and five nonpolar contacts with Tyr449, Leu452, Tyr489, Leu492 and Gly496.

**Table 2 The binding affinity between corilagin/TGG and ACE2/RBD.**

The second column shows the VINA binding affinity for the best pose found during docking. The third column shows the average VINA binding affinity computed over the interval of convergence (90-100 ns for RBD/E484K-N501Y with Corilagin and 75-100 ns for the rest) of the ligand-protein MD simulations. The fourth column shows the MMPBSA binding free energy computed over the same interval using the `g_mmpbsa` tools<sup>65</sup>. The average and standard deviation are computed using a 500-steps of bootstrap analysis and 40 ps snapshots on the interval of convergence (90-100 ns for RBD/E484K-N501Y with corilagin and 75-100 ns for the rest). For RBD/E484K-N501Y with corilagin, we compare the results for the two binding sites. Site 1 is the corilagin's localization before its disassociation (from 20-70 ns). Site 2 is the corilagin's localization after its reassociation (from 90-100 ns)

	Binding affinity (kcal/mol)		
	Docking VINA	MD VINA	MD MMPBSA
ACE2			
Corilagin	-7.3	$-6.1 \pm 0.5$	$-0.1 \pm 0.2$
TGG	-7.7	$-6.0 \pm 0.4$	$-14.4 \pm 0.2$
Wildtype RBD			
Corilagin	-8.1	$-5.0 \pm 0.5$	$-7.2 \pm 0.1$
TGG	-8.8	$-6.0 \pm 0.4$	$-12.8 \pm 0.4$
RBD/E484K			
Corilagin	-7.2	$-6.3 \pm 0.4$	$-7.0 \pm 0.2$
TGG	-7.8	$-4.5 \pm 0.4$	$-10.9 \pm 0.4$
RBD/N501Y			
Corilagin	-6.5	$-6.4 \pm 0.5$	$-11.3 \pm 0.1$
TGG	-7.1	$-3.8 \pm 0.5$	$-12.1 \pm 0.3$
RBD/E484K-N501Y			
Corilagin, Site 1	-7.5	$-4.3 \pm 0.7$	$-15.0 \pm 0.1$
Corilagin, Site 2	-7.5	$-4.1 \pm 1.0$	$-1.4 \pm 0.3$
TGG	-7.9	$-4.9 \pm 0.4$	$-15.1 \pm 0.1$

### 3.3.4 Corilagin ability to disrupt the ACE2/mutated RBD interactions

For **RBD/E484K**, the MMPBSA binding free energy of corilagin is almost identical to that of the WT. Yet, Table 1 shows that corilagin only interacts with three different interfacial residues in this case: Phe456; Lys458, which forms a salt-bridge with ACE2; and Tyr473, which forms a H-bond with ACE2. Thus, this association is most likely less efficient to prevent the ACE2/RBD complex formation than for the WT.

For **RBD/N501Y**, the MMPBSA binding free energy of corilagin

is significantly more negative than for the WT and E484K. However, in this binding site, corilagin is not involved in interactions with crucial residues of the interface (Table 1). It is rather located on the other side of the loop involved in the formation of the interface. It interacts with residues Glu471, Ile472 and Gln474, instead of the residues of the other side of the loop such as Tyr473, Ala475 or Gly476. Therefore, despite the increased binding affinity, corilagin would potentially be less effective when the N501Y mutation is present.

For **E484K/N501Y**, corilagin dissociates from its initial position at around 80 ns, then associates again at around 90 ns, but at a completely new location outside the RBM (Figure 1). This new docked position is characterized by a very small MMPBSA binding free energy. Moreover, Table 1 shows that, at this site, corilagin is not interacting with any RBD's residues found at the interface.

Overall, we find that corilagin binds to WT RBD and E484K with a high binding affinity in a position that could directly disrupt association of RBD with ACE2. However, when adding the N501Y mutation (alone or with E484K), corilagin binds to the RBD in a region outside of the interface and would likely have no direct effect on the RBD-ACE2 complex formation, reducing its interest as an inhibitor.

### 3.3.5 TGG ability to disrupt the ACE2/mutated RBD interactions

For **RBD/E484K**, the MMPBSA binding free energy is favorable, but TGG's location is different than in the WT as it interacts mainly with residues at the C-terminal end of the RBM (Table 1). In this position, TGG interacts with many residues forming H-bonds with ACE2 such as Gly496, Gln498, Thr500, Asn501 and Tyr505, and it even forms two H-bonds with Thr500 and Asn501.

For **RBD/N501Y**, TGG binds with an affinity similar to WT. Table 1 shows that, contrary to what is observed for corilagin, TGG takes position directly at the interface and makes contact with many nonpolar (Leu455, Phe456, Phe490, Pro491, Leu492 with H-bond formation) and polar (Lys458, Tyr473, Tyr489 and Gln493) residues.

For **E484K/N501Y**, TGG shows the highest binding affinity of all systems studied here with  $-15.1 \pm 0.1$  kcal/mol. At this position, TGG makes many nonpolar (Phe490 with H-bond, Leu492 and Gly496) and polar (Tyr449, Tyr489 and Gln493 with H-bond) contacts with residues found at the interface between ACE2 and RBD.

In summary, TGG binds to all RBD sequences studied here (without and with mutations) with high MMPBSA binding free energy. Moreover, for all these sequences, TGG binds to the protein at locations that could hinder the formation of ACE2/RBD complex and, potentially, reduce the ability of SARS-Cov-2's Spike protein to bind to the ACE2 protein on human cells.

## 4 Discussion

The current article is constructed in two parts. In the first part, we study the therapeutic potential of two small molecules, corilagin and TGG, to disrupt the association between ACE2 and the wild-type (WT) RBD; a crucial step of the infection by the virus. To do so, we used a combination of numerical tools (molecular dock-

ings, molecular dynamics and MMPBSA free energy calculations) and experimental tools (SPR and ELISA assay). In the second part, we study, using the same array of numerical tools, the impact of the main mutations (E484K and N501Y) of the variants of concerns (B.1.1.7, B.1.351, B.1.1.28).

### 4.1 Best ligand targeting ACE2 and the wild-type RBD

The first part of the article focus on the interactions between ACE2 and the WT RBD. Numerically, the ability of corilagin and TGG to potentially impair the association of ACE2 and WT RBD, is summarized in Figure S14 for the A1A2-RBM segments and Figure S15 for the HS-RBM segments; all segments containing residues involved in the ACE2/RBD association. We show that both ligands could interfere with the interaction between ACE2 and the WT RBD, more likely on the RBD's side as both ligands localization and binding energies are better than on the ACE2's side.

In order to validate our numerical conclusions and our numerical methodology, we tested the impact of corilagin and TGG on the association between ACE2 and the WT RBD using SPR and ELISA binding assay.

Our SPR experiments identify a strong association between ACE2 and the RBD, characterized by a dissociation constant of 41 nM (ACE2 with immobilized RBD) and 63 nM (RBD with immobilized ACE2) (Figure 5C). These values are compatible with recent experimental results from other research groups<sup>6,67,68</sup>. Both ligands are binding to RBD with dissociation constants in the low nanomolar range: 1.8 nM and 1.3 nM for corilagin and TGG, respectively (Figure 5A). Moreover, SPR shows that the incubation of corilagin or TGG with RBD fully inhibits its binding to immobilized ACE2 (Figure 5D). Consequently, our ligands have much more affinity compared to quercetin/ACE2 ( $K_D$  of 4.830  $\mu$ M) and quercetin/RBD ( $K_D$  of 2.210  $\mu$ M)<sup>68</sup>.

In addition to SPR experiments, we also test the binding of these ligands with RBD and ACE2 using biochemical inhibition assays. The results show that the addition of TGG or corilagin reduces from 45% (0.1  $\mu$ M) and up to 75% (10  $\mu$ M) the binding between ACE2 and RBD (Figure 6A-B). Moreover, a mixture of both ligands reduces the binding by 50%, independently of the ligand concentrations (between 0.1  $\mu$ M to 5  $\mu$ M) (Figure 6C); i.e., no synergy is observed when pre-mixing both ligands. More specifically, supporting the SPR results, we find that the ACE2/RBD binding inhibition by TGG and corilagin is mediated by their interactions with the RBD as there is no significant reduction in the ACE2/anti-ACE2 binding when adding TGG, corilagin or both (Figure 7A-C).

Interestingly, TGG and corilagin inhibit the interaction between the human ACE2 receptor and Spike protein RBD at 0.1  $\mu$ M, which is quite low as compared to the antiviral activities of tetra-TGG against SARS CoV (EC50 4.5  $\mu$ M)<sup>24</sup> and the inhibition of the binding of SARS CoV-2 spike protein RBD to ACE2 based on AlphaLISA assay (IC50 of 5.5  $\mu$ M)<sup>69</sup>. The concentration to inhibit the ACE2-RBD interaction found here is also lower than the EC50 of the polyphenol resveratrol (4.48  $\mu$ M) on SARS-CoV-2 replication in Vero cell culture<sup>70</sup> or the estimated concentration

of quercetin i.e., greater than  $25 \mu\text{M}$ <sup>71</sup>.

Natural polyphenolic compounds were reported to be sources of antiviral candidates against SARS-CoV-2, e.g. in terms of coronaviral entry inhibitors, protease inhibitors and coronavirus replication inhibitors<sup>72</sup>. However, it is important to note that some polyphenolic compounds such as magnolol and rosmarinic acid could increase the activity or expression of ACE2, and therefore aggravate SARS-CoV-2 infection<sup>73</sup>.

SARS CoV-2 induces death and injury of virus-infected cells and tissues which could be caused by high levels of inflammatory cytokines release as IL-1 $\beta$ , TNF- $\alpha$  and IL-6<sup>74,75</sup>. Interestingly, in addition to the inhibition of the binding ACE2-RBD, TGG and corilagin possess a wide range of biologic properties including anti-inflammatory, antioxidant and low toxicity. Corilagin could efficiently reduce inflammation with the reduction of the release of pro-inflammatory cytokines TNF- $\alpha$ , IL-1 $\beta$  and IL-6 through blocking the NF- $\kappa$ B pathway<sup>76-78</sup>.

Therefore, both SPR and biochemical inhibition essays show that corilagin and TGG bind to the RBD domain in a way that disrupt the ACE2/RBD interaction, in agreement with our modeling results. These findings suggest that corilagin and TGG can be useful as multi target treatment against the WT SARS CoV-2 infection.

## 4.2 The impact of the mutation of the RBD and therapeutic potential of the ligands

With the numerical protocol validated by our experimental results from SPR and ELISA assays, we also test the therapeutic potential of both molecules against three mutations (RBD/E484K, RBD/N501Y and RBD/E484K-N501Y) found in the main variants of concerns.

### 4.2.1 The impacts of RBD mutations on its structural ensemble

Recent results from experiments using pseudoparticles showed that the *Spike*-protein of the B.1.1.7, B.1.351 and B.1.1.248 mutants present no difference in terms of stability and cell entry kinetics compared to the SARS-CoV-2 WT<sup>79</sup>. Although mutant RBDs alone in solution have not been studied in the literature, they have been heavily studied in complex with ACE2. A recent Cryo-EM structure of the RBD with the N501Y mutation in complex with ACE2 showed that there was no significant changes in terms of secondary, quaternary and binding site structures compared to the RBD-WT/ACE2 complex<sup>80</sup>. The ACE2/RBD mutant complex was also studied using numerical techniques. MD simulations and principal component analysis were realized by Nelson *et al.*<sup>81</sup> on the ACE2 complex with RBD with N501Y mutation, E484K mutation and K417N+E484K+N501Y mutations. The authors showed that the complex with the E484K mutation adopts conformations that are mostly similar to the WT, while the complex with the N501Y mutation adopts conformations that are very different than the WT. The triple mutant conformational space more closely resembles the one from the WT or the E484K mutant than the N501Y mutant. Finally, Dehury *et al.*<sup>82</sup> studied the impact of multiple alanine point mutations on the ACE2/RBD complex using MD simulations. They found that the complex was sta-

ble during their simulation with no noteworthy changes in terms of secondary and quaternary structure for all mutants tested. The highest backbone RMSD they measured was  $0.33 \pm 0.09$  nm for the N501A mutant system compared to  $0.25 \pm 0.03$  nm for the WT. All these results taken together shows that the mutations have little impact on the stability, the cell's entry kinetics and the structure in the complex of RBD.

These results are compatible with what we observed for the RBD in solution; the mutations have very little impact on RBD's structure in terms of backbone RMSD, secondary structure and on the solvent accessibility of crucial residues involved in interactions with ACE2 (Figure 8). This last result suggests that the nature of the interface with ACE2 is probably similar for the WT and three mutant systems we tested. However, only extensive free-energy calculations, that are beyond the object of this work, could provide information on the impact of these mutations on the binding affinity between the two proteins. Here, these simulations are used to identify conformational ensembles representative of the mutants in order to evaluate their impact on the binding with corilagin and TGG.

### 4.2.2 Therapeutic potential of the ligands

Since the start of the pandemic, a multitude of therapeutic techniques were developed to fight against the virus; from vaccine<sup>83</sup>, to monoclonal antibody approved for emergency use by the FDA<sup>84,85</sup>. Although, the usage of small molecules as drugs that could be used against SARS-CoV-2 was heavily studied<sup>26,27</sup>, no drugs were able to be designed yet for widespread and efficient usage. However, the recent spread of multiple SARS-CoV-2 variants from the United Kingdom, South Africa and Brazil, imposes a reassessment of the efficacy of currently used treatment, as well as additional effort in drugs development. Recent results for the Novavax vaccine showed that its efficacy drop significantly for the variant from South-Africa (between 50% and 60%) compared to SARS-CoV-2 WT (89%)<sup>83</sup>. Additionally, a number of experimental<sup>79,86,87</sup> and numerical<sup>88</sup> results shows that SARS-CoV-2 mutants with the E484K mutation, like the one from Brazil and South-Africa, were partially, if not fully, resistant to the antibodies approved for emergency use by the FDA. Computed binding affinities between these antibodies and the RBD with the E484K mutation were heavily reduced compared to the WT<sup>88</sup>.

In our study, we find that the E484K mutation does not impact the therapeutic potential, contrary to what is found for the vaccine and antibodies, of either corilagin or TGG; both bind the RBD in a location that could prevent crucial interactions with ACE2 with affinities similar to the WT. On the other hand, we find that the N501Y mutation, present in the variants from the United-Kingdom, South-Africa and Brazil, highly impairs the therapeutic potential of corilagin. When only the N501Y mutation is present, corilagin binds to the RBD with a high affinity but doesn't interact with residues crucial for interaction with ACE2. The therapeutic potential of corilagin is even lower when both the N501Y and E484K mutations are present as it is unstable on the RBM. On the other hand, the N501Y mutation, alone or in pair with the E484K mutation, doesn't impact the therapeutic potential of TGG; TGG binds to the RBD at a relevant location to disrupt its interaction

with ACE2 and with a high binding affinity.

## Conclusion

In this study, the combination of numerical and experimental data shows that two natural polyphenols, corilagin and 1,3,6-tri-O-galloy- $\beta$ -D-glucose (TGG) could play a protective role in reducing the potency of WT SARS-CoV-2 by disrupting the S-protein-RBD/ACE2 receptor interface stability or the ability of the RBD of the S-protein to recognize the ACE2 receptor.

Combining molecular modelling, including molecular dynamics and protein-ligand docking, with SPR and ELISA assays, we demonstrate that the observed inhibition of the binding of Spike RBD to human ACE2 is caused mainly by the binding of these polyphenols to the RBD protein, with dissociation constants in the low nanomolar range: 1.8 nM and 1.3 nM for corilagin and TGG, respectively. Such preference would have the potential to limit physiological side-effects induced by the inhibition of ACE2.

In addition, we use the same numerical protocol to study the impact of RBD mutated sequences associated with three dominant variants — the B.1.1.7 variant, the B.1.351 variant, and the B.1.1.28, first identified in the United Kingdom, in South Africa and in Brazil respectively —, focusing on mutations affecting the interface : RBD/E484K, RBD/N501Y and RBD/E484K-N501Y.

Analysing the impact of the potential inhibitors by identifying docking sites using AutoDock VINA and further assessing the role of flexibility by running MD on the most stable ligand-protein configurations, using MMPBSA to compute the binding free-energy, we show that both molecules have the potential to bind more strongly to mutants RBD than ACE2, similarly to what is observed for WT RBD. Both also bind well to the RBD/E484K mutant compared to the WT, albeit with a significantly increased binding free energy for TGG compared to corilagin ( $-12.8 \pm 0.4$  kcal/mol vs  $-7.2 \pm 0.1$  kcal/mol for WT and  $-10.9 \pm 0.4$  vs  $-7.0 \pm 0.2$  kcal/mol for RBD/E484K), values that compare well with other potential inhibitors<sup>89</sup>. For the structures with the N501Y mutant (RBD/N501Y and RBD/E484K-N501Y), corilagin's binding localization is outside of the RBM's region that is interacting with ACE2, so that the recognition of ACE2 by SARS-CoV-2's Spike protein could still take place. On the other hand, TGG is as effective on these mutants than it is on the WT and RBD/E484K mutant.

This work strongly supports the need for further experimental assessments to evaluate the ligands' selectivity towards the virus vs. other binding sites as well as to establish their *in vivo* behavior.

## Author Contributions

**Vincent Binette:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; **Sébastien Côté:** Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; **Mohamed Haddad:** Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft; **Phuong Trang Nguyen:** Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft; **Sébastien Bélanger:** Investigation, Methodology; **Steve Bourgault:** Conceptualization, Funding acquisition, Investigation, Project administration,

Supervision, Visualization, Writing – original draft; **Charles Ramassamy:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing – original draft; **Roger Gaudreault:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing; **Normand Mousseau:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing;

## Acknowledgments

The authors acknowledge support from the Natural Science and Engineering Council of Canada (NSERC) and the Fonds de recherche du Québec - Nature et technologies (FRQ-NT #275304). They are grateful to Calcul Québec and Compute Canada for generous allocation of computer time. The authors would also like to thank La Fondation Famille Lemaire, the R. Howard Webster Foundation and Fruitomed Inc for their valuable support.

## Notes and references

- 1 World health organization *Weekly epidemiological update*, <https://www.who.int/publications/m/item/weekly-epidemiological-update---10-march-2021>, Accessed: 2021-03-13.
- 2 M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche *et al.*, *cell*, 2020, **181**, 271–280.
- 3 P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang *et al.*, *nature*, 2020, **579**, 270–273.
- 4 S. P. Singh, M. Pritam, B. Pandey and T. P. Yadav, *Journal of Medical Virology*, 2021, **93**, 275–299.
- 5 W. Ni, X. Yang, D. Yang, J. Bao, R. Li, Y. Xiao, C. Hou, H. Wang, J. Liu, D. Yang *et al.*, *Critical Care*, 2020, **24**, 1–10.
- 6 J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang *et al.*, *Nature*, 2020, **581**, 215–220.
- 7 A. Rambaut, N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D. L. Robertson and E. Volz, *Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations*, <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-563>.
- 8 H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E. J. San, N. Msomi *et al.*, *Nature*, 2021, 1–8.
- 9 P. C. Resende, J. F. Bezerra, R. H. T. de Vasconcelos, I. Arantes, L. Appolinario, A. C. Mendonça, A. C. Paixao, A. C. D. Rodrigues, T. Silva, A. S. Rocha *et al.*, *Spike E484K mutation in the first SARS-CoV-2 reinfection case confirmed in Brazil, 2020*, <https://virological.org/t/spike-e484k-mutation-in-the-first-sars-cov-2-reinfection-case-584>.
- 10 T. N. Starr, A. J. Greaney, S. K. Hilton, D. Ellis, K. H. Crawford,

- A. S. Dingens, M. J. Navarro, J. E. Bowen, M. A. Tortorici, A. C. Walls *et al.*, *Cell*, 2020, **182**, 1295–1310.
- 11 H. Zhang, J. M. Penninger, Y. Li, N. Zhong and A. S. Slutsky, *Intensive care medicine*, 2020, **46**, 586–590.
- 12 B. Qiao and M. Olvera de la Cruz, *ACS nano*, 2020, **14**, 10616–10623.
- 13 B. A. Johnson, X. Xie, A. L. Bailey, B. Kalveram, K. G. Lokugamage, A. Muruato, J. Zou, X. Zhang, T. Juelich, J. K. Smith *et al.*, *Nature*, 2021, **591**, 293–299.
- 14 C. Ramassamy, *European journal of pharmacology*, 2006, **545**, 51–64.
- 15 R. Gaudreault and N. Mousseau, *Current Alzheimer Research*, 2019, **16**, 529–543.
- 16 R. Gaudreault, V. Hervé, T. G. van de Ven, N. Mousseau and C. Ramassamy, *Journal of Alzheimer's Disease*, 2021, **81**, 1–23.
- 17 A. Durazzo, M. Lucarini, E. B. Souto, C. Cicala, E. Caiazzo, A. A. Izzo, E. Novellino and A. Santini, *Phytotherapy Research*, 2019, **33**, 2221–2243.
- 18 T.-c. Lin, F.-I. Hsu and J.-T. Cheng, *Journal of Natural Products*, 1993, **56**, 629–632.
- 19 J. Xu, G. Zhang, Y. Tong, J. Yuan, Y. Li and G. Song, *International journal of molecular medicine*, 2019, **43**, 967–979.
- 20 F. Notka, G. Meier and R. Wagner, *Antiviral research*, 2003, **58**, 175–186.
- 21 F. Notka, G. Meier and R. Wagner, *Antiviral Research*, 2004, **64**, 93–102.
- 22 L.-T. Lin, W.-C. Hsu and C.-C. Lin, *Journal of traditional and complementary medicine*, 2014, **4**, 24–35.
- 23 S.-G. Yeo, J. H. Song, E.-H. Hong, B.-R. Lee, Y. S. Kwon, S.-Y. Chang, S. H. Kim, S. won Lee, J.-H. Park and H.-J. Ko, *Archives of pharmacal research*, 2015, **38**, 193–202.
- 24 L. Yi, Z. Li, K. Yuan, X. Qu, J. Chen, G. Wang, H. Zhang, H. Luo, L. Zhu, P. Jiang *et al.*, *Journal of virology*, 2004, **78**, 11334–11339.
- 25 L. Du, Y. He, Y. Zhou, S. Liu, B.-J. Zheng and S. Jiang, *Nature Reviews Microbiology*, 2009, **7**, 226–236.
- 26 A. da Silva Antonio, L. S. M. Wiedemann and V. F. Veiga-Junior, *RSC Advances*, 2020, **10**, 23379–23393.
- 27 I. L. Paraiso, J. S. Revel and J. F. Stevens, *Current Opinion in Food Science*, 2020, **32**, 149–155.
- 28 K. Miroshnychenko and A. V. Shestopalova, [10.26434/chemrxiv.12377870.v1](https://doi.org/10.26434/chemrxiv.12377870.v1), 2020.
- 29 P. Pandey, J. S. Rane, A. Chatterjee, A. Kumar, R. Khan, A. Prakash and S. Ray, *Journal of Biomolecular Structure and Dynamics*, 2020, 1–11.
- 30 R. Gaudreault, T. G. van de Ven and M. A. Whitehead, *Molecular modeling annual*, 2002, **8**, 73–80.
- 31 R. Gaudreault, T. van de Ven and M. Whitehead, *Journal of Physical Chemistry A*, 2006, 3692–3702.
- 32 B. U. Reddy, R. Mullick, A. Kumar, G. Sharma, P. Bag, C. L. Roy, G. Sudha, H. Tandon, P. Dave, A. Shukla *et al.*, *Antiviral research*, 2018, **150**, 47–59.
- 33 X. Li, Y. Deng, Z. Zheng, W. Huang, L. Chen, Q. Tong and Y. Ming, *Biomedicine & Pharmacotherapy*, 2018, **99**, 43–50.
- 34 J.-T. Cheng, T.-C. Lin and F.-L. Hsu, *Canadian journal of physiology and pharmacology*, 1995, **73**, 1425–1429.
- 35 F. Jin, D. Cheng, J.-Y. Tao, S.-L. Zhang, R. Pang, Y.-J. Guo, P. Ye, J.-H. Dong and L. Zhao, *BMC gastroenterology*, 2013, **13**, 1–10.
- 36 V. Salmaso and S. Moro, *Frontiers in pharmacology*, 2018, **9**, 923.
- 37 J. Loschwitz, A. Jäckering, M. Keutmann, M. Olagunju, R. J. Eberle, M. A. Coronado, O. O. Olubiyi and B. Strodel, *Inorganic Chemistry*, 2021, **111**, 104862.
- 38 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Accounts of Chemical Research*, 2000, **33**, 889–897.
- 39 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19–25.
- 40 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *Journal of chemical theory and computation*, 2015, **11**, 3696–3713.
- 41 S. Nosé, *The Journal of chemical physics*, 1984, **81**, 511–519.
- 42 W. G. Hoover, *Physical review A*, 1985, **31**, 1695.
- 43 M. Parrinello and A. Rahman, *Journal of Applied physics*, 1981, **52**, 7182–7190.
- 44 T. Darden, D. York and L. Pedersen, *The Journal of chemical physics*, 1993, **98**, 10089–10092.
- 45 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of chemical physics*, 1995, **103**, 8577–8593.
- 46 B. Hess, H. Bekker, H. J. Berendsen and J. G. Fraaije, *Journal of computational chemistry*, 1997, **18**, 1463–1472.
- 47 S. Miyamoto and P. A. Kollman, *Journal of computational chemistry*, 1992, **13**, 952–962.
- 48 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. Van Gunsteren and A. E. Mark, *Angewandte Chemie International Edition*, 1999, **38**, 236–240.
- 49 Schrödinger, LLC, PyMOL: The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.
- 50 O. Trott and A. J. Olson, *Journal of computational chemistry*, 2010, **31**, 455–461.
- 51 Z. Wang, H. Sun, X. Yao, D. Li, L. Xu, Y. Li, S. Tian and T. Hou, *Physical Chemistry Chemical Physics*, 2016, **18**, 12964–12975.
- 52 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *Journal of cheminformatics*, 2011, **3**, 33.
- 53 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of computational chemistry*, 2004, **25**, 1157–1174.
- 54 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *The Journal of Physical Chemistry*, 1993, **97**, 10269–10280.
- 55 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *Journal of the American Chemical Society*, 2002, **115**, 9620–9631.
- 56 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *Journal of molecular graphics and modelling*, 2006, **25**, 247–260.
- 57 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A.

- Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian~16 Revision C.01*, 2016, Gaussian Inc. Wallingford CT.
- 58 A. W. S. Da Silva and W. F. Vranken, *BMC research notes*, 2012, **5**, 367.
- 59 W. Kabsch and C. Sander, *Biopolymers: Original Research on Biomolecules*, 1983, **22**, 2577–2637.
- 60 D. J. Barlow and J. Thornton, *Journal of molecular biology*, 1983, **168**, 867–885.
- 61 A. C. Wallace, R. A. Laskowski and J. M. Thornton, *Protein engineering, design and selection*, 1995, **8**, 127–134.
- 62 R. A. Laskowski and M. B. Swindells, *LigPlot+: multiple ligand–protein interaction diagrams for drug discovery*, 2011.
- 63 S. Genheden and U. Ryde, *Expert opinion on drug discovery*, 2015, **10**, 449–461.
- 64 E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. Zhang and T. Hou, *Chemical reviews*, 2019, **119**, 9478–9508.
- 65 R. Kumari, R. Kumar, O. S. D. D. Consortium and A. Lynn, *Journal of chemical information and modeling*, 2014, **54**, 1951–1962.
- 66 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proceedings of the National Academy of Sciences*, 2001, **98**, 10037–10041.
- 67 D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham and J. S. McLellan, *Science*, 2020, **367**, 1260–1263.
- 68 B. Pan, S. Fang, J. Zhang, Y. Pan, H. Liu, Y. Wang, M. Li and L. Liu, *Computational and structural biotechnology journal*, 2020, **18**, 3518–3527.
- 69 Q. M. Hanson, K. M. Wilson, M. Shen, Z. Itkin, R. T. Eastman, P. Shinn and M. D. Hall, *ACS Pharmacology & Translational Science*, 2020, **3**, 1352–1360.
- 70 M. Yang, J. Wei, T. Huang, L. Lei, C. Shen, J. Lai, M. Yang, L. Liu, Y. Yang, G. Liu *et al.*, *Phytotherapy Research*, 2020, **35**, 1127–1129.
- 71 G. Williamson and A. Kerimi, *Biochemical pharmacology*, 2020, 114123.
- 72 Y. S. Wijayasinghe, P. Bhansali, R. E. Viola, M. A. Kamal and N. K. Poddar, *Current pharmaceutical design*, 2020, **26**, 1–25.
- 73 A. Gasparotto, S. E. L. Tolouei, F. A. d. R. Lívero, F. Gasparotto, T. Boeing and P. de Souza, *Current Pharmaceutical Design*, 2021, **27**, 1588–1596.
- 74 M. Z. Tay, C. M. Poh, L. Rénia, P. A. MacAry and L. F. Ng, *Nature Reviews Immunology*, 2020, **20**, 363–374.
- 75 C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu *et al.*, *The lancet*, 2020, **395**, 497–506.
- 76 L. Zhao, S.-L. Zhang, J.-Y. Tao, R. Pang, F. Jin, Y.-J. Guo, J.-H. Dong, P. Ye, H.-Y. Zhao and G.-H. Zheng, *International immunopharmacology*, 2008, **8**, 1059–1064.
- 77 X.-R. Dong, M. Luo, L. Fan, T. Zhang, L. Liu, J.-H. Dong and G. Wu, *International journal of molecular medicine*, 2010, **25**, 531–536.
- 78 K. Youn, S. Lee, W.-S. Jeong, C.-T. Ho and M. Jun, *JOURNAL OF MEDICINAL FOOD*, 2016, **19**, 901–911.
- 79 M. Hoffmann, P. Arora, R. Groß, A. Seidel, B. F. Hörnich, A. S. Hahn, N. Krüger, L. Graichen, H. Hofmann-Winkler, A. Kempf, M. S. Winkler, S. Schulz, H.-M. Jäck, B. Jahrsdörfer, H. Schrezenmeier, M. Müller, A. Kleger, J. Münch and S. Pöhlmann, *Cell*, 2021, **184**, 2384–2393.e12.
- 80 X. Zhu, D. Mannar, S. S. Srivastava, A. M. Berezuk, J.-P. Demers, J. W. Saville, K. Leopold, W. Li, D. S. Dimitrov, K. S. Tuttle *et al.*, *PLoS biology*, 2021, **19**, e3001237.
- 81 G. Nelson, O. Buzko, P. R. Spilman, K. Niazi, S. Rabizadeh and P. R. Soon-Shiong, *BioRxiv*, 2021.
- 82 B. Dehury, V. Raina, N. Misra and M. Suar, *Journal of Biomolecular Structure and Dynamics*, 2020, 1–15.
- 83 M. Wadman and J. Cohen, *Science*, 2021.
- 84 A. Baum, D. Ajithdoss, R. Copin, A. Zhou, K. Lanza, N. Negrón, M. Ni, Y. Wei, K. Mohammadi, B. Musser *et al.*, *Science*, 2020, **370**, 1110–1115.
- 85 P. Chen, A. Nirula, B. Heller, R. L. Gottlieb, J. Boscia, J. Morris, G. Huhn, J. Cardona, B. Mocherla, V. Stosor *et al.*, *New England Journal of Medicine*, 2021, **384**, 229–237.
- 86 P. Wang, M. S. Nair, L. Liu, S. Iketani, Y. Luo, Y. Guo, M. Wang, J. Yu, B. Zhang, P. D. Kwong *et al.*, *Nature*, 2021, 1–6.
- 87 M. Widera, A. Wilhelm, S. Hoehl, C. Pallas, N. Kohmer, T. Wolf, H. F. Rabenau, V. M. Corman, C. Drosten, M. J. Vehreschild *et al.*, *medRxiv*, 2021.
- 88 L. Wu, C. Peng, Z. Xu and W. Zhu, *10.26434/chemrxiv.13897091.v1*, 2021.
- 89 S. K. Kwofie, E. Broni, S. O. Asiedu, G. B. Kwarko, B. Dankwa, K. S. Enninful, E. K. Tiburu and M. D. Wilson, *Molecules*, 2021, **26**, 406.