

IDENTIFICATION OF NOVEL COUMARIN BASED COMPOUNDS AS POTENTIAL INHIBITORS OF THE 3-CHYMOTRYPSIN-LIKE MAIN PROTEASE OF SARS-COV-2 USING DFT, MOLECULAR DOCKING AND MOLECULAR DYNAMICS SIMULATION STUDIES

G. SALGADO. MORAN ^a, WILSON CARDONA V. ^b, LORENA. GERLI CANDIA ^c, L.H. MENDOZA-HUIZAR ^d AND TOOBA ABDIZADEH ^{e*}

^aFacultad de Ciencias Químicas. Investigador Extramural, Universidad de Concepción, Concepción, Chile.

^bDepartamento de Ciencias Químicas, Facultad de Ciencias Exactas, Universidad Andrés Bello, Concepción, Chile.

^cDepartamento de Química Ambiental, Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Chile.

^dAutonomous University of Hidalgo State. Academic Area of Chemistry. Mineral de la Reforma, Hidalgo. México.

^eClinical Biochemistry Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran.

ABSTRACT

SARS-CoV-2 is the pandemic disease-causing agent COVID-19 with high infection rates. Despite the progress made in vaccine development, there is an urgent need for the identification of antiviral compounds that can tackle better the different phases of SARS-CoV-2. The main protease (Mpro or 3CLpro) of SARS-CoV-2 has a crucial role in viral replication and transcription. In this study, an in silico method was executed to elucidate the inhibitory potential of the synthesized 6-tert-octyl and 6-8-ditert-butyl coumarin compounds against the major protease of SARS-CoV-2 by comprehensive molecular docking and density functional theory (DFT), ADMET properties and molecular dynamics simulation approaches. Both compounds shown favorable interactions with the 3CLpro of the virus. From DFT calculations, HOMO-LUMO values and global descriptors indicated promising results for these compounds. Furthermore, molecular dynamics studies revealed that these ligand-receptor complexes remain stable during simulations and both compounds showed considerably high binding affinity to the main SARS-CoV-2 protease. The results of the study suggest that the coumarin compounds 6-tert-octyl and 6-8-ditert-butyl could be considered as promising scaffolds for the development of potential COVID-19 inhibitors after further studies.

Keywords: COVID-19, SARS-CoV-2, Coumarin, Molecular docking, Molecular dynamics, Density functional theory.

1. INTRODUCTION

COVID-19 is the contagious disease that is caused by the SARS-CoV-2 coronavirus. Since its appearance in Wuhan, China, in 2019, the infections and spread disease have been increasing steadily. COVID 19 is now a worldwide pandemic affecting more than 190 countries on all continents of the world, and the number of cases has been increasing daily and the number of deaths has risen. Several strategies have been implemented to prevent the spread of SARS-CoV-2, such as physical or social distancing, quarantine, indoor ventilation, and hand washing. The use of facemasks in public settings has also been recommended to minimize the risk of transmissions. But, the number of victims continues to rise and the symptoms caused by this virus have highlighted the urgent need for drugs to attack this disease [1]. However, to the best of our knowledge, there are no specific remedies for COVID-19 and research into its treatment is scarce [2]. In addition, we currently have no specific drug available for the treatment of patients with COVID-2019. In this sense, many governments, medical institutions and scientists have experimented with various treatments used for other diseases with very promising but so far inconclusive results. These treatments include chloroquine, hydroxychloroquine, camostat, nafamostat, umifenovir, tenofovir, remdesidir, sofosbuvir, galidesivir, lopinavir/ritonavir and indinavir, which are used to treat other diseases but which have shown some degree of SARS-CoV-2 inhibitory activity [3-5]. Thus, active and safe antiviral agents with broad-spectrum activity against this emerging and potentially fatal infection are being investigated [6], and numerous phytochemicals from therapeutic plants have been testified for antiviral activity [7-9]. Thus, it is well accepted that natural phytochemicals are a source that can provide an important and powerful resource of chemical compounds with antiviral properties [10, 11].

Also, it is important to mention that Methyltransferase (MTase), Endoribonuclease(endoU), ADP ribose Phosphatase and main protease enzymes are essential for the viability of SARS CoV-2 [12, 13], therefore, several attempts have been made to identify inhibitors of these enzymes. In this sense, some coumarin derivatives show antiviral activity and have shown promising results against MTase, endoU, ADP ribose phosphatase and major protein enzyme receptors. In addition, coumarins are easy to synthesize and derive, have high biological activity and low toxicity. In tandem, it is important to note that since coronavirus is known to cause pulmonary embolism, binding affinities against vitamin K epoxide reductase complex inhibitor have also been investigated, taking into account the anticoagulant property, which is an important property of coumarins.

However, coumarins have been scarcely investigated against SARS-CoV-2 [14, 15]. Also, it has been reported in the literature that some coumarin derivatives, both of natural and chemical origin have good antiviral activity. For examples, Maurya et al, performed in silico studies of two thousand seven hundred and fifty-five biologically active coumarin derivatives downloaded from the PubChem website, and they found five coumarin derivatives that are able to interact with MTase, endoU, ADP ribose phosphatase and major protein protease [12]. Molecular docking studies, which are the first step in drug design, have also been performed on the target coumarins, and the high-scoring results have allowed conducting in vitro and in vivo studies [14-18]. In this study, we have investigated two novel synthesized coumarin compounds to explore and identify the binding affinities and the interactions of these compounds against the coronavirus 3CLpro by molecular modeling approaches namely molecular docking, dynamics and DFT analysis. Also, the docking score for coumarin compounds was compared with antiviral drugs (Ritonavir and Lopinavir), which was found to have a significant outcome in the treatment of COVID-19.

2. MATERIAL AND METHODS

2.1 Protein structure

It has been reported that the 3CLpro cleavage sites on the polyproteins of coronaviruses are highly conserved, considering that SARS-CoV-2 belongs to the same family as of SARS-CoV and MERS-CoV. They share several similarities and their sequence and substrate specifications for coronaviruses of SARS-CoV-2, SARS-CoV, and MERS-CoV should be identical [19]. This similarity allows us to compare SARS-CoV-2 with its previous counterparts leading to the identification of compounds able to inhibit or control the replication of SARS-CoV-2. Therefore, the crystal structures of coronaviruses 3CLpro which were retrieved from the protein data bank (PDB) (<http://www.rcsb.org>) with the corresponding PDB identification codes [SARS-CoV-2 (6LU7), SARS-CoV (2DUC) and MERS-CoV (2YNA)]. 6LU7, 2DUC and 2YNA (PDB ID) were chosen as 3CLpro receptors because these have resolution values of 2.16 Å, 1.70 Å, and 1.50 Å, respectively. During the preparation process of the proteins using the Molecular Operation Environment (MOE) software [20], their water molecules and original ligands were removed, while polar hydrogen's and Gasteiger charges were added to each protein. The protein structures were minimized by the energy minimization algorithm of MOE using the MMFF94X force field with the conjugate gradient method. Then, the protein structures were saved for molecular docking studies.

2.2 Ligand preparation

6-tert-octyl and 6-8-ditert-butyl coumarin structures, two bioactive coumarins that are recently reported by Zárrega O. group [21] and the 3CLpro reference inhibitors (ritonavir and lopinavir) were drawn and converted into PDB format using Chemoffice Bio 3D ultra (version 12.0, Combridge Soft Corporation, Cambridge, UK, 2010) and were optimized at the AM1 semi-empirical level as implemented in Hyperchem package.

2.3 Density functional theory (DFT)

All the calculations were performed at the framework of the Density functional theory and carried with the program Gaussian 09 [22]. The electronic and structural properties of 6-tert-octyl and 6-8-ditert-butyl coumarin compounds were evaluated at the B3LYP/6-311++G(d-p) level of theory.

In order to analyze the ligands interaction in the binding site of SARS-CoV-2 main protease, we calculated molecular parameters such as: the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO), E_{HOMO} and E_{LUMO} , energy values, the energy gap (ΔE_{gap}), ionization potential (I), electron affinity (A), electronegativity (χ), chemical hardness (η), chemical potential (μ), chemical softness (S), electrophilicity index (ω), dipole moment (D) and molecular electrostatic potential (MEP) maps.

2.4 Molecular docking study

The molecular docking calculations of the 6-tert-octyl and 6-8-ditert-butyl coumarin molecules, reference inhibitors (ritonavir and lopinavir) and N3 inhibitor with the active site of coronaviruses 3CLpro were carried out using MOE software to analyze their interaction mode.

In the MOE (Inc 2016), the receptor–ligand binding affinities with all possible binding geometries are prioritized based on a numerical value called S-score. It may identify salt bridges, hydrogen bonds, hydrophobic interactions, sulfur-LP, cation- π , and solvent exposure. Thus, in this work, the interactions between inhibitors and receptor proteins are predicted based on this S-score. During the analysis the proteins were considered as a rigid structure, while the ligands were fully flexibility. By use of site-finder module implanted in MOE, the active site of the 3CLpro proteins was determined and the active site was defined with at least one atom within a distance cut off of 4.5 Å at ligand in the crystal structure of 3CLpro. The molecular docking was done using the triangle matcher placement algorithm in combination with the London dG scoring function to assign the theoretical free binding energies of the protein-ligand complexes and force field as the refinement method. The best conformation of the ligands was further evaluated by the binding energies (S-score, kcal/mol) and interactions between the ligands and proteins were analyzed by the LigX module in MOE and UCSF chimera software.

2.5 Validation of docking

Docking protocol was validated by re-docking of the co-crystallized ligand (N3) into the 3CLpro structure (6LU7). It can be observed in Figure 1, that the N3 molecule was bonded to similar positions of 3CLpro compared to its original crystallographic form and the docked structure had an RMSD of 1.669 Å after superimposing it on the co-crystallized native complex, indicating the validity of the method used. Here, the amino acids that are showed in the vicinity of 6.5 Å from the N3 inhibitor are considered as main bindings residues. Aldo, note that residues such as Leu27, His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166 and Gln189 are located within the catalytic pocket of the main protease.

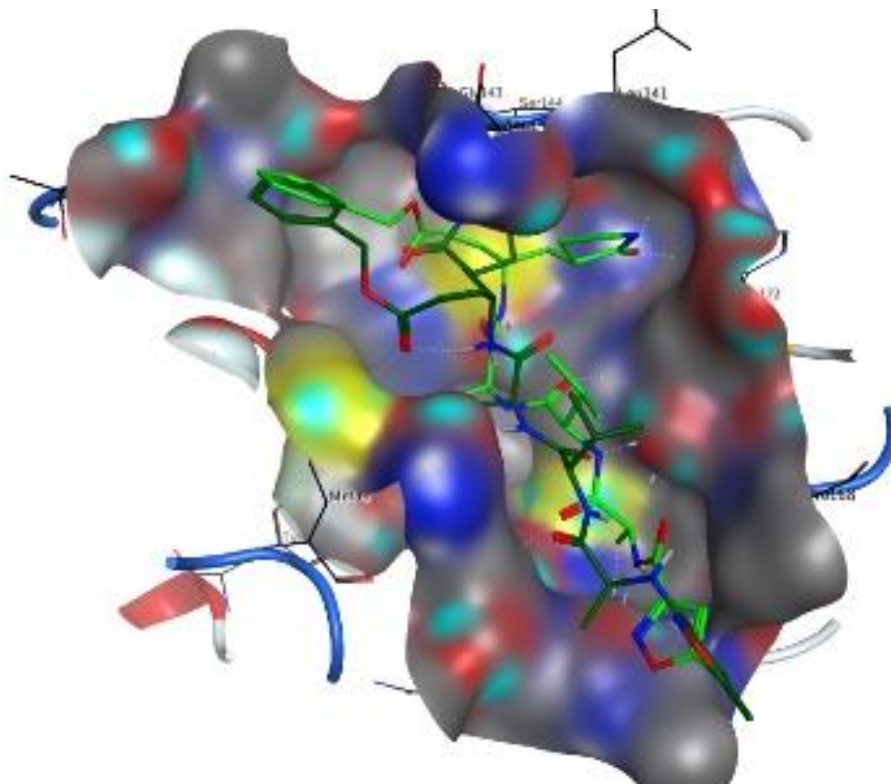


Figure 1. Re-docking validation for N3. Crystallographic (green) and re-docked (dark green) N3 inhibitor of the 6LU7 SARS-CoV-2 3CLpro inside its binding pocket.

2.6 In silico evaluation of physicochemical and pharmacokinetics properties

Using the Swiss and pkCSM-pharmacokinetics web tools, the pharmacokinetic properties of 6-tert-octyl and 6-8-ditert-butyl coumarins that exhibited significant binding affinity for 3CLpro of SARA-CoV-2 were evaluated. The analysis was performed by evaluating the pharmacokinetics and physicochemical characteristics using the drug likeness rules (Lipinski [23],

Veber [24], Egan [25], Ghose and Muegge [26], lipophilicity (Log Po/w), water solubility, Log S, polar surface area (TPSA), number of rotational bonds and medicinal chemistry methods (PAINS, Brenk, Lead likeness, synthetic accessibility). The canonical SMILES of the coumarin compounds were retrieved from Chem Draw to calculate ADMET and drug likeness properties using default parameters. Also, the PASS prediction and the molecular target studies were calculated using PASS-Way2Drug server and Swiss target prediction [27-29].

2.7 Molecular dynamics simulation of SARS-CoV-2 3CLpro

MD simulation was also used to analyze the dynamic behavior of the higher bonding complexes predicted by the molecular docking study [30, 31]. MD simulations were performed considering the SARS-CoV-2 free 3CLpro; 3CLpro complexed with the co-crystal inhibitor N3 and lopinavir as a control; and the complexes of 3CLpro with the 6-tert-octyl and 6-8-ditert-butyl coumarins obtained from molecular docking studies. This one allowed us to obtain more information about the protein and the docked complexes under biological conditions. The MD simulations were carried out using the GROningenMaChine for chemical simulations (GROMACS, version 2019.1) with amber99sb force field [32-34] for 50 ns at the real physiological condition and aqueous solution at T=310 K (37 °C) and P=1bar. The topology files of 6-tert-octyl and 6-8-ditert-butyl coumarins, N3, and lopinavir were obtained from the PRODRG server [35] while the topology file of 3CLpro protein was prepared by the GROMACS. The 3CLpro and 3CLpro-ligand systems were solvated and fully immersed in the cubic box with tip3p water model and neutralized by adding sodium or chloride ions with salt concentration of 0.15 M and energy minimized using the steepest descent algorithm [36, 37] for 50000 steps and energy tolerance of 500 kJ/mol in a periodic boundary condition and equilibrated to achieve the appropriate volume under NVT ensemble [38, 39]. The particle-Mesh Ewald (PME) method was used with a Fourier grid spacing of 1 Å to calculate the long-range electrostatic interactions. The short-range Lennard-Jones and Coulomb interactions were calculated by a cutoff value of 14 Å and the final MD production was carried out with a time step of 2 fs for 50 ns in the NPT ensemble [39, 40].

The simulation system comprised four phases. The first phase involved the energy minimization of the entire system employing the steepest descent integrator in preparation with the subsequent conjugate gradient algorithm integrator. The second phase involved the minimization and molecular dynamics of the NVT and NPT ensembles for 1000 ps and 5000 ps respectively, allowing the solvents and ions to evolve. The third phase involved heating of the systems, having a lower temperature coupling ($\tau=0.1$ ps) along with pressure coupling ($\tau=0.5$ ps) to achieve equilibrium at 310 K and 1 atm of temperature and pressure. In the equilibration phase, the thermostat and barostat were estimated through the Berendsen algorithm [41]. The hydrogen-containing bond lengths were repressed with the help of the LINCS algorithm [42]. The last and fourth phase, termed as the production step is where the MD simulations for 500 ns at 310 K temperature having 2 fs of time step were completed, and the final structures were attained. The Maxwell Boltzman distribution was employed in order to reassign the velocities at each step. Nose Hoover thermostat and Parrinello Rahman barostat were the respective thermostat and barostat for the final MD simulation [41]. The parameters of root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), solvent accessible surface area (SASA), secondary structure, and total intermolecular H bonds were ultimately calculated by MD simulations. Principle component analysis (PCA) was carried out using g-cover on the backbone atoms of the proteins.

2.8 The binding free energy calculations

Molecular Mechanic/Poisson-Boltzmann surface Area (MM-PBSA) method [43] was used to estimate the binding free energy of the ligand-protein complex in the explicit solvent [44] using the g-mmpbsa script program [45]. The MD simulation trajectory of 50 ns was considered for the calculation of different components of the binding energy of the ligand and 3CLpro complex. The free energy of binding was calculated using the following equation:

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand}$$

Here, ΔG_{bind} represent total binding free energy, while others show the free energy of ligand-protein complex, protein, and ligand, respectively.

$$G = E_{MM} - T\Delta S + \Delta G_{solv}$$

$$E_{MM} = E_{bonded} + E_{vdw} + E_{ele}$$

$$\Delta G_{solv} = G_{polar} + G_{non\ polar}$$

the E_{bonded} , E_{vdw} and E_{ele} represent interactions among bonded, van der Waals, and electrostatic states. In contrast, the polar and nonpolar interactions to the solvation free energy are presented by the G_{polar} and $G_{non\ polar}$, respectively and $-T\Delta S$ is the entropy contribution at temperature T.

To further confirm the key residues involved in the binding modes the binding free energy decomposition was calculated after the calculation of every term of binding energy contribution. The amino acid residues whose binding energy contribution is less than -2 KJ/mol and greater than 2 KJ/mol are selected to reveal their favorable and unfavorable contributions in the simulation process.

3. Results and discussion

The main aim of this study was to identify potential 6-tert-octyl and 6-8-ditert-butyl coumarins as inhibitors for the 3CLpro protein of SARS-CoV-2. 3CLpro was selected because of its important role in viral replication. The molecular docking of 6-tert-octyl and 6-8-ditert-butyl coumarins with 3CLpro protein was performed and these compounds that showed a strong binding affinity for 3CLpro were selected for further investigations. The ADMET properties of the 6-tert-octyl and 6-8-ditert-butyl coumarins were analyzed and then, these compounds were evaluated through MD simulations and calculated free energy of binding for the compounds using MM-PBSA. This *in silico* study was undertaken to identify potential antiviral compounds for COVID-19 (Figure 2).

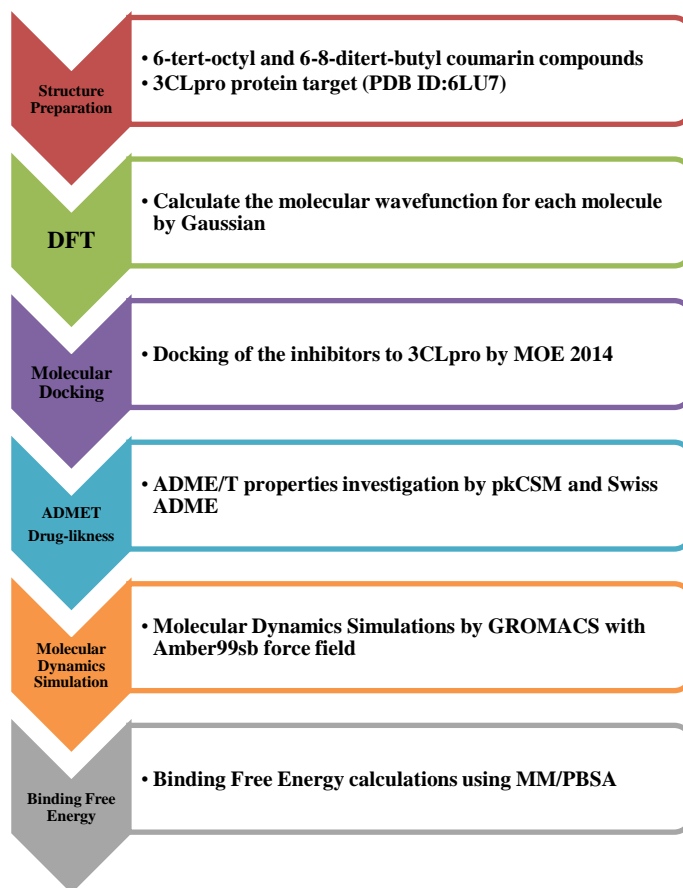


Figure 2. Schematic view of the computational approach used to identify 6-tert-octyl and 6-8-ditert-butyl coumarins as inhibitor for the 3CLpro protein.

3.1 Density functional theory

To determine the influence of molecular structure on the antiviral activity of the coumarin compounds, some of their corresponding theoretical properties were determined using B3LYP-6.311++G(d,p) method and their influence on COVID-19 main protease. The Frontier molecular orbitals (FMOs) of 6-tert-octyl and 6-8-ditert-butyl coumarin compounds specify an important role of charge-transfer interactions with the active site of SARS-CoV-2 main protease. HOMO orbitals as nucleophilic or electron donating and LUMO orbitals as electrophilic or electron accepting in 6-tert-octyl and 6-8-ditert-butyl coumarin compounds are spread on coumarin and 3-carboxylate moieties that interacts with Phe140, Gly143, Ser144, Cys145 and His41 (Figure 3). Also, HOMO orbitals are located on the 6-(2,4,4-trimethylpentan-2-yl) group of 6-tert-octyl coumarin and the 6-tert-butyl group of 6-8-ditert-butyl coumarin which interacts with Leu141, Met49 and Met165.

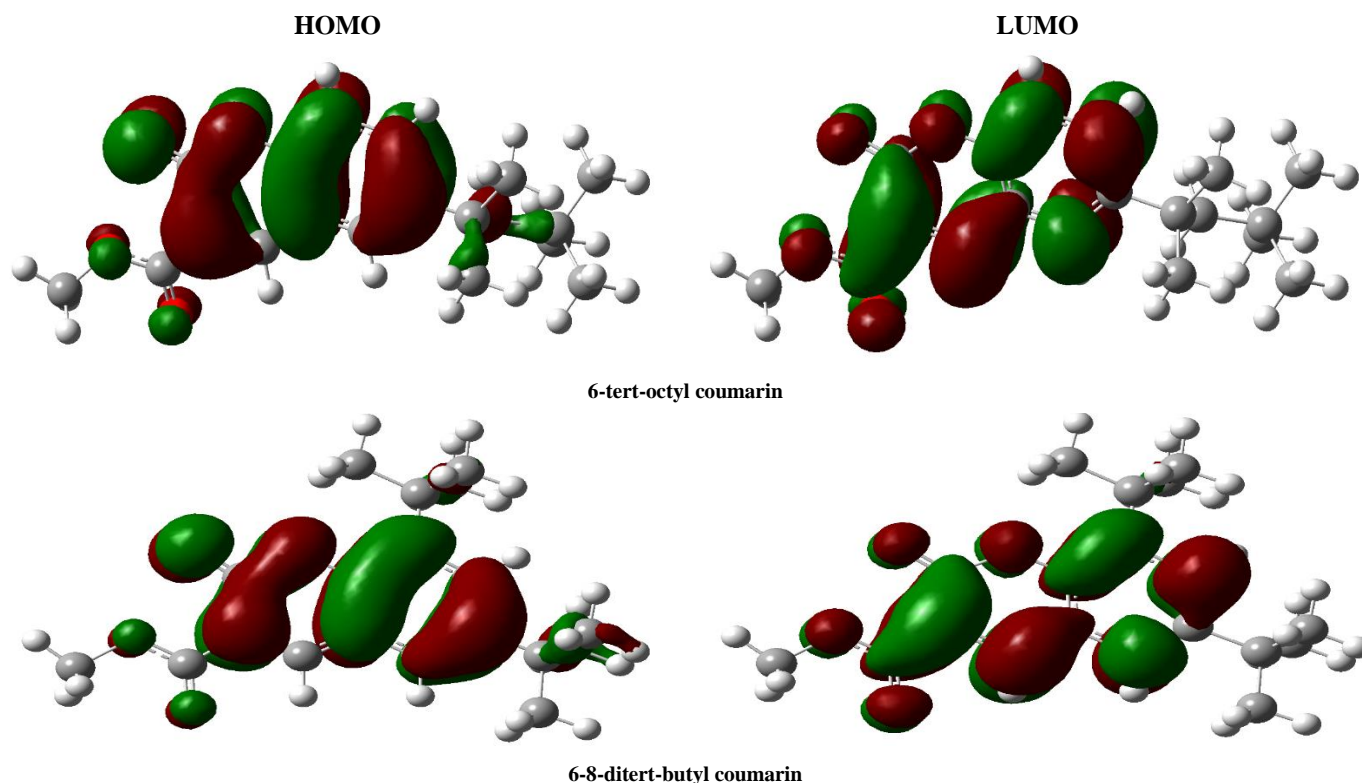


Figure 3. HOMO and LUMO plots of the 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin.

The calculated E_{HOMO} and E_{LUMO} of the coumarin compounds clearly explained the global reactivity descriptors. The negative values for E_{HOMO} and E_{LUMO} for 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds indicate their stability (Table 1). The energy gap values of 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds showed high affinity of these compounds for COVID-19 main protease. To get some conclusive evidence in determining the structure properties of these compounds, a set of chemical reactivity parameters such as electronegativity (χ), chemical potential (μ), global hardness (η), global softness (S) and global electrophilicity index (ω) were calculated (Table 1).

The chemical softness (S) values of 0.243 and 0.241 eV and the electrophilicity (ω) values of 5.20 and 5.38 eV explain the coumarin compounds stability and which correlated with the trend of molecular docking. The negative values of the chemical potential (μ) imply good stability, and the formation of a stable complex with the receptor [46]. Also, the high dipole moment 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds could show their binding pose within COVID-19 main protease and is an illustration of the molecular docking results.

Table 1. Frontier molecular orbital energies (ev) and global reactivity descriptors.

Property ^a	6-tert-octyl coumarin	6-8-ditert-butyl coumarin
E_{HOMO}	-6.69	-6.79
E_{LUMO}	-2.57	-2.64
ΔE_{Gap}	4.12	4.15
I (ev)	6.69	6.79
A (ev)	2.57	2.64
χ (ev)	4.63	4.72
μ (ev)	-4.63	-4.72
η (ev)	2.06	2.07
S (ev ⁻¹)	0.243	0.241
ω (ev)	5.20	5.38
Dipole moment (D)	5.87	7.50

$$^a|\Delta E| = E_{\text{HOMO}} - E_{\text{LUMO}}, I = -E_{\text{HOMO}}, A = -E_{\text{LUMO}}, \chi = (I + A)/2, \eta = (I - A)/2, \mu = -(I + A)/2, S = 1/(2\eta), \text{ and } \omega = \mu^2/2\eta.$$

The molecular electrostatic potential (MEP) is an important way to validate the reactivity of the drug as inhibitor. Although the MEP gives an indication about the molecular size and shape of the positive, negative as well as the neutral electrostatic potential. These could be a tool to predict physicochemical property relationships with the molecular structure of the drugs under investigation. Moreover, the molecular electrostatic potential is a useful tool to estimate the reactivity of the drugs toward electrophilic and nucleophilic attacks. The highest negative and positive potentials were characterized by red and blue colors [47], while other colors indicate their respective intermediate potentials on the total electron density surface. The MEP map of 6-tert-octyl coumarin range from -0.111e0 to 0.111e0 a.u. The highest negative (red) regions were spread over O11

atom of coumarin moiety and O13 and O22 atoms of the 3-methyl carboxylate moiety that interact with Ser144, Gly143 and Cys145. While the highest positive (blue) regions are located on aromatic ring of the coumarin moiety and 2,4,4-trimethylpentan-2-yl group of the compound that interact with His41 and Met165 (Figure 4a). The MEP map of 6-8-ditert-butyl coumarin range from -0.118e0 to 0.118e0 a.u. For 6-8-ditert-butyl coumarin, the red regions are located on O11 of coumarin moiety and O13 and O22 of the 3-methyl carboxylate that interact with Ser144 and Gly143. While the blue regions are found on aromatic ring of the coumarin moiety and tert-butyl groups of the 6-8-ditert-butyl coumarin that interact with His41 and Met49 (Figure 4b).

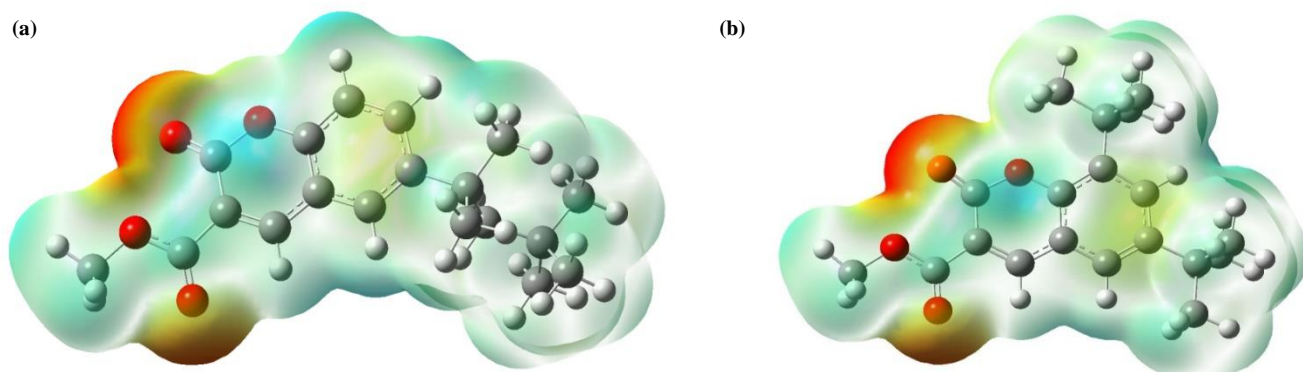


Figure 4. MEP map for (a) 6-tert-octyl coumarin and (b) 6-8-ditert-butyl coumarin.

The Mulliken's atomic charges (Figure 5) of the DFT calculation revealed a charge distribution in individual atoms. The results indicate that the oxygen atom has the most negative charge, which is due to the molecular relaxation. As well, the hydrogen atoms cover the positive charges and the charges on carbon atoms

exhibited either positive or negative values. Generally, the heteroatoms (O) can share their pairs of electrons with acceptor molecules. Based on the results, the coumarin compounds have good bioactivity and can be considered as potential COVID-19 main protease inhibitors.

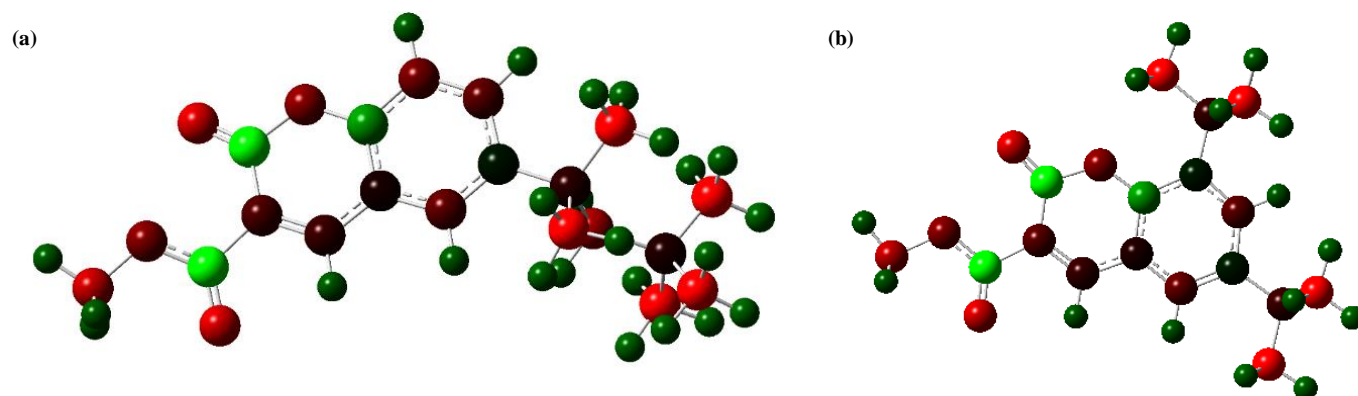


Figure 5. Mulliken atomic charge representation of (a) 6-tert-octyl coumarin and (b) 6-8-ditert-butyl coumarin at DFT/B3LYP/6-311++G(d-p) level.

3.2 Molecular docking

The molecular docking approach to identify potential hits has become one of the most popular methods for structure-based computer-aided drug discovery (SB-CADD). In this study, we investigate all the possible binding modes and mechanism of action of 6-tert-octyl and 6-8-ditert-butyl coumarins 3CLpro

proteins. The binding energy scores of the coumarin compounds to 3CLpro of SARS-CoV-2, along with their binding energy scores to 3CLpro of SARS-CoV and MERS-CoV were shown in Table 2. The results of the binding affinities from the docking analysis showed that 6-tert-octyl and 6-8-ditert-butyl coumarins can bind to 3CLpro of SARS-CoV-2 with the highest affinity of -11.54 and -11.41 kcal/mol, respectively.

Table 2. Binding affinities of co-crystal N3, reference inhibitors (Ritonavir and Lopinavir) and 6-tert-octyl and 6-8-ditert-butyl coumarins.

No	Bioactive Compounds	SARS-CoV-2	SARS-CoV	MERS-CoV
-	Ritonavir	-6.84	-8.10	-8.12
-	Lopinavir	-10.89	-9.97	-9.37
-	N3	-10.93	-9.59	-9.30
1	6-tert-octyl coumarin	-11.54	-10.18	-10.22
2	6-8-ditert-butyl coumarin	-11.41	-9.92	-10.74

As shown by Table 2, N3, lopinavir, and ritonavir that are considered as the co-crystal inhibitor of 3CLpro and reference inhibitors exhibited a docking score of -10.93 kcal/mol, -6.84 kcal/mol, and -10.89 kcal/mol for 3CLpro of SARS-CoV-2 whereas docking scores of N3, ritonavir, and lopinavir for 3CLpro of SARS-CoV equaled -9.59 kcal/mol, -8.10 kcal/mol and -9.97 kcal/mol and for 3CLpro of MERS-CoV equaled -9.30 kcal/mol, -8.17 kcal/mol and -9.37 kcal/mol. 6-tert-octyl and 6-8-ditert-butyl coumarins showed docking score against SARS-coronavirus, which surpassed that of N3 and reference inhibitors. The 6-tert-octyl and 6-8-ditert-butyl coumarins compounds were found to exhibit

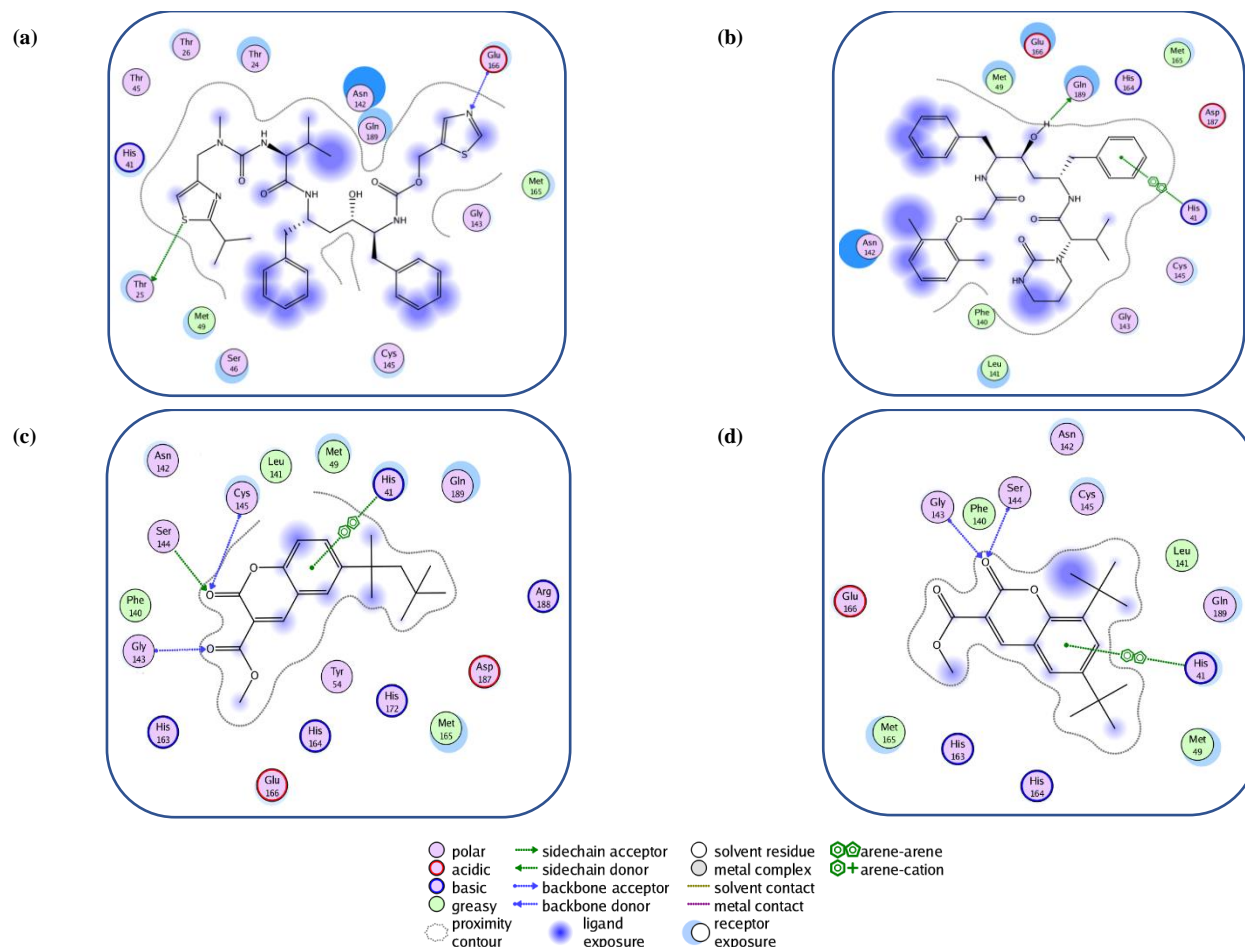
the best docking scores of -11.54 and -11.41, respectively, compared to N3, ritonavir and lopinavir inhibitors against 3CLpro of SARS-CoV-2, interestingly, 6-tert-octyl and 6-8-ditert-butyl coumarins had the highest binding affinity to that of SARS-CoV and MERS-CoV (Table 2). Analysis of the interactions of the 6-tert-octyl and 6-8-ditert-butyl coumarins and reference inhibitors with amino acid residues of 3CLpro of coronaviruses (Table 3) showed that these compounds majorly interacted with the hotspot residues through hydrophobic interactions and with hydrogen bonding (particularly with Cys145 and His41).

Table 3. Interacting amino acid (aa) residues of 3CLpro of coronaviruses with the 6-tert-octyl and 6-8-ditert-butyl coumarins.

Bioactive compound	Coronavirus	Interacted residues	aa residue involved in H-bonding (Bond Distance)
Ritonavir	SARS-CoV-2	His 41, Cys145, Gly143, Met165, His164, Glu166, Asn142, Met49, Gln189, Thr26, Thr24, Thr25, Thr45, Ser46	Glu166 (2.48), Thr25 (3.72)
Lopinavir		His 41, Cys145, Gly143, Met165, His164, Glu166, Asn142, Leu141, Phe140, Met49, Gln189, Asp187	Gln189 (2.08)
6-tert-octyl coumarin		His 41, Cys145, Ser144, Gly143, Met165, His163, His164, Glu166, Asn142, Leu141, Phe140, Met49, Gln189, Asp187, Arg188, His172, Tyr54	Gly143 (3.098), Ser144 (2.888), Cys145 (3.585)
6-8-ditert-butyl coumarin		His 41, Cys145, Gly143, Ser144, Met165, His163, His164, Glu166, Asn142, Leu141, Phe140, Met49, Gln189	Gly143 (2.721), Ser144 (2.916)
6-tert-octyl coumarin	SARS-CoV	Gln189, Met165, His164, Cys145, His41, Arg188, Asp48, Cys44, Thr25, Glu47, Thr24, Thr45, Ala46	Ser144 (2.48), Ser144 (2.78), Cys145 (2.42)
6-8-ditert-butyl coumarin		Gly143, Asp48, Cys145, His41, Cys44, Glu47, Ala46, Thr25, Thr26, Met49, His164, Asn142, Thr45	Thr25 (2.79), Glu47(2.65)
6-tert-octyl coumarin	MERS-CoV	Ser7, Ala8, Met6, Asp294, Gln299, Asp295, Met298, Asn156, Thr154, Glu155, Ser114, Ser116, Thr130	Thr130 (2.64), Thr130 (2.86), Ser116 (2.77)
6-8-ditert-butyl coumarin		Met6, Ala8, Gln299, Asp295, Met298, Phe291, Thr154, Glu155, Thr130, Ser114, Ala8, Ser116,	Gln299 (2.51), Gln299 (2.64)

The results of the molecular docking of the 6-tert-octyl and 6-8-ditert-butyl coumarins and reference inhibitors in the active site of SARS-CoV-2 3CLpro illustrated by their corresponding 2D interaction plots that the 6-tert-octyl and 6-8-ditert-butyl coumarins compounds interacted with either both (Cys145 and His41) or at least one catalytic dyad residue, detected by MOE (Figure 6) (Said et al. 2021; Ghosh et al. 2021). The 6-tert-octyl and 6-8-ditert-butyl coumarins compounds and ritonavir and lopinavir exhibit similar binding modes due to the parallel orientations of the ligands and their same key residues, such as His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, Met165, His164, Glu166, Gln189, and Thr190. The results of ligand-protein binding interaction showed that ritonavir and lopinavir as reference inhibitors were docked into the active site and catalytic dyad (Cys145 and His41) of SARS-CoV-2. Ritonavir could form two hydrogen bonds with the side chain of Thr25 and the backbone

of Glu166 (Figure 6a), while lopinavir with a considerably higher binding energy (-10.89 kcal/mol) than ritonavir showed significant π - π stacking interaction with His41 of the catalytic dyad and form one hydrogen bond with the side chain of Gln189 (Table 3, Figure 6b) and also, both of the inhibitors had hydrophobic interactions with surrendering residues. 6-tert-octyl coumarin had significant binding to the catalytic dyad of SARS-Cov-2 that interacted with His41 by π - π stacking interaction and form hydrogen bonds with the backbone of Cys145 and Gly143 and also side chain of Ser144 along with hydrophobic interactions with the other residues such as Met49, Met165, Leu141 and Phe140 (Table 3, Figure 6c). 6-8-ditert-butyl coumarin was stabilized through hydrogen bonding with the backbone of Gly143 and Ser144 and π - π stacking interaction with the catalytic residue of His41, and hydrophobic interactions with the other surrounding residues (Figure 6d).

**Figure 6.** 2D view of the binding conformation of ligands in binding pocket of SARS-CoV-2 3CLpro. (a) Ritonavir, (b) Lopinavir, (c) 6-tert-octyl coumarin, (d) 6-8-ditert-butyl coumarin.

Like the ligand-protein binding interaction of 6-tert-octyl coumarin to SARS-CoV-2 3CLpro that targeted the Cys-His catalytic dyad (Cys145 and His41) along with the other binding residues, the docking analysis showed that the SARS-CoV 3CLpro interacted with the same ligand differently. 6-tert-octyl coumarin interacted with Cys145 in catalytic dyad and Ser144 by hydrogen bonding interactions (Figure 7a) and for 6-8-ditert-butyl coumarin, hydrogen bonding interactions with Thr25 and Glu47 were observed (Figure 7b).

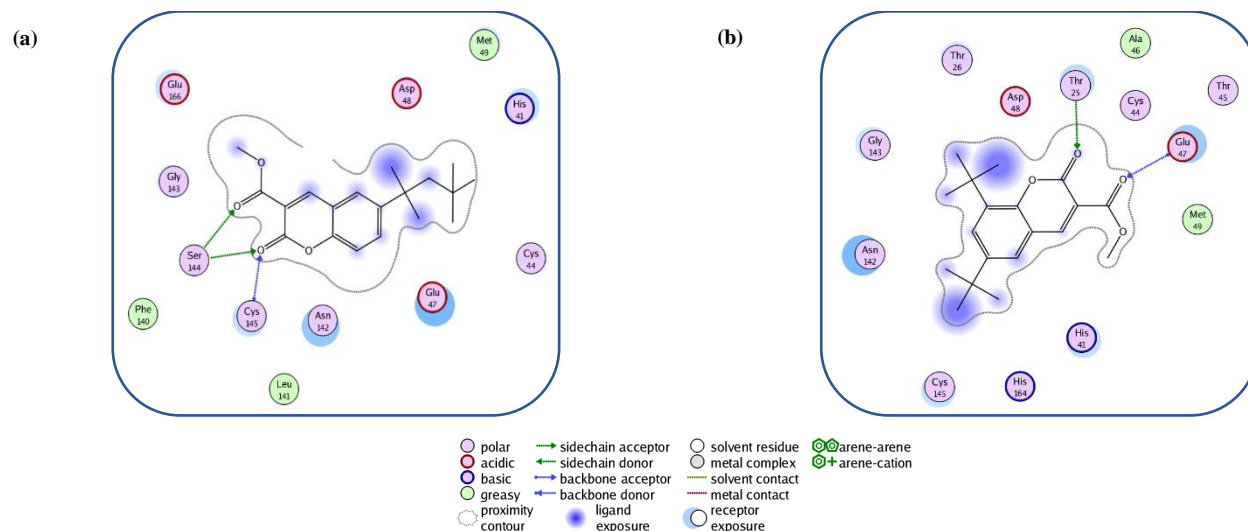


Figure 7. 2D representation of 3CLpro amino acid interactions of SARS-CoV with (a) 6-tert-octyl coumarin, (b) 6-8-ditert-butyl coumarin.

6-tert-octyl coumarin and 6-8-ditert-butyl coumarins interacted with MERS-CoV 3CLpro that is different from the other two coronaviruses. Hydrogen bonds were observed between 6-tert-octyl coumarin and Ser116 and Thr130 and 6-8-ditert-butyl coumarins interacted via hydrogen bonding interaction to Gln299. In addition to this, the other residues (Ala8, Asp295, Met6, Phe291, Asp294, and Met298) are involved in forming hydrophobic interactions with these compounds (Table 1 and Figure 8a and b).

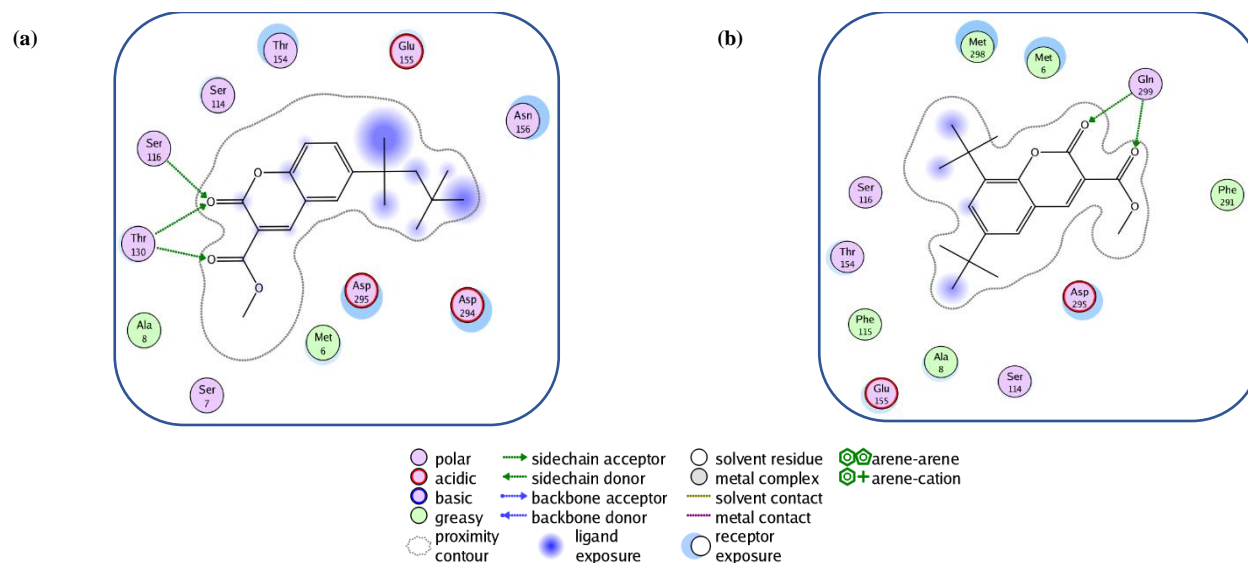


Figure 8. 2D representation of 3CLpro amino acid interactions of MERS-CoV with (a) 6-tert-octyl coumarin, (b) 6-8-ditert-butyl coumarin.

The results of this study showed that 6-tert-octyl and 6-8-ditert-butyl coumarin compounds with a considerable inhibitory tendency towards the SARS-coronavirus were docked into the active site and were interacting with the catalytic dyad (Cys-His) of 3CLpro protein of the coronaviruses in a similar pattern as ritonavir and lopinavir.

3.3 ADMET properties and drug likeness predictions

The pharmacokinetic and drug-likeness properties play a significant role towards the search and design of potential small drug-leads for a specific target [29]. All the ADMET properties of the 6-tert-octyl and 6-8-ditert-butyl coumarins are represented in Table 4 and 5.

According to the studies [48, 49], water solubility has been considered to be crucial to approximate the absorption of the medicines in the body, which for 6-tert-octyl and 6-8-ditert-butyl coumarins has been given in range -4.99 to -5.12. The intestinal absorption of 6-tert-octyl and 6-8-ditert-butyl coumarin compounds (95.63 and 95.61%) revealed an acceptable absorption feature.

Furthermore, the blood/brain partition coefficient (log BB) of the 6-tert-octyl and 6-8-ditert-butyl coumarin compounds indicated a lower opportunity for crossing the blood-brain barrier (BBB). Results have also shown that steady-state volume of distribution (VD_{ss}, log L/kg) value of 6-tert-octyl and 6-8-ditert-butyl coumarin showed more distribution in the tissues rather than in the plasma (Table 5). For metabolism, two compounds the plasma were predicted as the substrate for the CYP450 3A4 subtype, while these compounds not metabolized by CYP2D6. At the same time, the selected compounds could not inhibit the CYP450 2D6 and 3A4 subtypes; however, 6-8-ditert-butyl coumarin compound might inhibit CYP450 2C9 subtypes and 6-tert-octyl and 6-8-ditert-butyl coumarin compounds, could inhibit CYP450 1A2 and 2C19 subtypes. Based on

the prediction of the total clearance, hepatic and renal tissue can be used to clear such 6-tert-octyl and 6-8-ditert-butyl coumarin. The expected toxicity represents the fact that these compounds did not show any skin sensitization but has been detrimental to the liver. Moreover, Ames test has been used to reveal the anticipated toxicity, reflecting that 6-tert-octyl coumarin has been not mutagenic. 6-tert-octyl and 6-8-ditert-butyl coumarins obeyed the Lipinski's Rule of Five (Ro5) with a molecular weight < 500 g/mol, logP value < 5, hydrogen bond acceptors < 10, hydrogen bond donors < 5 and molecular refractivity < 140 (Table 4). The TPSA of the 6-tert-octyl and 6-8-ditert-butyl coumarins was less than 110 Å², indicating the potentiality of these compounds as favourable drug molecules [50]. The number of rotatable bonds for 6-tert-octyl and 6-8-ditert-butyl coumarins was ≤ 5, suggesting that these compounds are flexible in nature. Additionally, these compounds were moderately soluble and highly absorbable in the gastrointestinal tract. The synthetic accessibility score of the 6-tert-octyl and 6-8-ditert-butyl coumarins equaled 3.37 and 5.12, respectively. Its scale

demonstrated a score of 1 for a relatively simple synthetic route, whereas a score closer to 10 had a high structural complexity, and was therefore difficult to synthesize. Therefore, the synthetic accessibility value of the compounds was ≤ 6, which indicated their feasibility of synthesis. With regard to Table 2, drug-likeness filters such as, Veber, Ghose, Muegge, & Egan filters were in the reasonable ranges for the 6-tert-octyl and 6-8-ditert-butyl coumarins in solubility and lipophilicity.

Further analysis, a more illustrated and comprehensive study was done using bioavailability radar. Bioavailability radar is descriptive tool to investigate the drug-likeness of the ligands based on six physicochemical properties. The pink area shown in Table 4 represents most favorable area for each property like INSATU (unsaturation), INSOLU (insolubility), FLEX (rotatable bonds), LIPO (lipophilicity), SIZE (molecular weight) and POLAR (polar surface area). According to each parameter, it is possible to employ 6-tert-octyl and 6-8-ditert-butyl coumarins as the antiviral agents to treat COVID-19.

Table 4. Drug likeness properties of the top binding 6-tert-octyl and 6-8-ditert-butyl coumarins.

Drug Likeness Properties	6-tert-octyl coumarin	6-8-ditert-butyl coumarin
Molecular weight (g/mol) ^a	316.39	316.39
Log <i>P</i> _{ow} (ILOGP) ^b	3.30	3.40
Log S (ESOL) ^c	-5.18	-5.21
Num. H-bond acceptor ^d	4	4
Num. H-bond donor ^e	0	0
No of Rotatable Bonds ^f	5	4
Molar Refractivity	92.00	92.30
Lipinski	Yes	Yes
Ghose	Yes	Yes
Veber	Yes	Yes
Egan	Yes	Yes
Muegge	No (1 violation: XLOGP3>5)	No (1 violation: XLOGP3>5)
Bioavailability score	0.55	0.55
TPSA (Å ²) ^g	56.51	56.51
Synthetic accessibility (SA)	3.37	5.12
Solubility (mol/l)	4.89e-07	6.15e06
PAINS ^h	0 alert	0 alert
Brenk ⁱ	1 alert	1 alert
Leadlikeness	No	No
GI absorption	High	High
Bioavailability (Radar Plot) LIPO: lipophilicity FLEX: Flexibility POLAR: Polarity INSATU: Insaturation INSOLU: Insolubility SIZE: Size		

The general recommended ranges are as follows:

^aMolecular weight, <500

^bPredicted octanol/water partition coefficient, -0.4 to +5.6

^cPredicted aqueous solubility, <-5.0

^dNumber of hydrogen bond acceptors, <10

^eNumber of hydrogen bond donor, <5

^fRotatable bonds, <10

^gPolar surface area, <140 Å²

^hPan-Assay Interference

ⁱStructural Alert

Table 5. *In Silico* ADME/T prediction of the top binding 6-tert-octyl and 6-8-ditert-butyl coumarin.

Compound	Absorption		Distribution		Metabolism						Excretion	Toxicity	
	Water solubility	Intestinal absorption (human)	Blood brain barrier Permeability	VDss (human)	CYP						Total Clearance	AMES toxicity	
					2D6	3A4	1A2	2C19	2C9	2D6	3A4		
					substrate		inhibitor						
	(logmol/L)	Numeric (% Absorbed)	(log BB)	(log (L/kg))	Categorical		(Yes/No)				Numeric (log ml/min/kg)	(Yes/No)	
6-tert-octyl coumarin	-5.15	95.634	0.024	0.449	No	Yes	Yes	Yes	No	No	No	0.8	No
6-8-ditert-butyl coumarin	-4.99	95.61	-0.05	0.627	No	Yes	Yes	Yes	Yes	No	No	0.761	Yes

3.4 PASS prediction for antiviral activity

Studies have also considered PASS as a popular tool employed in nearly every drug industry with regard to the analyses of the structure-activity relationship [51]. It gives the prediction score for biological activities on the ratio of probability to be active (Pa) and probability to be inactive (Pi). A higher Pa means the biological activity is having more probability for a compound. Moreover, researchers identified the biological activity spectra of 6-tert-octyl and 6-8-ditert-butyl coumarin chemicals which are deposited [52] in the PASS database. Table 6 presents the prediction results of ten biological activities for

the 6-tert-octyl and 6-8-ditert-butyl coumarins. Notably, the present research results indicated the major utility of the PASS plan to predict the biological activities of 6-tert-octyl and 6-8-ditert-butyl coumarins based on the respective coumarin structure that has been shown by an average prediction coefficient equal to 0.80 (Pa ranging from 0.209 to 0.885 when Pa>Pi) for these compounds. Besides, the results revealed a number of biological activities of the 6-tert-octyl and 6-8-ditert-butyl coumarin compounds amongst the ten experimented activities.

Table 6. The PASS prediction results of the biological activities of the 6-tert-octyl and 6-8-ditert-butyl coumarins.

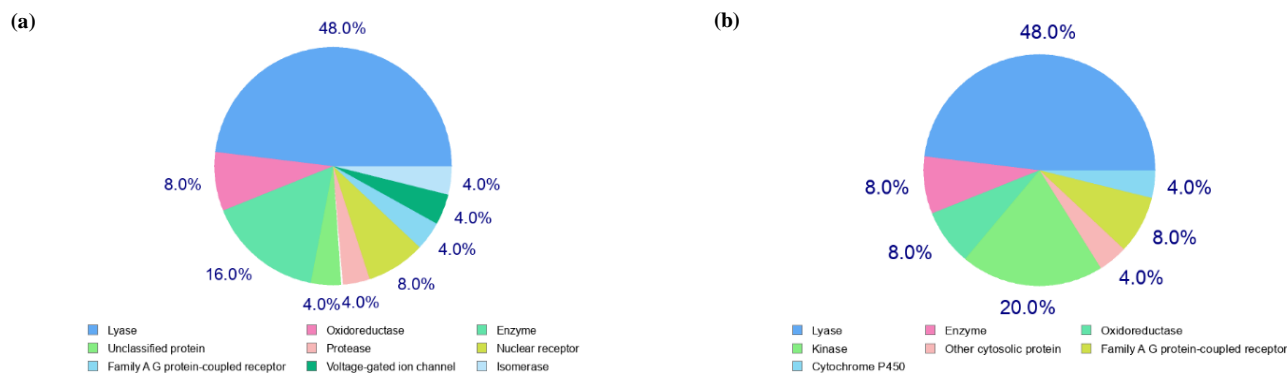
No	Biological activities	6-tert-octyl coumarin		6-8-ditert-butyl coumarin	
		Pa	Pi	Pa	Pi
1	CDP-glycerol glycerophosphotransferase inhibitor	0.393	0.198	0.655	0.066
2	Sugar-phosphatase inhibitor	0.480	0.091	0.530	0.074
3	Membrane integrity agonist	0.757	0.045	0.673	0.060
4	Aldehyde oxidase inhibitor	0.885	0.005	0.485	0.053
5	Antiinflammatory	0.588	0.034	0.846	0.005
6	Free radical scavenger	0.412	0.017	0.413	0.016
7	Chemopreventive	0.209	0.063	0.665	0.008
8	Membrane integrity agonist	0.757	0.045	0.673	0.060
9	Hepatoprotectant	0.379	0.036	0.448	0.025
10	Membrane permeability inhibitor	0.544	0.121	0.520	0.136

Pa: prediction of activity spectra for substances; Pi: probable inactivity.

3.5 Target prediction

Molecular target studies are important to find the phenotypic side effects or potential cross reactivity caused by the action of 6-tert-octyl and 6-8-ditert-butyl coumarins. Figure 9 showed the % bioactivity of these coumarins with respect to selected protein targets viz. enzymes, lyases, kinases, oxidoreductases, proteases, Family A G protein-coupled receptor, Cytochrome P450, and as a pie-chart. Analysis revealed that lyases, kinases, and enzymes were the main predicted targets for the proposed compounds. The pie chart of the 6-tert-octyl coumarin predicted 48% of Lyase, 16% of Enzyme, 4% of Unclassified protein,

4% of Protease, 8% of Oxidoreductase, 8% of Nuclear receptor, 4% of Isomerase, 4% of Family A G protein-coupled receptor and 4% of voltage-gated ion channel. Also, the pie chart of 6-8-ditert-butyl coumarin showed 48% of Lyase, 8% of Enzyme, 4% of other cytosolic protein, 20% of Kinase, 8% of Oxidoreductase, 8% of Cytochrome P450, and 8% of Family A G protein-coupled receptor. The possible sites of the target which the 6-tert-octyl and 6-8-ditert-butyl coumarins may bind were Lyases, Oxidoreductase and Family A G protein-coupled receptors which stimulate the drug reaction accordingly. Also, this analysis presents an explanation for the use of 6-tert-octyl and 6-8-ditert-butyl coumarins as SARS-CoV-2 main protease inhibitors.

**Figure 9.** Top-25 of target predicted for (a) 6-tert-octyl coumarin, (b) 6-8-ditert-butyl coumarin

3.6 Molecular dynamics simulation study

MD simulation is an imperative method to explore and confirm the structural rigidity and the contribution of key amino acid residues in proteins and validate the docking outcomes for the complexes. The MD simulations for 3CLpro-6-tert-octyl coumarin and 3CLpro-6-8-ditert-butyl coumarin complexes along with that of three other systems (ligand free 3CLpro, 3CLpro-N3, and 3CLpro-lopinavir) were done for 50 ns to analyze the stability of these docked compounds and evaluate the possible binding modes of the ligands. The stability of protein is the description of all the net forces to determine whether the protein will remain in folded state or assume non-native congregating structures. Therefore, the stability of protein is important to study the function of protein as alteration in protein stability would lead to misfolding or degradation of protein.

The RMSD analysis was performed on all understudy systems and their results are represented in Figure 10. In the case of free protein of 3CLpro after an initial jump due to the relaxation of the protein, the system reached equilibration after 10 ns and fluctuated around the mean RMSD value of 2.45 Å until the end of the simulation. This finding confirmed the sufficiency of simulation time, in addition to indicating that there is no significant change in protein structure during simulation. The RMSD value of the 3CLpro-N3 complex was stable throughout the simulation. However, a minor increase on random fluctuations of 2.87 Å was observed between 0 and 15 ns. Then, within the next 7 ns, the value was slightly

decreased (~ 2.51 Å) and the value increased progressively and reached ~ 2.70 Å (with some fluctuations) at 30 ns and remained almost the same till the end of the MD simulation. The RMSD value for 3CLpro-lopinavir was found to remain almost constant in a variation range of ~ 2.50 - 3.50 Å from 5 ns to 50 ns with some marginal fluctuations. For the 3CLpro-6-tert-octyl coumarin complex, the RMSD value from 3 ns to 19 ns maintained a constant value (~ 2.28 Å). Afterward, the value increased gradually and reached to ~ 3.71 Å at 22 ns. Then, the RMSD value slightly decreased and persisted at ~ 3.55 Å from 31 ns till the end of the MD simulation. Higher RMSD of 3CLpro-6-tert-octyl coumarin complex is mainly attributed to extended loop rearrangement (residues 185-200) and domain III (residues 200-306) fluctuations. For 3CLpro-6-8-ditert-butyl coumarin, the RMSD value from 2 ns to 8 ns maintained a constant value (~ 2.10 - 2.14 Å). Then, the RMSD value from 8 ns to 10 ns oscillated between ~ 2.53 - 3.64 Å. Thereafter a gradual decrease of RMSD value was observed upto 13 ns (~ 2.15 Å) and maintained equilibrium upto 50 ns of MD simulation. The average RMSD values for ligand free 3CLpro, 3CLpro-N3 and 3CLpro-lopinavir systems were found to be 2.45 Å, 2.73 Å, and 2.84 Å, respectively. Whereas, the average RMSD values of 3CLpro-6-tert-octyl coumarin, and 3CLpro-6-8-ditert-butyl coumarin complexes were 3.05 Å, and 2.00 Å, respectively. RMSD results showed that these two coumarin complexes were stable and the RMSD changes of 3CLpro with N3, lopinavir, and 6-tert-octyl coumarin were slightly more than ligand free 3CLpro that seemed that binding of the ligand with 3CLpro increased the conformational flexibility of 3CLpro.

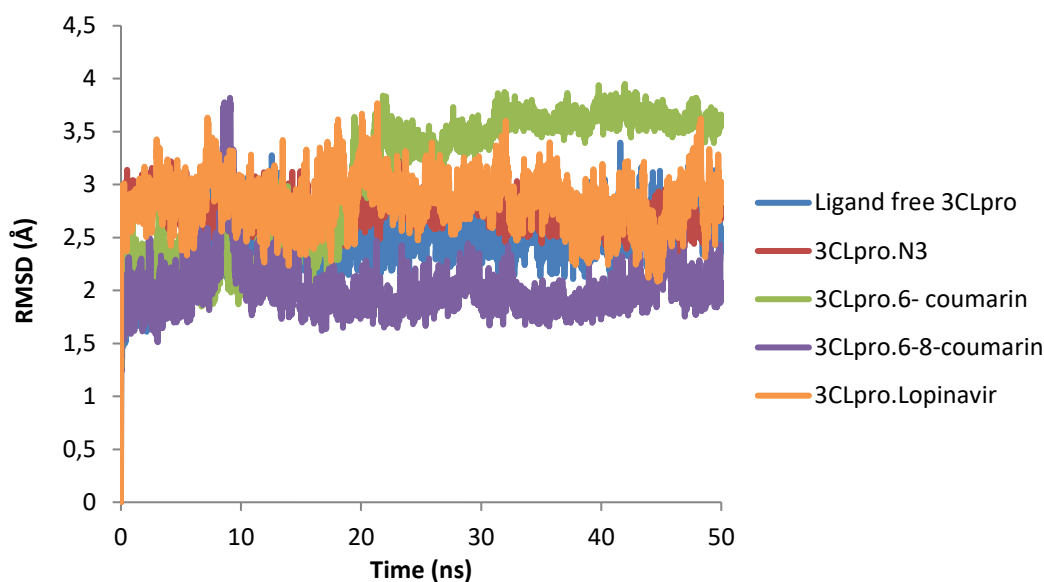


Figure 10. RMSD plots of ligand free 3CLpro, 3CLpro-N3, 3CLpro-6-tert-octyl coumarin, 3CLpro-6-8-ditert-butyl coumarin, and 3CLpro-lopinavir complexes of SARS-CoV-2.

The RMSF could be used to understand fluctuations as well as the flexibility of each residue in the different regions of simulated proteins. The average values of RMSF for ligand free 3CLpro, 3CLpro-N3, and 3CLpro-lopinavir were 1.36 Å, 1.80 Å, and 1.92 Å, respectively (Figure 11). In ligand free 3CLpro system, most of the amino acid residues within the domain I and II of this system had RMSF fluctuation below 3.0 Å and only, residues 47-52, 154, and 222 showed higher fluctuations (up to 3.0 Å). The RMSF plot of the 3CLpro-N3 complex showed that very few amino acid residues within domain I, II, and III (residues 155, 169, 195, 222, 236-238, 277-279, and 285) have an RMSF value of more than 3.0 Å. The RMSF plot of 3CLpro-lopinavir showed more or less similar conformational fluctuations that of the ligand free 3CLpro system. The fluctuations for many amino acid residues of domain I and II were reduced upon the binding of lopinavir to 3CLpro. Also, the RMSF plot in Figure 11 showed that 3CLpro-N3 and 3CLpro-lopinavir complexes suffer more conformational fluctuations in domain III. Analysis of RMSF plots showed that 3CLpro-6-tert-octyl coumarin and 3CLpro-6-8-ditert-butyl coumarin complexes had similar trends of dynamic fluctuation and RMSF distributions with average values of 1.56 Å, and 1.78 Å. These values indicated that 3CLpro-6-8-ditert-butyl coumarin and 3CLpro-6-8-ditert-butyl coumarin complexes showed lower conformational fluctuation as compared to ligand free 3CLpro and 3CLpro-N3/lopinavir complexes. The fluctuation of various specific amino acid residues

(Thr25, Thr26, His41, and Met49 in domain I; Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, and Glu166 of domain II and Gln189 and Thr190 of domain III) in 3CLpro-coumarin complexes was less than that in ligand free 3CLpro and 3CLpro-N3/lopinavir complexes indicating that these residues within the binding pocket of 3CLpro protein interacted with these coumarin compounds while the loop regions and the C-terminal and N-terminal of the protein were largely fluctuating in all systems. His41 and Cys145 residues of the catalytic dyad in SARS-CoV-2 3CLpro showed RMSF of 1.52 Å and 1.26 Å, respectively. The His41 residue showed RMSF of 1.48 Å and 1.50 Å for 3CLpro-N3/lopinavir complexes, respectively whereas in 3CLpro-6-tert-octyl coumarin and 3CLpro-6-8-ditert-butyl coumarin complexes His41 exhibited the RMSF of 1.03 Å, and 1.07 Å, respectively. The Cys145 residue exhibited the RMSF of 1.21 Å, 1.25 Å, 1.12 Å, and 1.13 Å for N3, lopinavir, 6-tert-octyl coumarin, and 6-8-ditert-butyl coumarin-protein systems, respectively indicating the stability of the target protein with smaller conformational changes and the lower fluctuations in binding residues of catalytic dyad for 3CLpro-coumarin complexes than that of the 3CLpro-N3/lopinavir complex. In fact, it was clearly evident from the RMSF plots that many key amino acid residues in the catalytic dyad/active site of 3CLpro (especially Cys145 and His41) were significantly reduced after binding to these two coumarins.

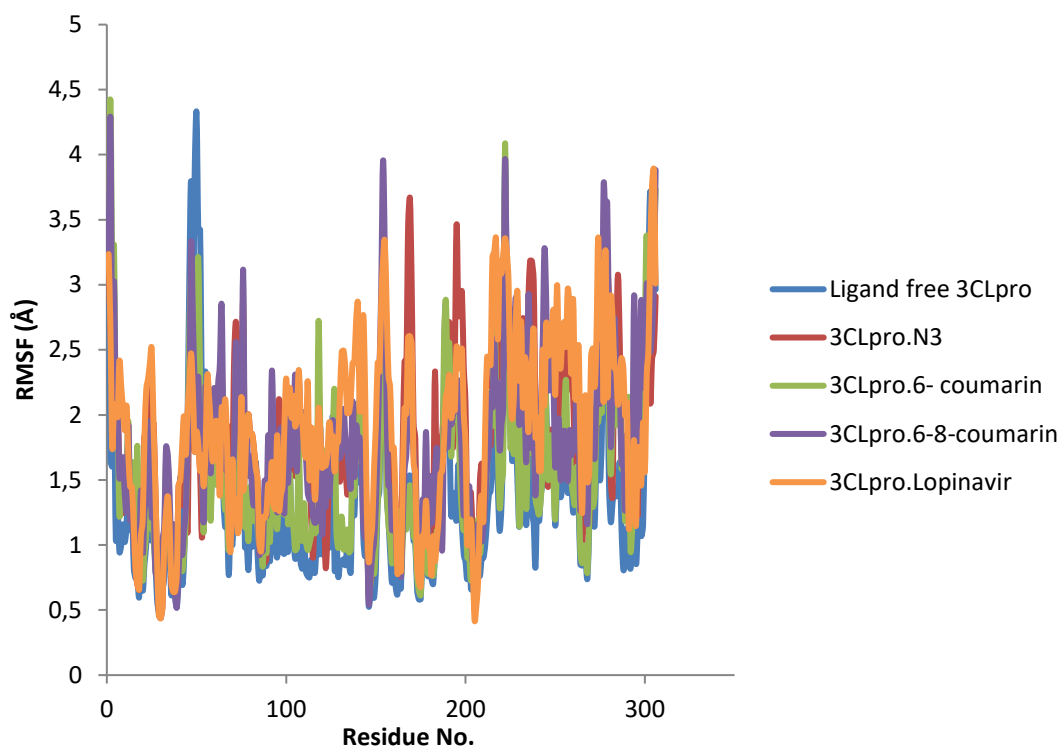


Figure 11. RMSF plot of ligand free 3CLpro and the 3CLpro-ligand complexes of SARS-CoV-2.

RMSD and RMSF analysis revealed that the SARS-CoV-2 3CLpro-6-tert-octyl coumarin, and 6-8-ditert-butyl coumarin docking complexes were highly stable during 50 ns simulations. Hydrogen bonding as a very specific bond plays a significant role in determining the stability of a ligand-receptor complex. Analysis of the main protease-ligand complexes revealed most of the compounds form Hydrogen bonds with the amino acid residues of the binding pocket (Figure 12). In the 3CLpro-N3 complex, the majority of conformations formed up to 3 hydrogen bonds during the MD simulation and a small number of conformations exhibited less than 1 and greater than 6 hydrogen bonds. For the 3CLpro-lopinavir complex, lopinavir formed 1 to 2 hydrogen bonds with residues of the

binding pocket. In 3CLpro-6-tert-octyl coumarin complex, the number of hydrogen bonds formed was between 2 to 5 in the whole simulation while 3CLpro- and 3CLpro-6-8-ditert-butyl coumarin complex showed changes in bonding. More hydrogen bonds (>5) were between 0 to 11 ns, after 11 ns the hydrogen bonds decreased to less than 5, and the last 13 ns, the hydrogen bond was between 2 to 3. This might suggest that there was a conformational change around 6-8-ditert-butyl coumarin in the binding site during simulation. In general, the results showed that these 3CLpro-coumarin complexes were highly stable.

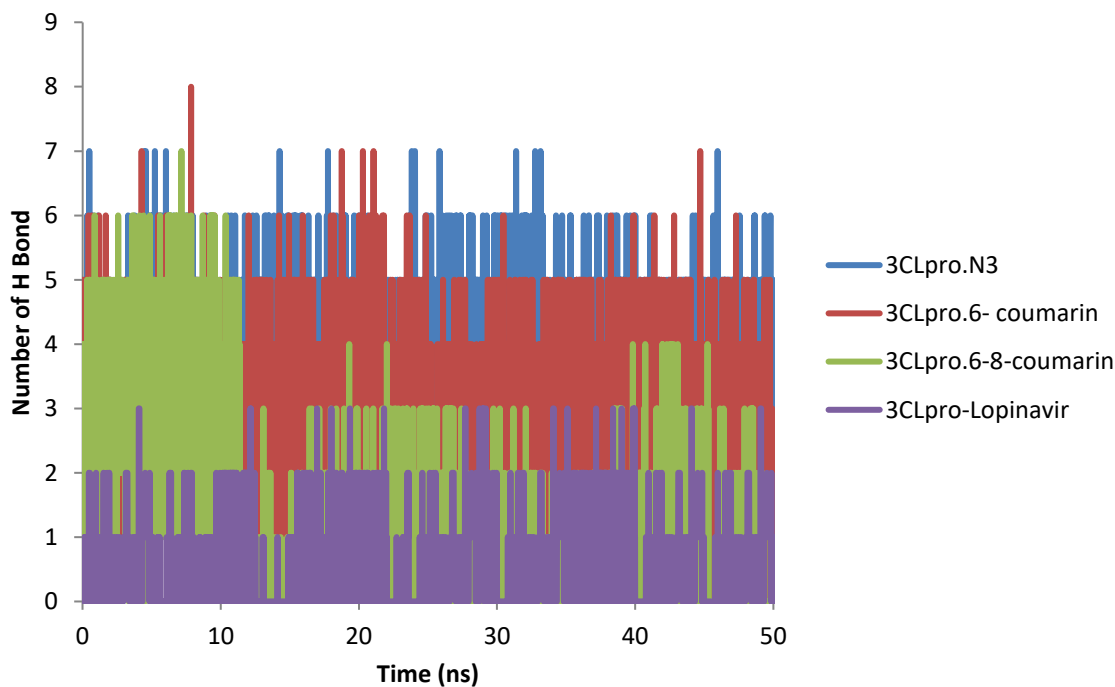


Figure 12. Number of intermolecular hydrogen bonds between 3CLpro of SARS-CoV-2 and N3, 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin and lopinavir.

The Rg parameter is often employed to describe the compactness and rigidity of the ligand-protein complex during MD simulations, in which less compactness (that is, being more unfolded) depicts higher Rg value with conformational entropy, while low Rg values explain strong compactness and higher structural stiffness (namely, being more folded). As shown in Figure 13, average Rg values of 3CLpro-N3 complex (22.16 Å) and 3CLpro-lopinavir complex (21.46 Å) were found to be in a similar range with ligand free 3CLpro (21.17 Å). The average Rg value for 3CLpro-6-tert-octyl coumarin (21.08 Å) and 3CLpro-6-8-ditert-

butyl coumarin (22.28 Å), systems was near or slightly lower than that of the other three systems (ligand free 3CLpro, 3CLpro-N3, and 3CLpro-lopinavir), which is a sign of compression in protein structure after binding to ligands. In an argument with the above observation, these molecules did not induce structural changes and were relatively rigid and both 3CLpro-coumarin complexes were compact throughout the simulation, indicating that the complexes were well converged.

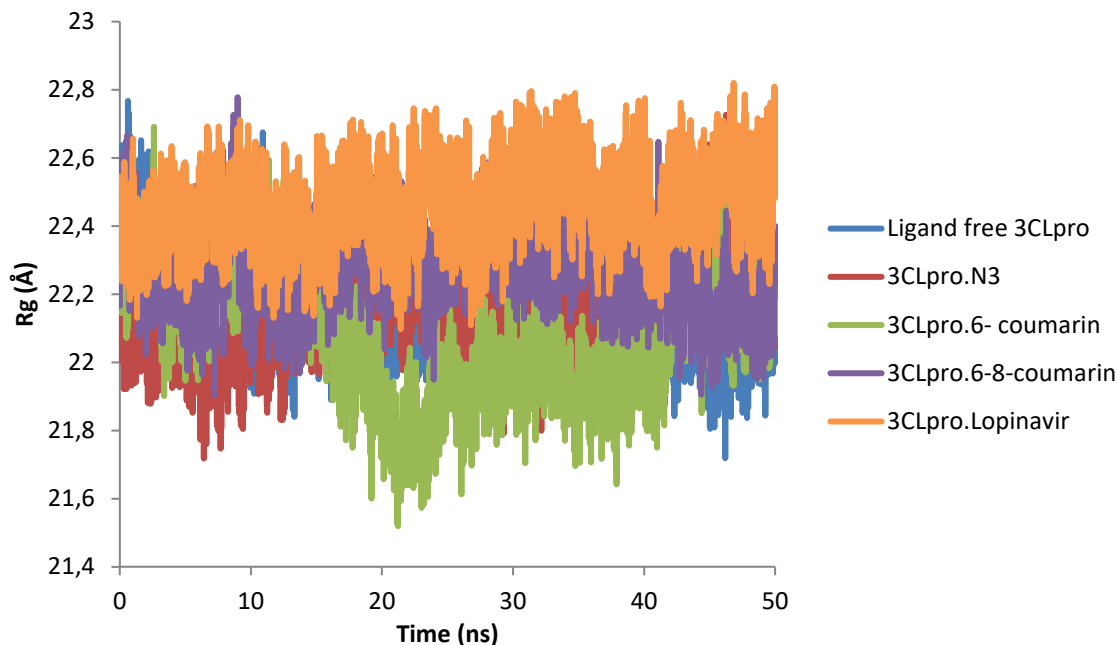


Figure 13. Radius of gyration (Rg) plot ligand free 3CLpro and the 3CLpro-ligand complexes of SARS-CoV-2.

Solvent Accessible Surface Area (SASA) value indicates the degree of expansion of protein volume in each system over the simulation time. The average SASA value of the 3CLpro-lopinavir complex ($\sim 14946.88 \text{ \AA}^2$) was higher than all other studied systems suggesting an expansion of 3CLpro during the interaction with lopinavir. The average SASA values of 3CLpro-6-tert-octyl coumarin, 3CLpro-6-8-ditert-butyl coumarin, and 3CLpro-N3 complexes were 14816.78 \AA^2 , 14944.73 \AA^2 , and 14840.58 \AA^2 , respectively. These values

indicated that the 6-tert-octyl coumarin complex was lower than that of the ligand free 3CLpro (14849.58 \AA^2) and 3CLpro-N3/lopinavir (Figure 14), suggesting that the binding of 6-tert-octyl coumarin potentially could reduce 3CLpro protein expansion. Also, these values indicated that 3CLpro-6-8-ditert-butyl coumarin complex experience slightly more expansion than that of ligand free 3CLpro and 3CLpro-N3 complex.

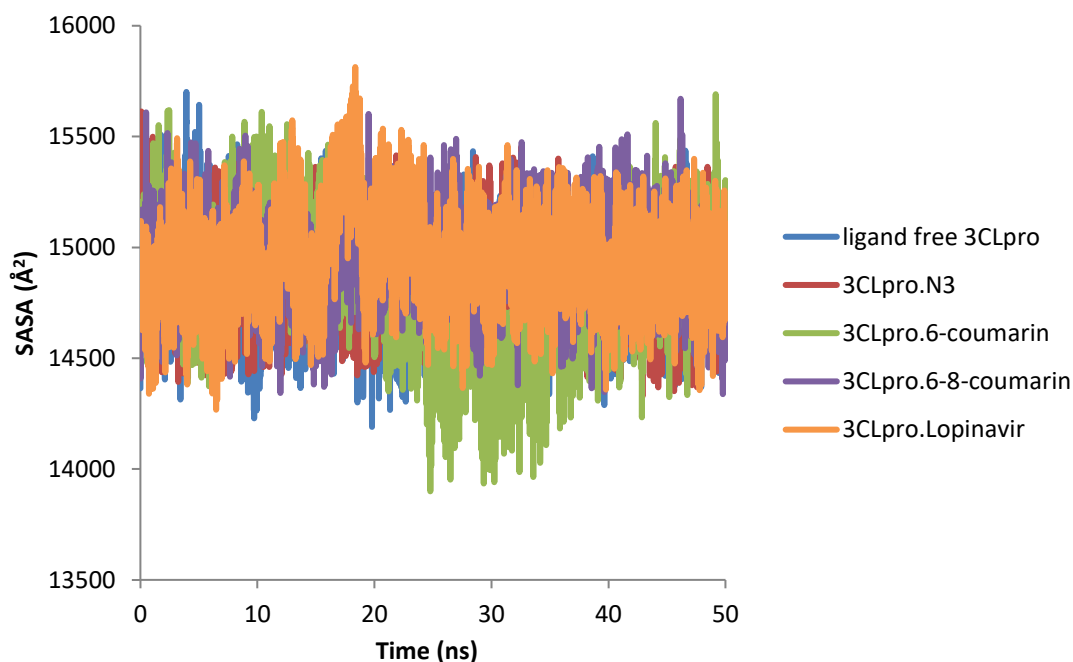


Figure 14. Solvent accessible surface area (SASA) plot ligand free 3CLpro and the 3CLpro-ligand complexes of SARS-CoV-2.

Principal component analysis (PCA) is used as a useful approach to extract the main component of protein movements. It plays an important role in identifying the configuration space of protein and function and any change in their patterns that contains a few degrees of freedom while the motion occurs. In this study, PCA is used to analyze the conformational structure of 3CLpro and the 3CLpro-ligand complexes via studying their collective motions through the MD simulation method. The 2D patterns of 3CLpro and 3CLpro-ligand complexes motions in distinct ligand-binding conditions were extracted by PCA and their results are demonstrated in Figure 15. The PCA plot in Figure 15a shows that

free 3CLpro protein contains almost two distinct movement clusters with a relatively medium range of motion from -8 to 6 nm. It has been found that 3CLpro-N3 and 3CLpro-lopinavir complexes have not altered both the pattern and these complexes occupied the same conformational subspace as 3CLpro in the free-state. The 3CLpro-6-tert-octyl coumarin and 3CLpro-6-8-ditert-butyl coumarin complexes have shown almost similar kind of correlated motions as compared to the free 3CLpro. Such data suggest that 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin have formed very stable complexes with 3CLpro (COVID-19 main protease) and can be considered as a lead compound.

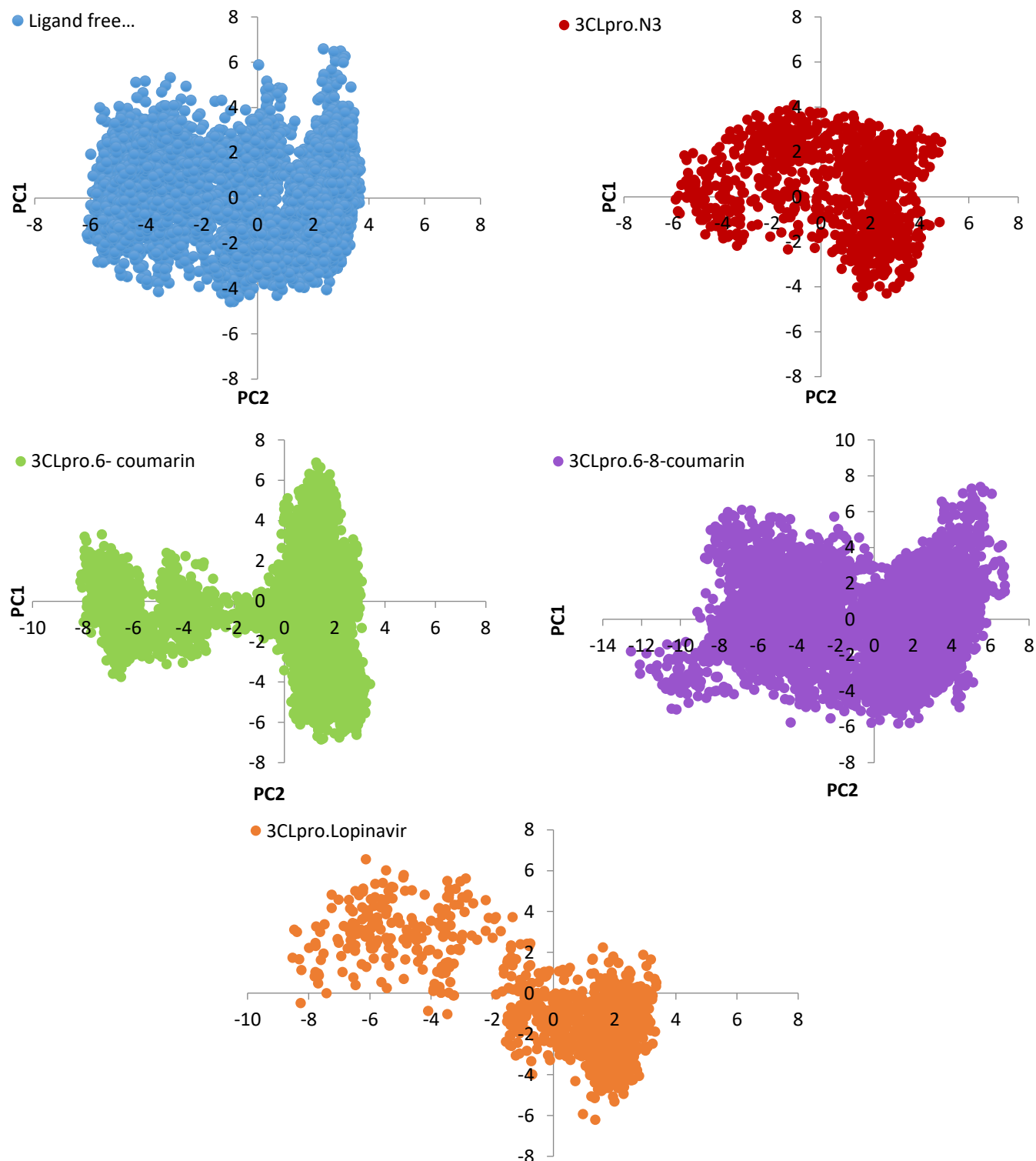


Figure 15. Comparison of changes in PCA pattern of protein in interaction with different ligands for function, free 3CLpro protein (blue), 3CLpro-N3 complex (red), 3CLpro-6-tert-octyl coumarin complex (green), 3CLpro-6-8-ditert-butyl coumarin (purple), 3CLpro-lopinavir (orange).

The define secondary structure of proteins (DSSP) is a standard method for determining secondary structure of proteins and investigating their alterations as a result of diverse conditions, such as the binding of a ligand. The conformational variations in the 3CLpro protein can be further elucidated throughout the MD simulation process by estimating the secondary structure content of the protein alone and in complexes with 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin, N3 and lopinavir. The content of coil, β -sheet, β -bridge, bend, turn, α -helix, and 3-helix in ligand free 3CLpro was 26%, 26%, 2%, 8%, 16%, 18% and 4%, respectively (Figure 16 and Supplementary Figure. S1). According to these results, the coil and β -sheet contents of the 3CLpro protein secondary structure are the most proportion. It was noticed that all these secondary structural components of 3CLpro undergo no significant alterations upon binding to 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin, N3 and lopinavir. In the presence of the N3 and lopinavir, the coil proportion of the 3CLpro protein was reduced to

25% compared with the free 3CLpro protein and for 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin, while, the coil and bend proportions of the 3CLpro protein were similar to free 3CLpro protein. In contrast, the turn and 3-helix structures were slightly decreased in the presence of the 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin, whereas the α -helix content increased to 19%, indicating that the 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin binding influenced the secondary structure of the 3CLpro protein. Hence, the 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin binding-induced increase in the α -Helix and decreased turn and 3-Helix structure of 3CLpro. Thus, the binding of coumarins to 3CLpro has no effect on the rigidity of 3CLpro structure and is consistent with the results of RMSD RG, RMSF, and H-bond studies. Altogether, these findings suggested that overall structural conformation including secondary conformation of 3CLpro is unaltered even when complexed with coumarins (6-tert-octyl coumarin and 6-8-ditert-butyl coumarin).

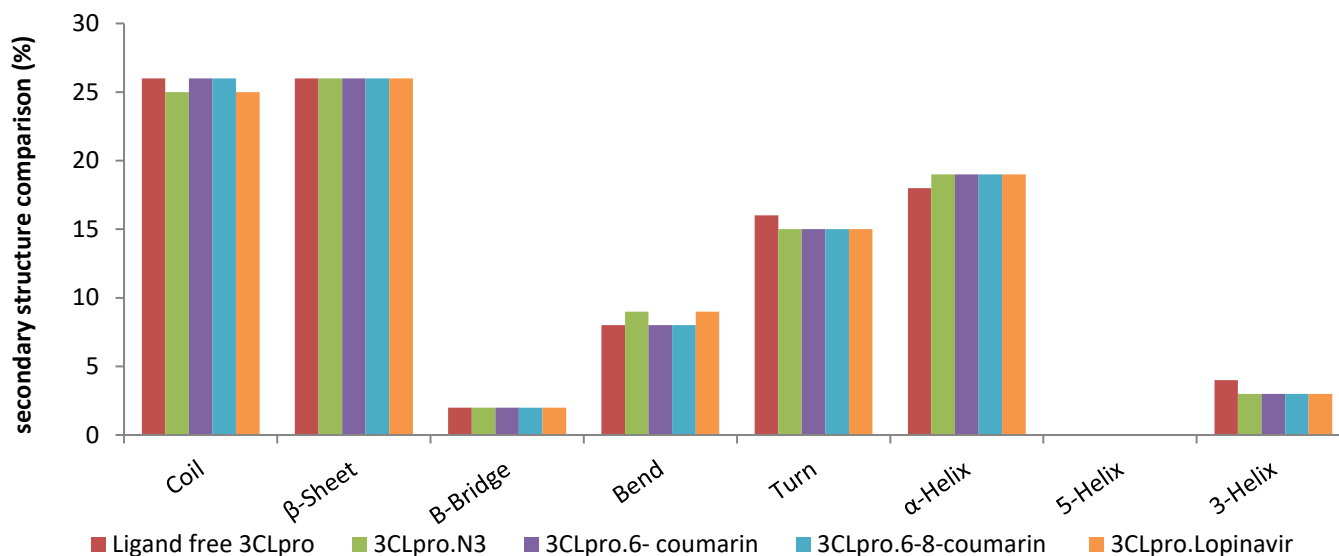


Figure 16. Secondary structure content of the free 3CLpro protein and in the complex with N3, 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin and lopinavir.

3.7 MM-PBSA binding free energy calculation

The binding free energy of the ligand-protein complex was performed to revalidate the ligand affinity to the target receptor for the ligand-protein complex predicted by the molecular docking studies. The MM-PBSA free energy values of the 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin complexes, as well as 3CLpro-N3 and 3CLpro-lopinavir, were calculated from 50 ns trajectories corresponding to every 5 ns time interval. The summation of the non-bonded interaction energies (viz. Van der Waals and electrostatic interaction), polar solvation energy, and SASA energy, were then calculated for both complexes, as shown in Table 7. The calculated ΔG binding energy values of 3CLpro-N3 and 3CLpro-lopinavir complexes were found to be -46.33 kJ/mol and -41.93 kJ/mol (Table 7). On the contrary, the binding free energy values of 3CLpro-6-tert-octyl

coumarin and 6-8-ditert-butyl coumarin complexes were -62.26 kJ/mol and -48.10 kJ/mol and these negative values of ΔG binding energy indicated that the coumarin compounds favorably interact with the target protein of 3CLpro. Among the 3CLpro-coumarin complexes, the 3CLpro-6-tert-octyl coumarin complex exhibited the higher binding free energy, while the 3CLpro-6-8-ditert-butyl coumarin complex showed the lower binding free energy. According to the results of Table 7, the major favorable contributors were van der Waals (ΔE_{vdw}) and electrostatic (ΔE_{elec}) interactions and SASA energy while the polar component of solvation (ΔG polar) contributed unfavorably to the binding of 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin to 3CLpro that the positive value for the polar energy was due to the solvation of ligands by the water molecules, thereby reducing their interactions with the protein.

Table 7. Binding free energy for 3CLpro-6-tert-octyl and 6-8-ditert-butyl coumarins, 3CLpro-N3/lopinavir of SARS-CoV-2 calculated by MM-PBSA analysis.

Complex	Van der Waal energy (ΔE_{vdW}) (Kj/mol)	Electrostatic energy (ΔE_{elec}) (Kj/mol)	Polar solvation energy (ΔG polar) (Kj/mol)	SASA energy (Kj/mol)	Binding energy (Kj/mol)
3CLpro-N3	-74.56	-17.62	57.58	-11.73	-46.33
3CLpro-Lopinavir	-59.05	-15.94	43.76	-10.70	-41.93
3CLpro-6-tert-octyl coumarin	-92.16	-4.24	43.75	-9.61	-62.26
3CLpro-6-8-ditert-butyl coumarin	-67.86	-6.64	33.33	-6.93	-48.10

The MM/PBSA energy components additionally explained the importance of aromatic rings, hydrophobic cores, and H-bond donor-acceptor groups in ligands as well as at the binding site. For determining the key residues involved in the ligand activities as well as understanding the interactions of the ligand with the 3CLpro protein residues, total binding free energy decomposed into the contribution energy of diverse residues at the active site of 3CLpro protein with all the four ligands has been computed and showed that three important regions:

amino acids 26 to 56, amino acids 136 to 148 and amino acids 162 to 194 including Met49, His41, Gly143, Asn142, Cys145, Ser144, Glu166, Gln189, and Met165 were the most, contributive residues. Figure 17 depicts the respective energy contribution. These findings agree with, and mutually support, previously reported results of the main interacting residues within the 3CLpro active site that are deemed critical for effective ligand binding.

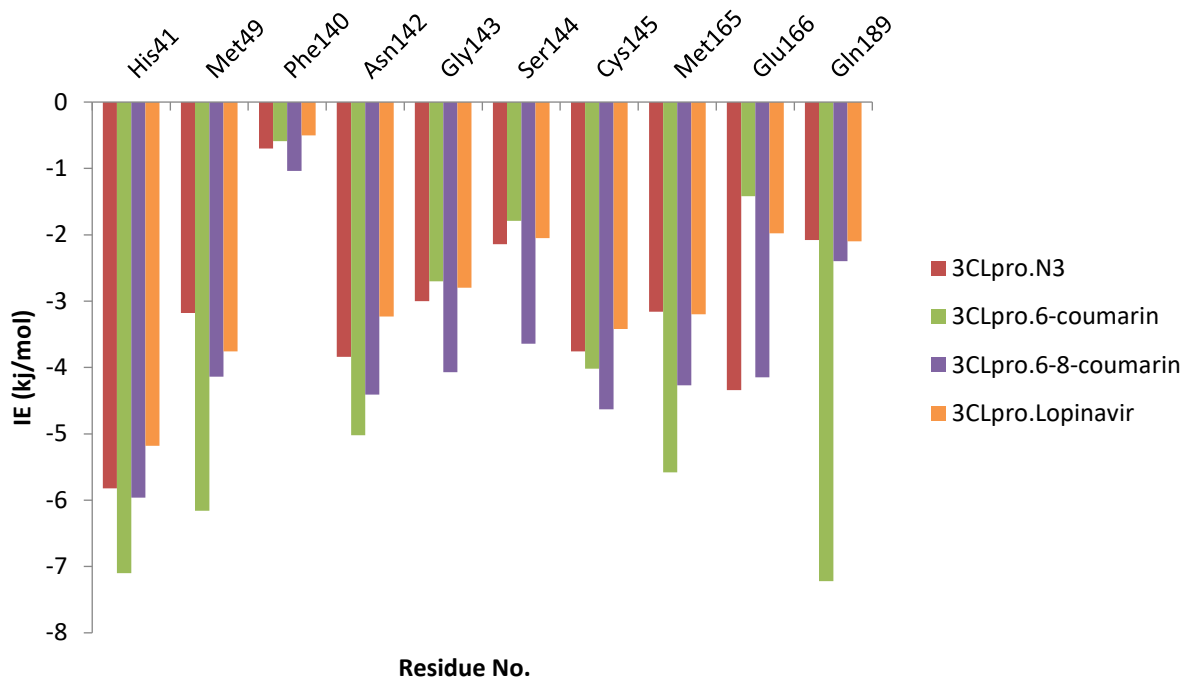


Figure 17. Binding free energies of the residues which have considerable interactions with N3, 6-tert-octyl coumarin, 6-8-ditert-butyl coumarin and Lopinavir.

The results indicated that catalytic dyad (His41 and Cys145) in the 3CLpro-coumarin complexes had a major energy contribution in binding affinity of 3CLpro compared to that of the 3CLpro-N3/lopinavir complexes. From Figure 17, it was found that interactions of Cys 145 (-4.02 and -4.63 kJ/mol) by hydrogen bonds with 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds had the big energy contribution to these complexes; and Asn142 (-5.02 and -4.41 kJ/mol) was also found to be one of the most important residues for activity by hydrogen bonds with 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds, respectively. While the binding free energy values of Cys145 and Asn142 residues in the 3CLpro-N3 complex were -3.76 kJ/mol and -3.84 kJ/mol and for 3CLpro-lopinavir complex were found to be -3.42 kJ/mol and -3.23 kJ/mol, respectively. In addition, Ser144, Gly143, Glu166, and Gln189 promoted the interaction of energy with 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin compounds in a significant manner by establishing electrostatic interactions and H bond formation with these compounds; hence, the

contributions of these residues were advantageous for affinity binding. His41, Met49, and Met165 had the biggest energy contribution to these coumarin complexes that could have hydrophobic interactions as well as π - π stacking with the coumarin compounds and make significant positive contributions to the binding of ligands with 3CLpro. In addition to the catalytic dyad, key residues of Asn142, Gly143, Glu166, and Gln189 favorably contribute to the binding affinity and verify the reliability of the molecular docking results. In general this analysis shows the better binding interactions of 6-tert-octyl coumarin compound on the active site of 3CLpro than 6-8-ditert-butyl coumarin compound.

According to the simulation results, the initial docked structure and the ultimate structure of the 3CLpro-6-tert-octyl coumarin complex had been in a similar binding pocket and showed that ligand-protein conformation was stable after the simulation and docking results of the 6-tert-octyl coumarin with 3CLpro was reliable (Figure 18).

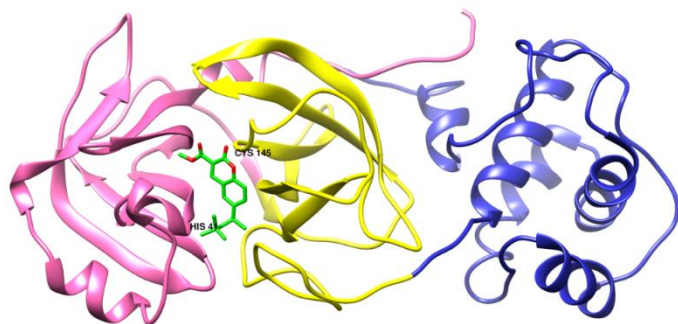


Figure 18. A diagrammatic representation of interaction between 6-tert-octyl coumarin with 3CLpro receptor after 50 ns MD simulations.

With regard to the 3D conformation result of 6-8-ditert-butyl coumarin (Figure 19), the interactions of catalytic dyad residues (His41 and Cys145) and 6-tert-octyl coumarin in the initial docked and ultimate 6-8-ditert-butyl coumarin complex after 50 ns simulations did not change. Also, for the investigation of the system condition during simulation, the 3CLpro-6-tert-octyl coumarin and 3CLpro-6-8-ditert-butyl coumarin structures were extracted from trajectories for every 10 ns (Figure S2 and S3). These snapshots proved the fixed orientation of 6-tert-octyl coumarin and 6-8-ditert-butyl coumarin at the active site of the 3CLpro throughout the simulation.

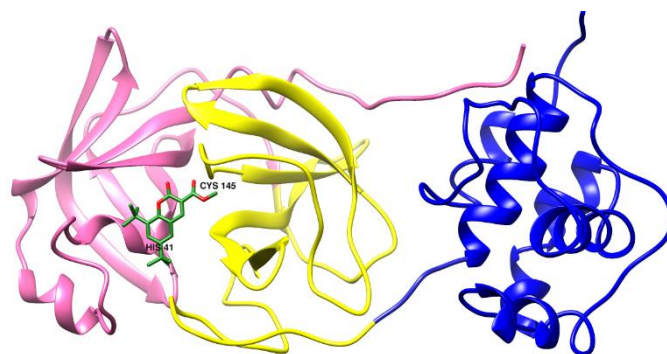


Figure 19. A diagrammatic representation of interaction between 6-8-ditert-butyl coumarin with 3CLpro after 50 ns MD simulations.

According to the present research, 6-tert-octyl and 6-8-ditert-butyl coumarin compounds displayed the molecular docking/ dynamics results and comparable binding energy values with that of N3 and lopinavir. Therefore, further experimental *in vitro/in vivo* studies are suggested to explore probable preclinical and clinical efficiency of these compounds to inhibit main protease protein and treat COVID-19.

CONCLUSION

Although FDA has issued an emergency use authorization on vaccines against COVID-19, drugs for anti SARS-CoV-2 are still extremely important [53]. 3-Chymotrypsin-like main protease (3CLpro) is an attractive target for the inhibition of the viral replication cycle and the treatment of SARS-CoV-2 infections. This study was aimed to investigate the synthesized coumarin compounds as potential inhibitors of coronavirus 3CLpro using *in silico* approaches. These inhibitors could inhibit the 3CLpro with a highly conserved inhibitory effect to both SARS-CoV2 and SARS-CoV. The 6-tert-octyl and 6-8-ditert-butyl coumarin compounds showed significant interactions with one or both of the catalytic residues (His41 and Cys145) of 3CLpro through hydrophilic and hydrophobic bonding. MD results revealed the stability of docked 6-tert-octyl and 6-8-ditert-butyl coumarin structures within the active site of 3CLpro of SARS-CoV-2 with significant binding free energies of -62.26 and -48.10 kJ/mol and also, the pharmacokinetics and ADMET evaluation indicate their efficiency as drug molecules. Based on these findings, these coumarin compounds can be further evaluated against *in vitro* SARS-CoV-2 3CLpro inhibition and viral infection for the drug formulation against SARS-CoV-2 infection.

ACKNOWLEDGMENT

We are grateful to Clinical Biochemistry Research Center, Basic Health Sciences Institute; Shahrekord University of Medical Sciences and Facultad de Ciencias Químicas. Investigador Extramural, Universidad de Concepcion, Concepcion, Chile for financial support of this research. The study was supported financially by the grant (No. 3364) provided by Shahrekord University of Medical Sciences and also, was approved by the Ethics Committee of Shahrekord University of Medical Sciences (IR.SKUMS.REC.1399.117).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. A. Enríquez, C. Sáenz, Serie Estudios y Perspectivas. Sede Subregional de la CEPAL en México. Naciones Unidas. CEPAL. 189, (2021).
2. AJ. Rodríguez-Morales, K. MacGregor, S. Kanagarajah, D. Patel, P. Schlagenhaut, *Travel. Med. Infect. Dis.* **33**, 101578, (2020).
3. Z. Wang, X. Chen, Y. Lu, F. Chen, W. Zhang, *Biosci. Trends.* (2020).
4. O. Yañez, MI. Osorio, E. Uriarte, C. Areche, W. Tiznado, JM. Pérez-Donoso, O. García-Beltrán, F. González-Nilo, *Front. Chem.* **8**, 1273 (2021).
5. S. Hiremath, HV. Kumar, M. Nandan, M. Mantesh, K. Shankarappa, V. Venkataravanappa, CJ. Basha, CL. Reddy, *3Biotech.* **11**, 1-18 (2021).
6. JR. Mora, SA. Cuesta, A. Belhassan, G. Salgado Morán, T. Lakhlifi, M. Bouachrine, C. Peña, L. Gerli, LH. Mendoza-Huizar, (2022).
7. H. Zakaryan, E. Arabyan, A. Oo, K. Zandi, *Arch. Virol.* **162**, 2539-2551 (2017).
8. SM. Thayil, S. Thyagarajan, J. Pharmacogn. *Phytochem. Res.* **8**, 1020-1024 (2016).
9. S. Jo, S. Kim, DH. Shin, M-S. Kim, *J. Enzyme. Inhib. Med. Chem.* **35**, (1):145-151, (2020).
10. IE. Orhan, FSS. Deniz, *Nat. Prod. Bioprospect.* **10**, 171-186 (2020).
11. JS. Mani, JB. Johnson, JC. Steel, DA. Broszczak, PM. Neilsen, KB. Walsh, M. Naiker, *Virus. Res.* **284**, 197989 (2020).
12. AK. Maurya, N. Mishra, *J. Biomol. Struct. Dyn.* **39**, 7306-7321, (2021).
13. SK. Chidambaram, D. Ali, S. Alarifi, S. Radhakrishnan, I. Akbar, *J. Infect. Public. Health.* **13**, 1671-1677 (2020).
14. M. Özdemir, B. Köksoy, D. Ceyhan, M. Bulut, B. YALCIN, *J. Turkish. Chem. Soc.* **7**, 691-712, (2020).
15. M. Özdemir, B. Köksoy, D. Ceyhan, K. Sayin, E. Erçağ, M. Bulut, B. Yalçın, *J. Biomol. Struct. Dyn.* 1-16, (2020)
16. UR. Abdelmohsen, A. Albohy, BS. Abdulrazik, SA. Bayoumi, LG. Malak, IS. Khallaf, G. Bringmann, SF. Farag, *RSC. Advances.* **11**, 16970-16979, (2021).
17. S. Chidambaram, MA. El-Sheikh, AH. Alfarhan, S. Radhakrishnan, I. Akbar, *Saudi. J. Biol. Sci.* **28**, 1100-1108, (2021).
18. R. Abdizadeh, F. Hadizadeh, T. Abdizadeh, *Mol. Divers.* 26, 1053-1076, (2022).
19. YW. Chen, C-PB. Yiu, K-Y. Wong, *F1000Research.* **9**, (2020).
20. TW. Shattuck, Montreal, QC: Chemical Computing Group ULC, 2011.
21. OM. Zárraga, M. Darouch, E. Lisboa, PP. Arroyo, MA. Miranda, *J. Chil. Chem. Soc.* **66**, 5220-5222 (2021).
22. M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, Gaussian 09 revision D. 01,(Gaussian Inc., 2009).
23. CA. Lipinski, F. Lombardo, BW. Dominy, PJ. Feeney, *Adv. Drug. Deliv. Rev.* **23**, 3-25, (1997)
24. DF. Veber, SR. Johnson, H-Y. Cheng, BR. Smith, KW. Ward, KD. Kopple, *J. Med. Chem.* **45**, 2615-2623, (2002).
25. WJ. Egan, KM. Merz, JJ. Baldwin, *J. Med. Chem.* **43**, 3867-3877, (2000).
26. I. Muegge, SL. Heald, D. Brittelli, *J. Med. Chem.* **44**, 1841-1846, (2001).
27. RK. Goel, D. Singh, A. Lagunin, V. Poroikov, *Med. Chem. Res.* **20**, 1509-1514, (2011).
28. N. Khurana, MPS. Ishaq, A. Gajbhiye, RK. Goel, *Eur. J. Pharmacol.* **662**, 22-30, (2011).
29. A. Srivastava, S. Siddiqui, R. Ahmad, S. Mehrotra, B. Ahmad, A. Srivastava, *J. Biomol. Struct. Dyn.* 1-51, (2020).
30. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, AE. Mark, HJ. Berendsen, *J. Comput. Chem.* **26**, 1701-1718, (2005).
31. J. Prajapati, R. Patel, D. Goswami, M. Saraf, RM. Rawal, *Comput. Biol. Med.* 104568, (2021).
32. MJ. Abraham, T. Murtola, R. Schulz, S. Páll, JC. Smith, B. Hess, E. Lindahl, *SoftwareX.* **1**, 19-25, (2015).
33. S. Keretsu, SP. Bhujbal, SJ. Cho, *Sci. Rep.* **10**, 1-14, (2020).
34. S. Verma, AK. Pandey, *3Biotech.* **11**, 1-10, (2021).
35. S. Pant, M. Singh, V. Ravichandiran, U. Murty, HK. Srivastava, *J. Biomol. Struct. Dyn.* **39**, 2904-2913, (2020).
36. R. Ghosh, A. Chakraborty, A. Biswas, S. Chowdhuri, *J. Mol. Struct.* **1229**, 129489, (2021).
37. MA. Said, A. Albohy, MA. Abdelrahman, HS. Ibrahim, *Eur. J. Pharm. Sci.* **160**, 105744, (2021).
38. P. Yadav, M. Rana, P. Chowdhury, *J. Mol. Struct.* **1246**, 131253, (2021).
39. S. Keretsu, SP. Bhujbal, SJ. Cho, *Sci. Rep.* **9**, 1-14, (2019).
40. D. Mishra, RR. Maurya, K. Kumar, NS. Munjal, V. Bahadur, S. Sharma, P. Singh, I. Bahadur, *J. Mol. Liq.* **335**, 116185, (2021).
41. HJ. Berendsen, Jv. Postma, WF. van Gunsteren, A. DiNola, JR. Haak, *J. Chem. Phys.* **81**, 3684-3690, (1984).
42. B. Hess, H. Bekker, HJ. Berendsen, JG. Fraaije, *J. Comput. Chem.* **18**, 1463-1472, (1997).
43. M. Aldeghi, MJ. Bodkin, S. Knapp, PC. Biggin, *J. Chem. Inf. Model.* **57**, 2203-2221, (2017).
44. S. Genheden, U. Ryde, *Expert. Opin. Drug. Discov.* **10**, 449-461, (2015).
45. R. Kumari, R. Kumar, *J. Chem. Inf. Model.* **54**, 10.1021, (2014).
46. TI. Adelusi, A-QK. Oyedele, OE. Monday, ID. Boyenle, MO. Idris, AT. Ogunlana, AM. Ayoola, JO. Fatoki, OE. Kolawole, KB. David, *J. Mol. Struct.* 131879, (2021).
47. T. Topal, Y. Zorlu, N. Karapinar, *J. Mol. Struct.* **1239**, 130514, (2021).
48. A. Daina, O. Michielin, V. Zoete, *Sci. Rep.* **7**, 1-13, (2017).
49. MY. Al-Nour, MM. Ibrahim, T. Elsaman, *Curr. Pharmacol. Rep.* **5**, 255-280, (2019).
50. M. Pooja, GJ. Reddy, K. Hema, S. Dodoala, B. Koganti, *Eur. J. Pharmacol.* **890**, 173688, (2021).
51. P. Chinnasamy, R. Arumugam, India. *Egypt. J. Basic. Appl. Sci.* **5**, 265-279, (2018).
52. P. Sethi, Y. Bansal, G. Bansal, *Med. Chem. Res.* **27**, 61-71, (2018).
53. S. Lau, P. Wang, B. Mok, A. Zhang, H. Chu, A. Lee, S. Deng, P. Chen, K. Chan, W. Song, *Emerg. Microbes. Infect.* **9**, 837-842, (2020).