

## Molecular Docking Studies of Curcumin Analogues against SARS-CoV-2 Spike Protein

Jéssica R. Nogueira,<sup>a,b</sup> Flávia A. Verza,<sup>a</sup> Felipe Nishimura,<sup>a</sup> Umashankar Das,<sup>c</sup> Ícaro P. Caruso,<sup>d,e,f</sup>  
Ana L. Fachin,<sup>a,g</sup> Jonathan R. Dimmock<sup>\*c</sup> and Mozart Marins<sup>†a,g,h</sup>

<sup>a</sup>Unidade de Biotecnologia, Universidade de Ribeirão Preto, 14096-900 Ribeirão Preto-SP, Brazil

<sup>b</sup>Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas (IFSULDEMINAS),  
37903-358 Passos-MG, Brazil

<sup>c</sup>College of Pharmacy and Nutrition, University of Saskatchewan, 110 Science Place, Saskatoon,  
Saskatchewan S7N 5C9, Canada

<sup>d</sup>Centro Multiusuário de Inovação Biomolecular (CMIB), Instituto de Biociências, Letras e Ciências Exatas,  
Universidade Estadual Paulista (Unesp), 15054-000 São José do Rio Preto-SP, Brazil

<sup>e</sup>Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas,  
Universidade Estadual Paulista (Unesp), 15054-000 São José do Rio Preto-SP, Brazil

<sup>f</sup>Centro Nacional de Ressonância Magnética Nuclear Jiri Jonas,  
Instituto de Bioquímica Médica Leopoldo de Meis (IBqM) e  
Centro Nacional de Biologia Estrutural e Bioimagem (CENABIO),  
Universidade Federal do Rio de Janeiro (UFRJ), 21941-590 Rio de Janeiro-RJ, Brazil

<sup>g</sup>Curso de Medicina, Universidade de Ribeirão Preto, 14096-900 Ribeirão Preto-SP, Brazil

<sup>h</sup>Curso de Ciências Farmacêuticas, Universidade de Ribeirão Preto, 14096-900 Ribeirão Preto-SP, Brazil

Severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) is the etiologic agent of the current pandemic of coronavirus disease 2019 (COVID-19) that has inflicted the loss of thousands of lives worldwide. The coronavirus surface spike (S) glycoprotein is a class I fusion with a S1 domain which is attached to the human angiotensin converting enzyme 2 (ACE2) receptor, and a S2 domain which enables fusion with the host cell membrane and internalization of the virus. Curcumin has been suggested as a potential drug to control inflammation and as a potential inhibitor of S protein, but its therapeutic effects are hampered by poor bioavailability. We performed a molecular docking and dynamic study using 94 curcumin analogues designed to have improved metabolic stability against the SARS-CoV-2 spike protein and compared their affinity with curcumin and other potential inhibitors. The docking analysis suggested that the S2 domain is the main target of these compounds and compound 2606 displayed a higher binding affinity ( $-9.6 \text{ kcal mol}^{-1}$ ) than curcumin ( $-6.8 \text{ kcal mol}^{-1}$ ) and the Food and Drug Administration (FDA) approved drug hydroxychloroquine ( $-6.3 \text{ kcal mol}^{-1}$ ). Further additional validation *in vitro* and *in vivo* of these compounds against SARS-CoV-2 may provide insights into the development of a drug that prevents virus entry into host cells.

**Keywords:** COVID-19, coronavirus, SARS, curcumin, molecular docking

### Introduction

The SARS-associated coronavirus (SARS-CoV) is the causative agent of severe acute respiratory syndrome (SARS), which emerged as an epidemic from 2002 to 2003. More than 8000 cases were reported worldwide and before

being brought under control, it took the lives of about 10% of those infected.<sup>1</sup> We are now facing a new SARS-associated coronavirus pandemic and the first cases were detected in Wuhan City, Hubei Province of China during December 2019. It is transmitted from human-to-human via sneezing, coughing and respiratory droplets and until February 2021, it has infected in excess of 107 million people around the world and caused over 2.3 million deaths.<sup>2,3</sup>

\*e-mail: jr.dimmock@usask.ca; mmarins@gmb.bio.br

The new virus was named by World Health Organization (WHO) as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the associated disease Coronavirus Disease-2019, also known as COVID-19.<sup>4</sup>

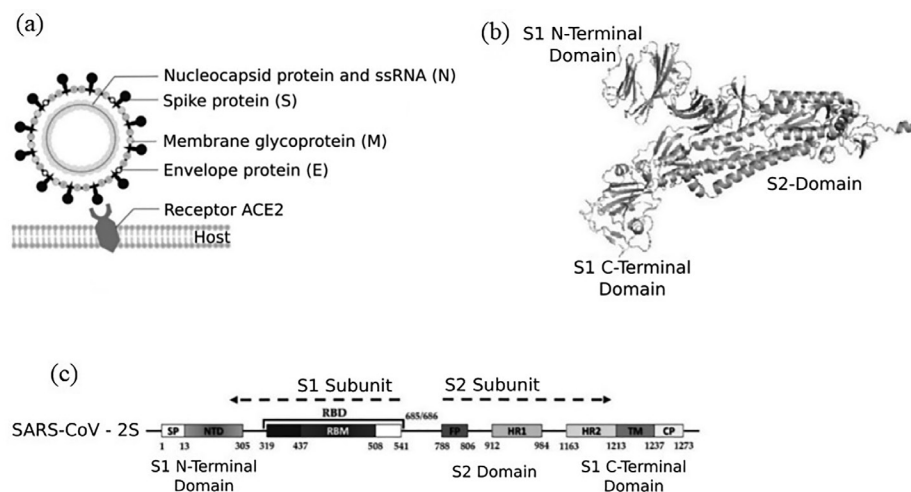
Coronaviruses are part of the Coronaviridae family of the order Nidovirales. Within this subfamily there are four genera, *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* and *Deltacoronavirus*.<sup>5-8</sup> These viruses can be found in several animals, such as birds and mammals, which are the possible transmitters of the virus to humans.<sup>9</sup> Phylogenetic studies revealed that SARS-CoV-2 is a  $\beta$  coronavirus of 2B group related to other bat SARS-related coronaviruses such as SARS-CoV and the Middle Eastern respiratory syndrome coronavirus (MERS-CoV) which frequently has been associated with fatal diseases.<sup>4,10</sup>

SARS-CoV-2 is a positive-sense single-stranded ribonucleic acid (RNA) virus with a genome of 29.891 kb in size, enclosed by a 5'-cap and 3' poly-A tail, coding 10 open reading frames (ORF). It has unique genomic features that determines diversification into at least three different genetic clades,<sup>5</sup> but maintains structural proteins like those of other members of the Coronavirus family of which the major ones are the spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins.<sup>6</sup> One of the main differences between the virus found in humans and the ones found in animals, is in the spike (S) protein.<sup>7</sup> This protein has two subunits, S1 and S2, the first being responsible for the adherence of viruses to the host cells and the second for the fusion of the virus and cell membranes.<sup>8</sup> The adhesion process of the S1 subunits takes place through the receptor binding domains (RBD, receptor-binding domain),

which in the case of SARS-CoV and SARS-CoV-2, interact with ACE2 human cell receptors (angiotensin, converting enzyme 2),<sup>9</sup> Figure 1. The structure of the ACE2 ectodomain,<sup>11</sup> presents a claw-shaped N-terminal peptidase, with the active site at the base of a deep groove and a C-terminal "collectrin" domain.<sup>12</sup> A fragment of the S1 region, residues 318 to 510, is sufficient for binding to the peptidase domain of ACE2.<sup>13,14</sup> The receptor-binding domain (RBD) determines the virus-receptor interaction as well as the viral host range.<sup>15</sup> Changes in some residues in the RBD can cause efficient transmission between species and is important for determining host specificity.<sup>16</sup>

The huge negative health and economic impact of COVID-19 demands intensive research on treatment and though different vaccines have reached human trials and commercialization,<sup>17</sup> results in animals indicate that they protect against development of COVID-19 but not of infection by SARS-CoV-2.<sup>18</sup> Moreover, the genetic diversification of SARS-CoV-2 points out that, as for other corona viruses, it is unlikely that there will be vaccines for life-long protection.<sup>19</sup> Therefore, chemotherapeutics to stop virus replication and spread is still an important objective. The S protein is the major antigen responsible for inducing effective neutralizing antibodies, and therefore not only a target for vaccines but also for antiviral drugs.

Natural compounds such as flavonoids, alkaloids, lipids, benzenoids, lignans, neolignans and synthetic analogues have been investigated for their potential therapeutic effects against SARS-CoV-2.<sup>20,21</sup> Among their diverse bioactive activities they have antiviral properties and have demonstrated significant interaction with TMPRSS2 residues (essential for viral dissemination



**Figure 1.** (a) Schematic diagram of the structure of SARS-CoV-2 and host cell membrane with ACE2 receptor. (b) Schematic representation of SARS-CoV-2 spike protein with domains and subunits. S1 subunit contains signal peptide (SP, 1-13 aa), N-terminal domain (NTD, 14-305 aa), receptor binding domain (RBD, 319-541 aa) and receptor binding motif (RBM, 437-508 aa). S2 subunit contains fusion peptide (FP, 788-806 aa), heptad repeat 1 (HR1, 912-984 aa), heptad repeat 2 (HR2, 1163-1213 aa), transmembrane domain (TM, 1214-1237 aa) and cytoplasmic domain (CP, 1238-1273 aa). (c) Monomer of spike glycoprotein with different domains.

and pathogenicity) at the active site of S proteins.<sup>21</sup> Among 33 molecules studied, a natural compound (rutin) showed greater efficiency in binding to the active site of the SARS-CoV-2 protease, than a control antiviral drug (ritinovir), an antiprotozoal drug (emetin) and hesperidin which is a natural compound.<sup>22</sup>

Another natural compound studied for COVID-19 treatment is curcumin [(1*E*,6*E*)-1,7-bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadiene-3,5-dione], a dietary phytochemical constituent of turmeric powder derived from the rhizomes of *Curcuma longa*.<sup>23</sup> It has several benefits in controlling inflammation,<sup>24</sup> which is a common reaction caused by SARS-CoV-2.<sup>25</sup> In addition, curcumin has been reported to have a good binding affinity and efficient pharmacokinetic parameters, suggesting its use as a potential inhibitor of the S protein.<sup>26,27</sup> On the other hand the therapeutic effects of curcumin are hampered by its poor absorption after ingestion, quick degradation under physiological pH conditions, as well as rapid metabolism and elimination.<sup>23</sup>

Molecular docking is a computational technique which is used to access information about the interaction between proteins and small ligand molecules.<sup>28</sup> Since this interaction occurs in several ways, molecular docking performs simulations that can predict the best conformations and orientations of the ligands in the binding sites of the target protein.<sup>29</sup> Hence, this technique has been used to design new drugs. Recently, molecular docking studies conducted by Patel *et al.*<sup>27</sup> have shown that curcumin and especially its derivative bisdemethoxycurcumin showed good binding affinity to the spike protein of both the SARS-CoV and SARS-CoV-2 viruses. Curcumin and this analog have a  $\beta$ -diketone moiety which is substrate for liver aldoketo reductases and may be one of the reasons for the rapid metabolism of curcumin *in vivo*.<sup>30</sup>

Many research groups have circumvented this problem by generating analogous compounds with a mono-ketone moiety to improve the stability and metabolic profiles of curcumin.<sup>31</sup> In the present study, we conducted molecular docking studies to assess the interaction of a series of such compounds with the S1 and S2 subunit of the SARS-CoV-2 spike protein. We found that compound 2606 has potential as an inhibitor of the S2 domain and could be explored to inhibit the fusion of SARS-CoV-2 and host cell membranes, which are a potential therapeutic target for micromolecules.

## Methodology

Molecular docking analysis of complexes of X-ray crystallographic structure of COVID-19 spike protein from 6YVB access code,<sup>8</sup> with inhibitors, were performed using

AutoDock Vina version 1.1.2 and AutoDock 4.<sup>32,33</sup> The spike protein and ligand molecules were prepared through AutoDockTools,<sup>34</sup> for molecular docking simulations. In a first blind docking experiment, the binding affinity of the whole spike protein and of the 94 compounds (Supplementary Information (SI) section, Figure S1) used in this work were analyzed (Table 1). This analysis was also performed for 15 inhibitors described in the literature (structures available in PubChem). The three-dimensional structure of the curcumin analogues was designed with the Gaussian 09 program,<sup>35</sup> which was optimized using the semi-empirical method PM6.<sup>36</sup> Hydrogens from the protein and ligands attached to non-polar atoms were removed and the protonation status of the histidine residues of the spike protein was defined according to the PROPKA 3.0 program,<sup>37</sup> of the server PDB2PQR version 2.0.0,<sup>38</sup> considering pH 7.0. The rigid root of the ligands was identified automatically, defining all possible rotatable covalent connections as active to perform torsions in the conformational search at the connection site.

The search space in the blind docking simulations covering the whole spike protein was included in a box of  $76 \times 92 \times 126$  Å, 1 Å spacing, for the AutoDock and AutoDock Vina programs. In the AutoDock Vina program, 100 exhaustiveness was used. In the AutoDock program, 100 runs were carried out, a population of 250 was used, and the maximum number of evaluations were equal to 25,000,000. Remaining parameters were set to program standards. In this program, Lamarckian genetic algorithm (LGA), root-square-mean deviation tolerance (RMSD) was used for the 2.0 Å cluster analysis and the internal electrostatic energy was calculated. The coordinates of the center of the conformational search box at the protein binding site were defined according to the position of the protein. After the blind docking simulation, the ten ligands with the highest binding affinity were selected for a simulation of local molecular docking. In this simulation, the AutoDock Vina program was used, considering the domain found in the poses of the greatest binding affinities of the blind docking simulation. In this second experiment, local docking was carried out in the S1 domain of the C-terminal and in the S2 domain of the protein, since no ligand showed greater binding affinity in the S1 portion of the N-terminal of the protein. The local docking using the spike protein C-terminal domain was performed with a box of  $58 \times 54 \times 50$  Å. A box of  $44 \times 44 \times 118$  Å was used for the remaining compounds which displayed the highest binding affinity to the S2 domain. A spacing of 1 Å and 100 exhaustiveness was considered for both domains. The LigPlot+ program<sup>39</sup> was used to calculate the residues formed between the spike protein and selected ligands.

The bioavailability was also calculated for the ligands. The SwissADME program<sup>40</sup> was used for these calculations, which uses the bioavailability score described by Martin.<sup>41</sup>

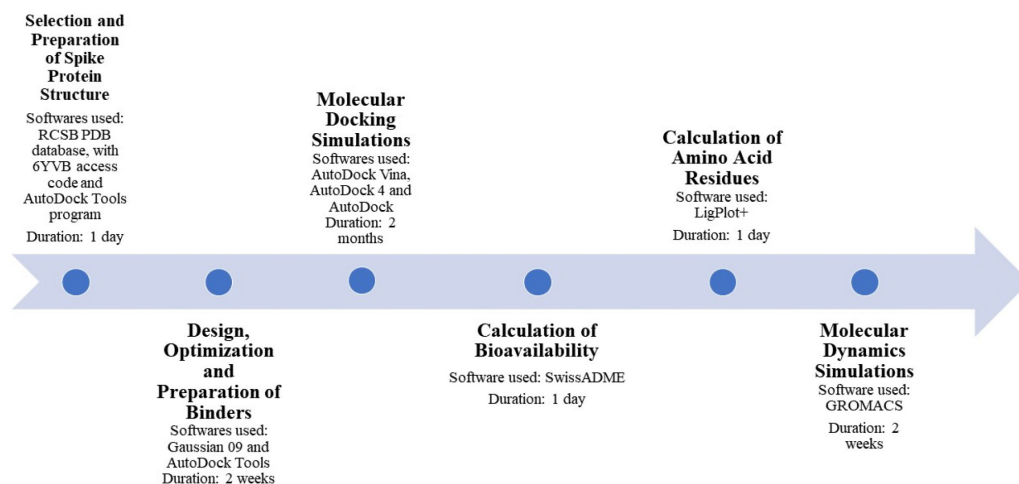
Afterwards molecular dynamics (DM) simulations were performed using the GROMACS version 5.0.1 program.<sup>42</sup> The DM simulations occurred with the spike protein S2 domain, since the ligands selected in the docking assays were part of this domain. The trajectories had 50 ns and used the CHARMM force field and the TIP3P water model.<sup>43</sup> The initial position of the compounds used in the DM simulations was obtained from the output of the AutoDock Vina program.

In the simulations, a 74 Å cubic box and solvation in a solution of 100 mmol L<sup>-1</sup> of NaCl in water was considered. The protonation status of the residues was defined considering the results obtained in PROPKA 3.0,<sup>37</sup> using a pH of 7.0. Periodic boundary conditions and isothermal-isobaric (NPT) ensemble were used in all simulations. For the system, a temperature of 298 K (25 °C) and pressure of 1.0 bar was used, through the modified Berendsej thermostat ( $\tau_t = 0.1$  ps) and the Parrinello-Rahman barostat ( $\tau_p = 2.0$  ps and compressibility  $4.5 \times 10^{-5}$  bar<sup>-1</sup>). A cutoff point of 14 Å for the Lennard-Jones and Coulomb potentials was used. To measure electrostatic interactions, the particle mesh Ewald (PME) algorithm was used. In DM simulations, a time interval of 2.0 fs was also used and all covalent bonds involving hydrogen atoms were restricted to their equilibrium distances. The conjugated gradient minimization algorithm was used to relax the atoms, in order to avoid the overlaps that occur at the beginning of the box construction process. To minimize the energy of the system, the conjugate gradient and steepest descent integrator algorithm was used, with maximum force criterion equal to 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>. After performing the tests, the trajectories obtained were analyzed according to the following parameters: (i) RMSD

of non-hydrogen atoms, for the S2 domain of the spike protein and ligands; (ii) number of hydrogen bonds, using cutting distance = 3.5 Å and the maximum angle = 45°; and (iii) number of contacts for distances < 0.6 nm. All molecular dynamics simulations were performed in duplicate and the average RMSD values obtained for protein and ligand, hydrogen bonds and number of contacts were averaged. In the SI section, in Figures S2, S3 and S4, the independent results obtained in these simulations can be observed. In Figure 2, a flowchart displays the methodology and the time spent to carry out the simulations in this work.

## Results and Discussion

The interactions of the compounds and the spike protein are listed in Table 1 according to their AutoDock Vina and AutoDock docking scores. This table also shows the bioavailability of the compounds evaluated. The highest binding affinity were observed for compounds 2606 (-9.6 kcal mol<sup>-1</sup>) and 2608 (-9.5 kcal mol<sup>-1</sup>), followed by the ligands rutin (-9.2 kcal mol<sup>-1</sup>), compound 2065 (-9.1 kcal mol<sup>-1</sup>) and hesperidin (-9.0 kcal mol<sup>-1</sup>) which displayed the highest binding affinities, considering the AutoDock Vina program. In the AutoDock program, the highest binding affinities were obtained for binders 2576 (-6.51 kcal mol<sup>-1</sup>), 2610 (-5.67 kcal mol<sup>-1</sup>), 2557 (-5.34 kcal mol<sup>-1</sup>), 2062 (-5.33 kcal mol<sup>-1</sup>) and 2606 (-5.33 kcal mol<sup>-1</sup>). It was found that, as many compounds used in simulation have a large number of torsions, the results obtained in the AutoDock Vina program may be more suitable for these simulations than those obtained in AutoDock. As an example, the rutin and hesperidin binders have the binding affinities 0.28 and -1.68 kcal mol<sup>-1</sup>, respectively, in the AutoDock program. These results are inferior when compared to those



**Figure 2.** Methodology and time used in computer simulations.

**Table 1.** AutoDock Vina, AutoDock and bioavailability scores for binding affinity between spike protein of the compounds (ligands) analyzed. The best results are highlighted in clear gray and control molecules are in bold

Binding affinity and bioavailability score							
Molecule	Binding affinity / (kcal mol <sup>-1</sup> )		Bioavailability score	Molecule	Binding affinity / (kcal mol <sup>-1</sup> )		Bioavailability score
	AutoDock Vina	Auto Dock			AutoDock Vina	Auto Dock	
<b>Rutin</b>	-9.2	0.28	0.17	<b>Resveratrol</b>	-7.9	-4.33	0.55
<b>Hesperidin</b>	-9.0	-1.68	0.17	<b>Apigenin</b>	-7.7	-4.15	0.55
<b>Genistein</b>	-8.7	-4.62	0.55	<b>Kaempferol</b>	-7.4	-3.69	0.55
<b>Fisetin</b>	-8.5	-4.18	0.55	<b>Bisdemethoxy- curcumin</b>	-7.0	-3.84	0.55
<b>Isorhamnetin</b>	-8.4	-3.44	0.55	<b>Pterostilbine</b>	-6.9	-4.16	0.55
<b>Quercetin</b>	-8.4	-3.86	0.55	<b>Curcumin</b>	-6.8	-3.33	0.55
<b>Emetine</b>	-8.3	-4.85	0.55	<b>Hydroxy-chloroquine</b>	-6.3	-2.52	0.55
<b>Luteolin</b>	-8.2	-4.00	0.55				
2606	-9.6	-5.33	0.55	2569	-7.6	-4.37	0.17
2608	-9.5	-4.5	0.55	2085	-7.6	-4.65	0.55
2065	-9.1	-5.23	0.55	2144	-7.6	-4.73	0.55
2578	-8.8	-4.95	0.55	2557	-7.6	-5.34	0.55
2610	-8.7	-5.67	0.17	2568	-7.6	-4.40	0.17
2582	-8.7	-4.96	0.17	2561	-7.5	-3.11	0.55
1833	-8.6	-4.47	0.55	2603	-7.5	-5.06	0.55
2605	-8.6	-4.63	0.55	2616	-7.5	-4.70	0.55
2138	-8.5	-4.84	0.55	2069	-7.5	-4.40	0.55
2444	-8.5	-4.51	0.55	2094	-7.5	-4.40	0.55
2581	-8.5	-4.28	0.55	2101	-7.5	-3.67	0.55
2604	-8.4	-5.05	0.55	2102	-7.5	-3.94	0.55
2614	-8.4	-5.09	0.55	2376	-7.5	-4.93	0.55
2624	-8.4	-4.86	0.55	2089	-7.4	-4.58	0.55
2576	-8.4	-6.51	0.17	2448	-7.4	-5.10	0.55
2137	-8.3	-4.49	0.55	2548	-7.4	-4.67	0.55
2553	-8.3	-4.64	0.55	2549	-7.4	-3.93	0.55
2609	-8.2	-5.13	0.55	2560	-7.4	-3.56	0.55
2086	-8.2	-4.78	0.17	2615	-7.4	-4.08	0.55
2135	-8.2	-4.97	0.17	2068	-7.7	-4.78	0.17
2567	-8.2	-4.33	0.17	2554	-7.4	-4.90	0.55
2552	-8.1	-4.85	0.55	2566	-7.4	-4.29	0.17
2083	-8.0	-4.85	0.55	2570	-7.4	-3.93	0.17
2082	-8.0	-4.70	0.55	2067	-7.3	-4.67	0.55
2091	-8.0	-4.38	0.17	2090	-7.3	-4.57	0.17
2565	-8.0	-5.08	0.55	1831	-7.2	-4.62	0.55
2062	-7.9	-5.33	0.55	2572	-7.2	-3.47	0.55
2571	-7.9	-4.24	0.17	2136	-7.2	-4.59	0.55
2573	-7.9	-5.02	0.55	1834	-7.1	-4.32	0.55
2331	-7.9	-4.56	0.55	2551	-7.1	-2.98	0.55
2556	-7.9	-4.07	0.55	2559	-7.1	-3.91	0.55
2558	-7.9	-4.96	0.55	2562	-7.1	-3.48	0.55
2443	-7.8	-4.99	0.55	2162	-7.0	-4.46	0.55
2453	-7.8	-4.90	0.55	2550	-7.0	-4.02	0.55
2369	-7.8	-4.74	0.55	2611	-6.9	-4.33	0.55
2446	-7.8	-4.51	0.55	2258	-6.9	-3.02	0.55
2423	-7.7	-4.20	0.55	2291	-6.9	-3.68	0.55
2574	-7.7	-5.32	0.55	2529	-6.9	-3.88	0.55
2607	-7.7	-4.58	0.55	2527	-6.8	-4.02	0.55
2617	-7.7	-4.39	0.55	2580	-6.8	-2.84	0.17
2064	-7.7	-4.60	0.55	2118	-6.8	-3.80	0.55
2087	-7.7	-5.28	0.17	2612	-6.7	-3.82	0.55
2288	-7.7	-4.11	0.55	2613	-6.7	-3.35	0.55
2424	-7.7	-4.70	0.55	2134	-6.6	-2.24	0.17
2555	-7.7	-4.07	0.55	2564	-6.5	-2.14	0.17
2575	-7.7	-5.05	0.55	2096	-6.5	-1.91	0.17
2579	-7.7	-4.56	0.55	2105	-6.3	-3.89	0.55

obtained by other binders, such as for genistein, which has an affinity of  $-4.62 \text{ kcal mol}^{-1}$ . However, it is verified that the ligands rutin and hesperidin have 16 and 15 points of torsion, respectively, while genistein has 4 points, which can interfere in the binding affinity. Considering the bioavailability of the binders, all calculations showed values of 0.17 or 0.55 and, for this reason, were not considered for the selection of compounds of greatest interest. After the initial blind docking simulation, the ten ligands with the highest binding affinity were selected for a simulation of local molecular docking. In this simulation, the AutoDock Vina program was used considering the domain found in the poses of the greatest binding affinity of the initial simulation.

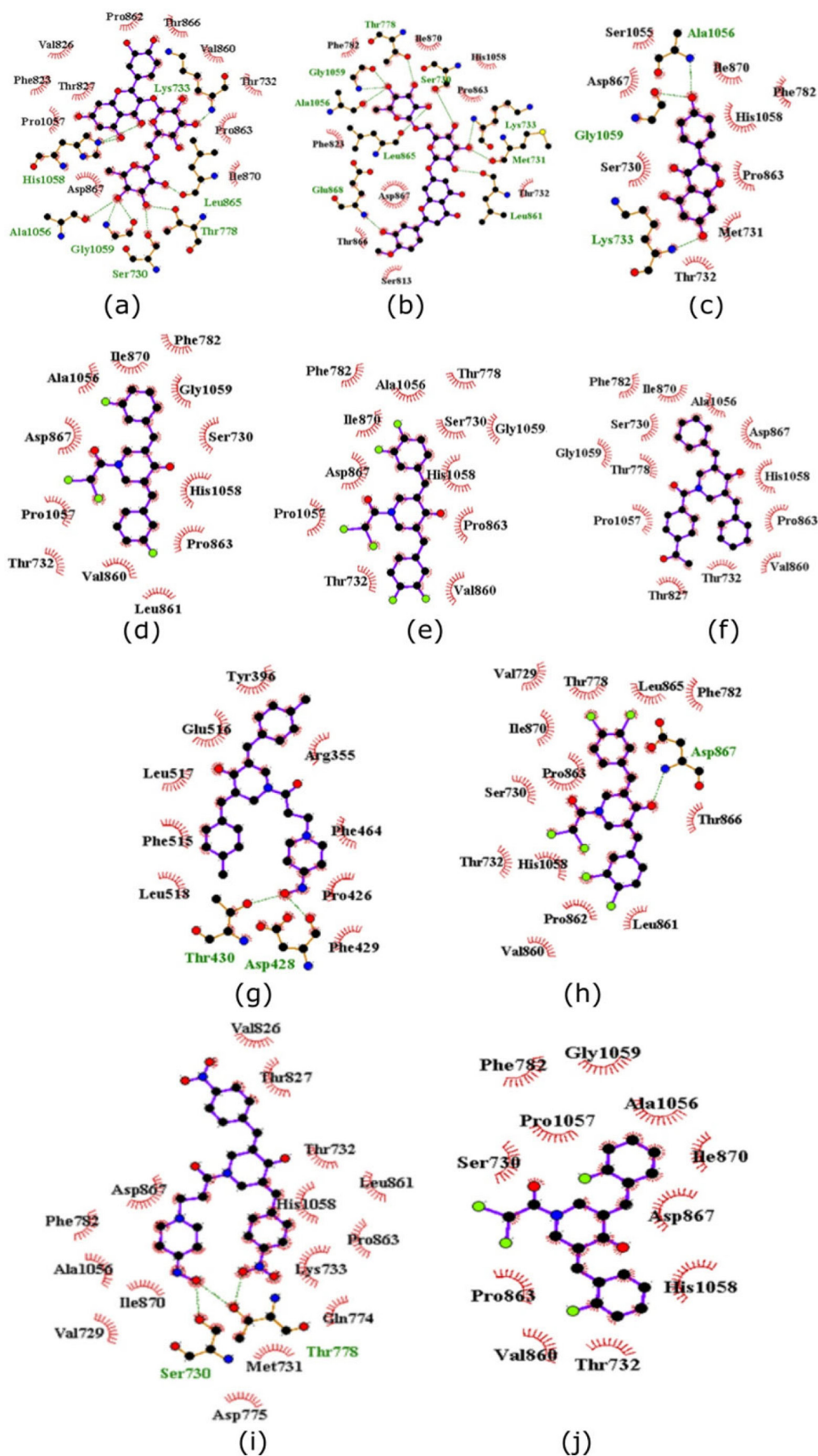
In this second experiment, local docking was carried out in the S1 domain of the C-terminal and in the S2 domain of the protein. The residues formed in this interaction were calculated (Table 2 and Figure 3) and no ligand showed greater binding affinity in the S1 portion of the N-terminal of the protein than the selected compounds. The compound 2578 showed the highest binding affinity when only the S1 domain of the spike protein C-terminal is used, while the other compounds showed better results with the S2 domain of the protein. The following residues were not

included: (i) the Asp775 residue, identified only for ligand 2582; (ii) residues Glu868 and Ser813, found only for the compound hesperidin; (iii) the Gln774 residue, identified only for linker 2582; (iv) the Phe823 residue, verified only for the compounds hesperidin and rutin; (v) residue Pro862, identified only for compounds rutin and 2610; (vi) the Ser1055 residue, found for the genistein ligand; (vii) the Thr827 residue, found for compounds 2065 and 2582; (viii) residue Val729, identified for binders 2610 and 2582; (ix) the Val826 residue, identified for the compounds rutin and 2582.

The residues Asp867, His1058, Ile870, Pro863, Ser730, Thr732 are present in all the bonds formed between ligands and the S2 domain of the spike protein, Table 3. The Ala1056 and Phe782 residues are identified for eight evaluated ligands and Gly1059 residue for seven. The remaining residues (Thr778 and Val860, for 6 ligands; Pro1057, for 5 compounds; Leu861 and Lys733, for 4 ligands; Leu865, Met731 and Thr866 for 3 compounds) were found for a smaller number of bonds. These results show that part of the domain is coincident for all ligands that showed the highest binding affinity in the spike protein S2 domain. The portion of the spike protein which is common to the compounds is featured in Figure 4.

**Table 2.** AutoDock Vina scores for blind and local docking scores of selected compounds. Spike protein domain and identified residues of selected compounds are also listed

Molecule	Binding affinity		Spike domain	Identified residues
	Blind docking / (kcal mol <sup>-1</sup> )	Local docking / (kcal mol <sup>-1</sup> )		
Rutin	-9.2	-9.1	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Leu865, Lys733, Phe823, Pro1057, Pro862, Pro863, Ser730, Thr732, Thr778, Thr827, Thr866, Val826 and Val860
Hesperidin	-9.0	-9.2	S2	Ala1056, Asp867, Glu868, Gly1059, His1058, Ile870, Leu861, Leu865, Lys733, Met731, Phe782, Phe823, Pro863, Ser730, Ser813, Thr732, Thr778 and Thr866
Genistein	-8.7	-8.7	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Lys733, Met731, Phe782, Pro863, Ser1055, Ser730 and Thr732
2606	-9.6	-9.6	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Leu861, Phe782, Pro1057, Pro863, Ser730, Thr732 and Val860
2608	-9.5	-9.5	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Phe782, Pro1057, Pro863, Ser730, Thr732, Thr778 and Val860
2065	-9.1	-9.2	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Phe782, Pro1057, Pro863, Ser730, Thr732, Thr778, Thr827 and Val860
2578	-8.8	-8.8	S1 C-terminal	Arg355, Asp428, Glu516, Leu517, Leu518, Phe429, Phe464, Phe515, Pro426, Thr430 and Tyr396
2610	-8.7	-8.8	S2	Asp867, His1058, Ile870, Leu861, Leu865, Phe782, Pro862, Pro863, Ser730, Thr732, Thr778, Thr866, Val729 and Val860
2582	-8.7	-8.8	S2	Ala1056, Asp775, Asp867, Gln774, His1058, Ile870, Leu861, Lys733, Met731, Phe782, Pro863, Ser730, Thr732, Thr778, Thr827, Val729 and Val826
2605	-8.6	-8.7	S2	Ala1056, Asp867, Gly1059, His1058, Ile870, Phe782, Pro1057, Pro863, Ser730, Thr732 and Val860



**Figure 3.** Binding interactions of spike protein and rutin (a); hesperidin (b); genistein (c); compounds 2606 (d); 2608 (e); 2065 (f); 2578 (g); 2610 (h); 2582 (i); 2605 (j).





**Table 4.** Variation in the binding affinities of the compounds

Number	R (substituent)	Binding affinity using AutoDock Vina / (kcal mol <sup>-1</sup> )
2603	H	-7.5
2605	2-F	-8.6
2606	3-F	-9.6
2607	4-F	-7.7
2608	3,4-F <sub>2</sub>	-9.5
2609	4-Cl	-8.2
2610	3,4-Cl <sub>2</sub>	-8.7
2611	4-OCH <sub>3</sub>	-6.9
2612	3,4-(OCH <sub>3</sub> ) <sub>2</sub>	-6.7
2613	3,4,5-(OCH <sub>3</sub> ) <sub>3</sub>	-6.7
2614	3,4-OCH <sub>2</sub> O	-8.4
2604	4-CH <sub>3</sub>	-8.4
2616	4-NO <sub>2</sub>	-7.5
2615	4-N(CH <sub>3</sub> ) <sub>2</sub>	-7.4

molecules are 2606 and 2608 which are identified as lead molecules for analog development.

(iii) In contrast, compounds with aryl methoxy substituents (2611-2613) have lower binding affinities than 2603. Somewhat surprising, the structurally related 3,4-methylenedioxy and 4-methyl analogs, namely 2614 and 2604, respectively, have high binding affinities.

(iv) Compounds containing a strongly electron-attracting (2616) and a strongly electron-releasing (2615) groups in the aryl rings have similar binding affinities.

(v) Most of the compounds are more potent than curcumin and bisdemethoxycurcumin which have binding affinity figures of -6.8 and -7.0 kcal mol<sup>-1</sup>, respectively.

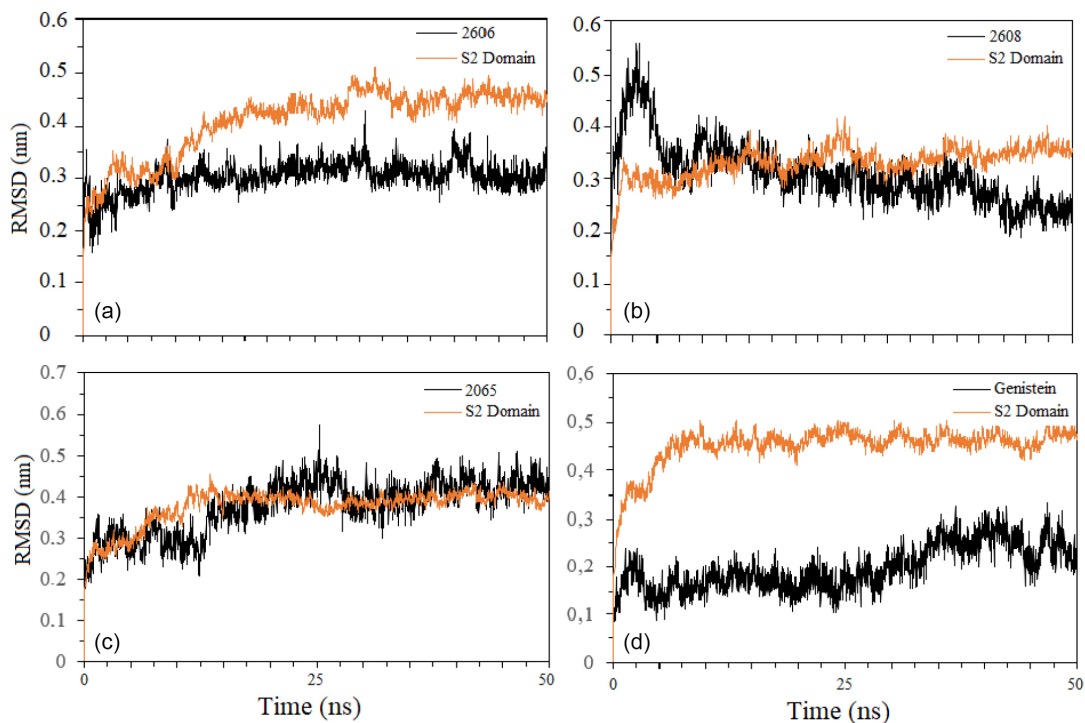
The average binding affinities of the compounds with halogens in the aryl rings namely 2605-2610 is -8.7 kcal mol<sup>-1</sup>. Hence in the future, a number of analogs with halogens in different locations in the aryl rings should be prepared and their binding affinities determined with the objective of finding novel compounds with high binding affinities.

Considering the affinities of ligation obtained in the docking programs, it appears that the curcumin analogues 2606, 2608 and 2605 showed the best affinities in the AutoDock Vina program and competitive results for the AutoDock program. For this reason, these ligands have been used in molecular dynamics simulations. In addition, it was found that the ligand genistein obtained the results of greatest interest among the compounds described in the literature and that were evaluated in this work, since this compound showed competitive affinities in the two docking programs used. Thus, this ligand was used as

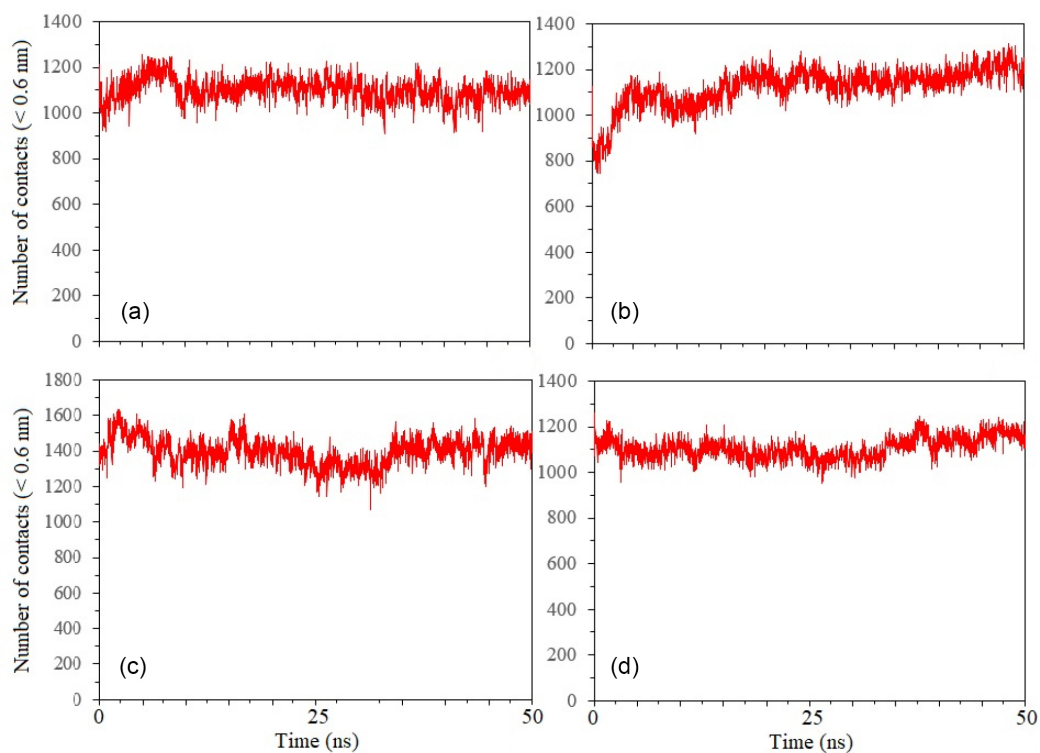
a reference in DM simulations. RMSD values of non-hydrogen atoms of S2 spike protein and ligands from the starting structure (docking pose) for the complex are reported as a function of the time at the bottom of Figure 5. In these simulations, the averages of values obtained in the duplicate simulations were calculated, from which the following RMSD values were obtained: (i) for ligand 2606 ( $0.30 \pm 0.03$ ) and S2 domain of the spike protein ( $0.40 \pm 0.06$ ), those results are shown in Figure 5a; (ii) for ligand 2608 ( $0.30 \pm 0.05$ ) and S2 domain of the spike protein ( $0.33 \pm 0.03$ ), the results of which are shown in Figure 5b; (iii) for ligand 2065 ( $0.37 \pm 0.06$ ) and S2 domain of the spike protein ( $0.37 \pm 0.04$ ), the results of which are shown in Figure 5c; (iv) for the genistein ligand ( $0.19 \pm 0.04$ ) and S2 domain of the spike protein ( $0.45 \pm 0.04$ ), those results are shown in Figure 5d. It is observed that all ligands remain stable during the DM simulations, especially the ligand 2065, which does not show considerable changes in its initial conformation along the 50 ns dynamic trajectory.

The results obtained for the number of contacts, for distances < 0.6 nm, are presented in Figure 6. In this simulation, the following numbers of contacts were obtained: (i) 2606, with  $1100 \pm 54$  contacts (Figure 6a); (ii) 2608, with  $1122 \pm 87$  contacts (Figure 6b) (iii), 2065, with  $1395 \pm 80$  contacts (Figure 6c) and (iv) genistein, with  $1107 \pm 47$  contacts (Figure 6d). In this simulation, it was found that the curcumin analog 2065 presented the largest number of contacts, among the evaluated ligands.

The average number of hydrogen bonds formed between the S2 domain of the spike protein and the selected ligands is verified in Figure 7. In this simulation,  $1 \pm 1$  hydrogen bonds were obtained for ligands 2606 (Figure 7a) and 2608 (Figure 7b). The reference ligand genistein, which had the highest number of hydrogen bonds in the simulations performed, a value of  $4 \pm 1$  bonds was obtained. The average of hydrogen bonds obtained for ligand 2065 was 0 (Figure 7c). The highest percentage of trajectory occupation in relation to the hydrogen bridges formed between the S2 domain of the spike protein and the ligands 2606, 2608 and 2065 (in the first DM simulation) occurred between His1077, which donates nitrogen and whose acceptor is O<sub>2</sub>, available in these binders. In the second DM simulation for the ligand 2065, the longest hydrogen bridge occurred between Asp849, which donates N and the 2065 which receives O<sub>3</sub>. In the first dynamics simulation of genistein ligand, it was found that the hydrogen bridge that occurs the longest in the simulation occurs between genistein, which donates O<sub>2</sub> and receives O, available at Gly1078. For the second simulation, the longest bridge occurs between the



**Figure 5.** Analysis of the parameters of 50 ns molecular dynamic simulations of the structural models of the complexes formed with 2606 (a), 2608 (b), 2065 (c) and genistein (d). At the bottom, the RMSDs of non-hydrogen atoms of the ligands (black line) and backbone atoms of the S2 domain spike protein (orange line) from the starting structure (AutoDock Vina model) for the complexes are reported as a function of time.

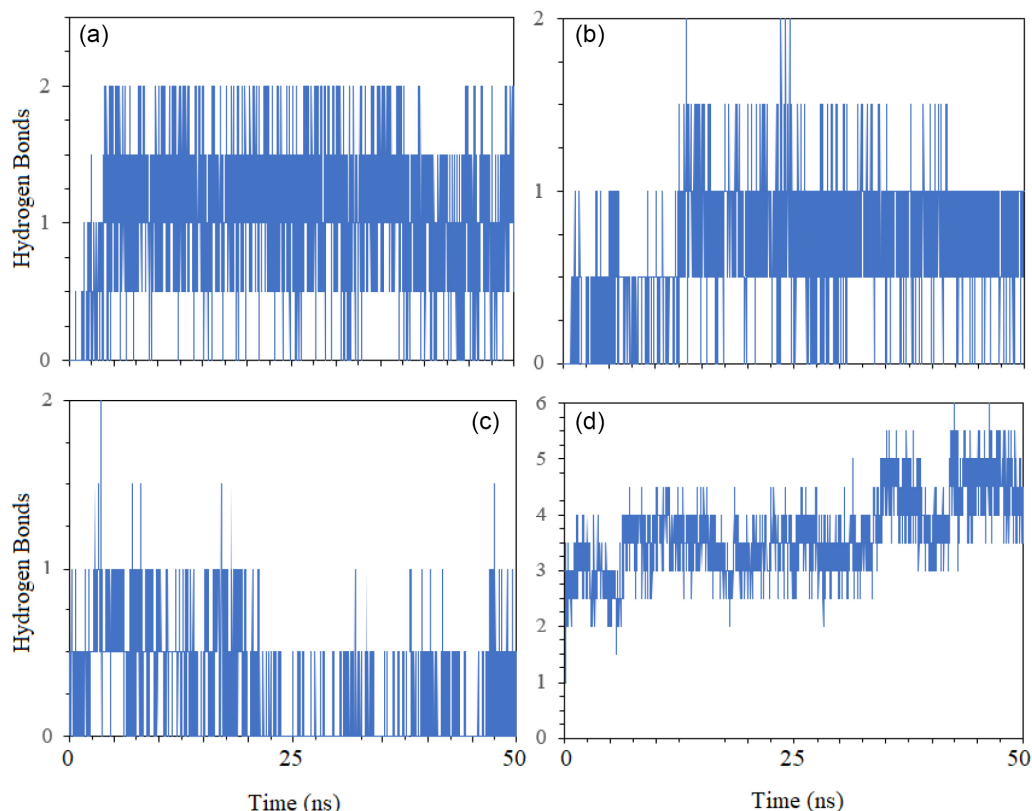


**Figure 6.** Analysis of the parameters of 50 ns molecular dynamic simulations, considering number of contacts between atoms of S2 domain spike protein and ligands 2606 (a), 2608 (b), 2065 (c) and genistein (d) for distances < 0.6 nm.

nitrogen donated by Lys752 and which has genistein's  $O_1$  as a receptor. Detailed information can be found in the SI section, in Table S1.

## Conclusions

We have shown by molecular docking that a series of



**Figure 7.** Analysis of the parameters of 50 ns molecular dynamic simulations, considering hydrogen bonds of non-hydrogen atoms between atoms of S2 domain spike protein and ligands 2606 (a), 2608 (b), 2065 (c) and genistein (d) for distances < 0.6 nm.

curcumin analogues with a mono-keto moiety have the potential to inhibit SARS-CoV-2 spike protein interaction with the host receptor. Compound 2606 displayed higher binding affinity ( $-9.6 \text{ kcal mol}^{-1}$ ) to spike protein than bisdemethoxycurcumin ( $-6.8 \text{ kcal mol}^{-1}$ ) and curcumin ( $-7.0 \text{ kcal mol}^{-1}$ ), using AutoDock Vina program. Regarding the AutoDock program, all binders with higher affinities are curcumin analogues. In the results of molecular dynamics, it was also observed that the compounds used showed competitive results, compared to the reference ligand. With our parameters it also displayed better results than the Food and Drug Administration (FDA) approved drug hydroxychloroquine.<sup>45</sup> Compound 2606 binds preferentially to the C-terminal region of the spike protein S2 domain which is responsible for the formation of an helical coiled structure that participates in the virus-host cell membrane fusion.<sup>8</sup> The fusion of SARS-CoV-2 with host cell membrane relies on activation by surface proteases, such as TMPRSS2, which induces S1 dissociation and structural changes of the S2 domain.<sup>46</sup> Therefore, future work will consider if the compounds described here can interfere with the binding entry of SARS-CoV-2 into the host cell by blocking the interaction of the spike protein and TMPRSS2. Although the work in progress is computational and requires additional validation

*in vitro* and *in vivo*, our results demonstrate the potential of the curcumin analogues described here for the purpose of finding an anti-COVID-19 drug.

## Supplementary Information

Supplementary data with chemical structure of compounds and molecular dynamics simulations are available free of charge at <http://jbcns.sbq.org.br> as PDF file.

## Acknowledgments

This research was funded by FAPESP, grant numbers 19/03074-1, 18/50008-1, Maunders McNeil Foundation Inc. and CAPES. We acknowledge the Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas. The authors are grateful for the access to the Santos Dumont supercomputer (NVIDIA V100) at the National Laboratory of Scientific Computing (LNCC), Santos Dumont, Rio de Janeiro, Brazil.

## Author Contributions

Jéssica R. Nogueira, Flávia A. Verza, Felipe Nishimura and Icaro P. Caruso were responsible for the

conceptualization, data curation, investigation, software, writing original draft and writing-review and editing; Umashankar Das, Ana L. Fachin, Jonathan R. Dimmock and Mozart Marins for the project administration, resources, writing original draft and writing-review and editing.

## References

1. *Summary of Probable SARS Cases with Onset of Illness from 1 November 2002 to 31 July 2003*; <https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003>, accessed in June 2021.
2. <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf>, accessed in June 2021.
3. *Coronavírus em: Brasil-Rastreador no Mapa Dinâmico do Microsoft Bing*; <https://www.bing.com/covid/local/brazil>, accessed in June 2021.
4. Gorbalenya, A. E.; Baker, S. C.; Baric, R. S.; de Groot, R. J.; Drosten, C.; Gulyaeva, A. A.; Haagmans, B. L.; Lauber, C.; Leontovich, A. M.; Neuman, B. W.; Penzar, D.; Perlman, S.; Poon, L. L. M.; Samborskiy, D. V.; Sidorov, I. A.; Sola, I.; Ziebuhr, J.; *Nat. Microbiol.* **2020**, *5*, 536.
5. Castells, M.; Lopez-Tort, F.; Colina, R.; Cristina, J.; *J. Med. Virol.* **2020**, *92*, 2165.
6. Satarker, S.; Nampoothiri, M.; *Arch. Med. Res.* **2020**, *51*, 482.
7. Cui, J.; Li, F.; Shi, Z. L.; *Nat. Rev. Microbiol.* **2019**, *17*, 181.
8. Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Velesler, D.; *Cell* **2020**, *181*, 281.
9. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X.; *Nature* **2020**, *581*, 215.
10. Choudhry, H.; Bakhrebah, M. A.; Abdulaal, W. H.; Zamzami, M. A.; Baothman, O. A.; Hassan, M. A.; Zeyadi, M.; Helmi, N.; Alzahrani, F.; Ali, A.; Zakaria, M. K.; Kamal, M. A.; Warsi, M. K.; Ahmed, F.; Rasool, M.; Jamal, M. S.; *Future Virol.* **2019**, *14*, 237.
11. Hofmann, H.; Pyrc, K.; van der Hoek, L.; Geier, M.; Berkhout, B.; Pöhlmann, S.; *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 7988.
12. Xiao, X.; Chakraborti, S.; Dimitrov, A. S.; Gramatikoff, K.; Dimitrov, D. S.; *Biochem. Biophys. Res. Commun.* **2003**, *312*, 1159.
13. Wong, S. K.; Li, W.; Moore, M. J.; Choe, H.; Farzan, M.; *J. Biol. Chem.* **2004**, *279*, 3197.
14. Babcock, G. J.; Esshaki, D. J.; Thomas, W. D.; Ambrosino, D. M.; *J. Virol.* **2004**, *78*, 4552.
15. Li, W.; Zhang, C.; Sui, J.; Kuhn, J. H.; Moore, M. J.; Luo, S.; Wong, S. K.; Huang, I. C.; Xu, K.; Vasilieva, N.; Murakami, A.; He, Y.; Marasco, W. A.; Guan, Y.; Choe, H.; Farzan, M.; *EMBO J.* **2005**, *24*, 1634.
16. Song, H. D.; Tu, C. C.; Zhang, G. W.; Wang, S. Y.; Zheng, K.; Lei, L. C.; Chen, Q. X.; Gao, Y. W.; Zhou, H. Q.; Xiang, H.; Zheng, H. J.; Chern, S. W. W.; Cheng, F.; Pan, C. M.; Xuan, H.; Chen, S. J.; Luo, H. M.; Zhou, D. H.; Liu, Y. F.; He, J. F.; Qin, P. Z.; Li, L. H.; Ren, Y. Q.; Liang, W. J.; Yu, Y. D.; Anderson, L.; Wang, M.; Xu, R. H.; Wu, X. W.; Zheng, H. Y.; Chen, J. D.; Liang, G.; Gao, Y.; Liao, M.; Fang, L.; Jiang, L. Y.; Li, H.; Chen, F.; Di, B.; He, L. J.; Lin, J. Y.; Tong, S.; Kong, X.; Du, L.; Hao, P.; Tang, H.; Bernini, A.; Yu, X. J.; Spiga, O.; Guo, Z. M.; Pan, H. Y.; He, W. Z.; Manuguerra, J. C.; Fontanet, A.; Danchin, A.; Niccolai, N.; Li, Y. X.; Wu, C. I.; Zhao, G. P.; *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2430.
17. *Press Updates|COVID-19*; <https://covid19vaccintrial.co.uk/press-updates>, accessed in June 2021.
18. van Doremalen, N.; Lambe, T.; Spencer, A.; Belij-Rammerstorfer, S.; Purushotham, J. N.; Port, J. R.; Avanzato, V. A.; Bushmaker, T.; Flaxman, A.; Ulaszewska, M.; Feldmann, F.; Allen, E. R.; Sharpe, H.; Schulz, J.; Holbrook, M.; Okumura, A.; Meade-White, K.; Pérez-Pérez, L.; Edwards, N. J.; Wright, D.; Bissett, C.; Gilbride, C.; Williamson, B. N.; Rosenke, R.; Long, D.; Ishwarbhai, A.; Kailath, R.; Rose, L.; Morris, S.; Powers, C.; Lovaglio, J.; Hanley, P. W.; Scott, D.; Saturday, G.; de Wit, E.; Gilbert, S. C.; Munster, V. J.; *Nature* **2020**, *586*, 578.
19. Tse, L. V.; Meganck, R. M.; Graham, R. L.; Baric, R. S.; *Front. Microbiol.* **2020**, *11*, 658.
20. Elfiky, A. A.; *J. Biomol. Struct. Dyn.* **2020**, *39*, 3204.
21. Rahman, N.; Basharat, Z.; Yousuf, M.; Castaldo, G.; Rastrelli, L.; Khan, H.; *Molecules* **2020**, *25*, 2271.
22. Das, S.; De, A.; Saha, R.; Sharma, N.; Khemka, M.; Singh, S.; Hesanoor Reja, A.; Kumar, P.; *Indian J. Dermatol.* **2020**, *65*, 118.
23. Hewlings, S.; Kalman, D.; *Foods* **2017**, *6*, 92.
24. Hassanzadeh, S.; Read, M. I.; Bland, A. R.; Majeed, M.; Jamialahmadi, T.; Sahebkar, A.; *Pharmacol. Res.* **2020**, *159*, 104921.
25. Li, G.; Fan, Y.; Lai, Y.; Han, T.; Li, Z.; Zhou, P.; Pan, P.; Wang, W.; Hu, D.; Liu, X.; Zhang, Q.; Wu, J.; *J. Med. Virol.* **2020**, *92*, 424.
26. Zahedipour, F.; Hosseini, S. A.; Sathyapalan, T.; Majeed, M.; Jamialahmadi, T.; Al-Rasadi, K.; Banach, M.; Sahebkar, A.; *Phytother. Res.* **2020**, *34*, 2911.
27. Patel, A.; Rajendran, M.; Shah, A.; Patel, H.; Pakala, S. B.; Karyala, P.; *J. Biomol. Struct. Dyn.* **2021**, DOI: 10.1080/07391102.2020.1868338.
28. Ferreira, L. G.; dos Santos, R. N.; Oliva, G.; Andricopulo, A. D.; *Molecules* **2015**, *20*, 13384.
29. Salmaso, V.; Moro, S.; *Front. Pharmacol.* **2018**, *9*, 923.
30. Rosemond, M. J. C.; St. John-Williams, L.; Yamaguchi, T.; Fujishita, T.; Walsh, J. S.; *Chem.-Biol. Interact.* **2004**, *147*, 129.
31. Das, S.; Das, U.; Michel, D.; Gorecki, D. K. J.; Dimmock, J. R.; *Eur. J. Med. Chem.* **2013**, *64*, 321.

32. Trott, O.; Olson, A. J.; *J. Comput. Chem.* **2009**, *31*, 455.
33. Goodsell, D. S.; Olson, A. J.; *Proteins: Struct., Funct., Bioinf.* **1990**, *8*, 195.
34. Sanner, M. F.; *J. Mol. Graphics Modell.* **1999**, *17*, 57.
35. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; *Gaussian 09, Revision E.01*; Gaussian, Inc., Wallingford CT, 2009.
36. Stewart, J. J. P.; *J. Mol. Model.* **2007**, *13*, 1173.
37. Olsson, M. H. M.; SØndergaard, C. R.; Rostkowski, M.; Jensen, J. H.; *J. Chem. Theory Comput.* **2011**, *7*, 525.
38. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A.; *Nucleic Acids Res.* **2004**, *32*, 665.
39. Laskowski, R. A.; Swindells, M. B.; *J. Chem. Inf. Model.* **2011**, *51*, 2778.
40. Daina, A.; Michielin, O.; Zoete, V.; *Sci. Rep.* **2017**, *7*, 42717.
41. Martin, Y. C.; *J. Med. Chem.* **2005**, *48*, 3164.
42. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E.; *Bioinformatics* **2013**, *29*, 845.
43. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.; *J. Chem. Phys.* **1983**, *79*, 926.
44. Delano, W. L.; *The PyMOL Molecular Graphics System*; Delano Scientific, San Carlos, 2002.
45. Chary, M. A.; Barbuto, A. F.; Izadmehr, S.; Hayes, B. D.; Burns, M. M.; *J. Med. Toxicol.* **2020**, *16*, 284.
46. Hofmann, H.; Pöhlmann, S.; *Trends Microbiol.* **2004**, *12*, 466.

Submitted: February 16, 2021

Published online: June 11, 2021

