In a search for potential drug candidates for combating COVID-19: computational study revealed salvianolic acid B as a potential therapeutic targeting 3CLpro and spike proteins

Ayman Abo Elmaaty \(^{a,b}\), Khaled M. Darwish \(^{b,1}\), Muhammad Khattab \(^{c}\), Sameh S. Elhady \(^{d}\), Mohammed Salah \(^{e}\), Mohammed I. A. Hamed \(^{d}\), Ahmed A. Al-Karmalawy \(^{g}\) and Moustafa M. Saleh \(^{e}\)

\(^{a}\)Department of Medicinal Chemistry, Faculty of Pharmacy, Port Said University, Port Said, Egypt; \(^{b}\)Department of Medicinal Chemistry, Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt; \(^{c}\)Department of Chemistry of Natural and Microbial Products, Division of Pharmaceutical and Drug Industries, National Research Centre, Cairo, Egypt; \(^{d}\)Department of Natural Products, Faculty of Pharmacy, King Abdulaziz University, Jeddah, Saudi Arabia; \(^{e}\)Department of Microbiology and Immunology, Faculty of Pharmacy, Port Said University, Port Said, Egypt; \(^{f}\)Department of Organic and Medicinal Chemistry, Faculty of Pharmacy, Fayoum University, Fayoum, Egypt; \(^{g}\)Department of Pharmaceutical Medicinal Chemistry, Faculty of Pharmacy, Horus University-Egypt, New Damietta, Egypt

Communicated by Ramaswamy H. Sarma

ABSTRACT

The global prevalence of COVID-19 disease and the overwhelming increase in death toll urge scientists to discover new effective drugs. Although the drug discovery process is a challenging and time-consuming, fortunately, the plant kingdom was found to have many active therapeutics possessing broad-spectrum antiviral activity including those candidates active against severe acute respiratory syndrome coronavirus (SARS-CoV). Herein, nine traditional Chinese medicinal plant constituents from different origins (Glycyrrhizin \(^{1}\), Lycorine \(^{2}\), Puerarin \(^{3}\), Daidzein \(^{4}\), Daidzin \(^{5}\), Salvianolic acid B \(^{6}\), Dihydrotanshinone I \(^{7}\), Tanshinone I \(^{8}\), Tanshinone Ia \(^{9}\)) previously reported to exhibit antiviral activity against SARS-CoV were virtually screened in silico (molecular docking) as potential inhibitors of SARS-CoV-2 target proteins. The tested medicinal plant compounds were in silico screened for their activity against two key SARS-CoV-2 target proteins; 3CLpro, and Spike binding-domain proteins. Among the tested medicinal plant compounds, Salvianolic acid B \(^{6}\) (Sal-B) showed promising binding affinities against the two specified SARS-CoV-2 target proteins compared to the reference standards used. Hence molecular dynamics simulations followed by calculating the free-binding energy were carried out for Sal-B providing information on its affinity, stability, and thermodynamic behavior within the two SARS-CoV-2 target proteins as well as key ligand-protein binding aspects. Besides, the quantum mechanical calculations showed that Sal-B can adopt different conformations due to the existence of various rotatable bonds. Therefore, the enhanced antiviral activity of Sal-B among other studied compounds can be also attributed to the structural flexibility of Sal-B. Our study gives an explanation of the structure activity relationship required for targeting SARS-CoV-2 3CLpro and Spike proteins and also facilitates the future design and synthesis of new potential drugs exhibiting better affinity and specificity. Besides, an ADME study was carried out on screened compounds and reference controls revealing their pharmacokinetics properties.

1. Introduction

By late of December 2019 in Wuhan City of China, an extraordinary outbreak of pneumonia of unknown cause has arisen. In January 2020, the Chinese Centre for Disease Control and Prevention (CCDC) identified the causative microorganism, and was then named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) or (2019-nCoV). The disease was subsequently given the name COVID-19 by the World Health Organization (WHO) (Sohrabi et al., 2020). This unique coronavirus is related to the previously reported Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) and Middle East Respiratory Syndrome coronavirus (MERS-CoV) based on its genetic similarity (Zhu, 2020).

The 2019-nCoV is extremely contagious and can lead to severe respiratory tract infections with high morbidity and mortality rates (Alnajjar et al., 2020; Yu et al., 2020; Zaki et al., 2020). By early January of 2021, the time finishing this manuscript, there are more than 86.6 million reported cases and about 1.87 million deaths all over the world (Culp, 2020).

The coronavirus genome is composed mainly of two large open reading frames, 1a and 1b (ORF1a and ORF1b) that are

CONTACT Ayman Abo Elmaaty (ayman.mohamed@pharm.psu.edu.eg) Department of Medicinal Chemistry, Faculty of Pharmacy, Port Said University, Port Said 42526, Egypt; Ahmed A. Al-Karmalawy (akarmalawy@horus.edu.eg) Department of Pharmaceutical Medicinal Chemistry, Faculty of Pharmacy, Horus University-Egypt, New Damietta 34518, Egypt

†These authors equally contributed to this work.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2021.1918256.

© 2021 Informa UK Limited, trading as Taylor & Francis Group
protease (3CLpro) (Konno, 2017; Yu et al., 2020). Among the processing of polyproteins (pp1a and pp1ab) are 3CL protein (N) (Forni et al., 2017). Namely, spike (S), membrane (M), envelope (E) and nucleoprotein (N) (Forni et al., 2017).

Moreover, one of the main key proteases responsible for the processing of polyproteins (pp1a and pp1ab) are 3CL protease (3CLpro) (Konno, 2017; Yu et al., 2020). Among the four structural proteins, (S) protein which exists on the virus outer surface is considered the most essential recognition factor for virus binding and entrance binding interactions with host cells receptors (Belouzard et al., 2012). Hence, the two important proteins in 2019-nCoV, 3CLpro and S-protein are considered as potential active targets for many antiviral drugs (Yu et al., 2020).

In the running battle against 2019-nCoV, one strategy to find a solution for this notorious virus is to screen for molecules that are commonly used which may have a therapeutic effect on 2019-nCoV (Wu, 2020).

Meanwhile, medicinal plants have been widely used to treat several infectious diseases. It is estimated that about 25% of the commonly used medicines contain compounds obtained from plants (Mukhtar et al., 2008). Besides, some other medicinal plants were also found to have broad-spectrum antiviral effects; hence they can be screened for newly emerging viral diseases. For example, glycyrrhizin, a bioactive component of liquorice (Glycyrrhiza uralensis), and lycorine isolated from Lycoris radiate were previously reported having strong anti-SARS-CoV activity. In addition, some other medicinal plants were also found to have an anti-SARS activity such as Salvia miltiorrhiza (with the main active constituents, salvianolic acid B, dihydrotanshinone, tanshinone I, and tanshinone Il a) and kudzu (Pueraria Montana) with the main active constituents, puerarin, daidzein, and daidzin) (Al-Karmalawy & Khattab, 2020; Buhner, 2013; Eliaa et al., 2020; Ghanem et al., 2020; Mukhtar et al., 2008).

So, in continuation to our previous work (Al et al., 2021; Elmaaty et al., 2021; Khattab & Al-Karmalawy, 2021; Samra, 2021), and concerning the aforementioned facts, nine medicinal plant constituents of drug origin (glycyrrhizin, lycorine, salvianolic acid B, dihydrotanshinone, tanshinone I, tanshinone Il a, puerarin, daidzein, and daidzin) were screened in silico by our group against the main proteins of SARS-CoV-2 (3CLpro, and S-protein). Molecular docking study highlighted that Salvianolic acid-B (Sal-B) is the most promising candidate amongst other studied drugs exhibiting inhibitory activity against SARS-CoV-2 main protease and spike binding-domain. Therefore, quantum chemical calculations were also utilized herein to study the structural and electronic configurations of Sal-B in detail. Chemical structure of Sal-B (CAS 121521-90-2) is given in Figure 1 along with other synonyms of Sal-B. It is obvious that Sal-B is a big-sized molecule (MWt= 718.6 Da) possessing a number of rotatable single bonds and chiral centers in addition to four aromatic rings including three catechol moieties. Hence, Sal-B is expected to exhibit conformational flexibility along with absorbptive properties of the UV-Vis light and the plane-polarized light.

2. Materials and methods

2.1. Docking studies

2.1.1. Preparation of the tested compounds

The tested medicinal plant compounds (Glycyrrhizin 1, Lycorine 2, Puerarin 3, Daidzein 4, Daidzin 5, Salvianolic acid B 6, Dihydrotanshinone I 7, Tanshinone I 8, and Tanshinone Il a 9), Figure 2, were downloaded from (https://pubchem.ncbi.nlm.nih.gov/) page. Their structures and the formal charges on atoms were checked by the 2D depiction, subjected to energy minimization and the partial charges were automatically calculated. The medicinal plant compounds under investigation, as well as the co-crystallized ligand and control reference, N3 inhibitor and a reference control; umifenovir were used as reference standards for docking of 3CLpro, and S-protein, respectively. The COVID-19 clinically promising anti-viral agent, umifenovir, was assigned as reference control for S-protein as it acts through inhibition of membrane fusion and was useful in managing influenza viruses in Russia and China, while significantly improved the patient conditions suffering from COVID-19 pneumonia (Aktas et al., 2020; Wang, 2017; Zhang et al., 2020).

2.1.2. Preparation of the target receptors (3CLpro, and spike receptor-binding domain)

Protein Data Bank was used to download the crystal structure of SARS-CoV-2 main protease (3CLpro) with (PDB code 6LU7, resolution: 2.16 Å) (Jin et al., 2020), and Spike receptor-binding domain with (PDB code 6ZCZ, resolution: 2.65 Å) (Zhou, 2020). Sequence editor was used for proper protein chain selection;
(Chains A and C for 3CLpro, and Chain E for spike receptor-binding domain). The crystal structures were protonated and hydrogen atoms were added with their standard 3D geometry, automatic correction for any errors in the atom's connection and the type was applied followed by energy minimization, and potential fixation of the receptor and its atoms were done. Site Finder was applied for selection of the same pocket site of the co-crystallized N3 inhibitor for 3CLpro, and NAG for Spike receptor-binding domain, docking was run. Using all default items and dummy atoms of the pocket were created. It is worth noting that Spike receptor-binding domain docking was carried out at the same co-crystallized NAG pocket owing to the belief that NAG is the main component of glycosylated coronavirus spike protein and the presence of NAG at SARS-CoV-2 Spike may hamper antibodies attachment due to glycan shielding and possible epitope masking of an HCoV-NL63, thus deterring immune responses (Banerjee et al., 2020). Therefore, the rationale for exploring the affinity of Sal-B at S-protein NAG site could provide a way for depriving the membrane-bounded protein from its dodging mechanism making the virus more identified throughout targeted immune responses.

2.1.3. Validation of the applied MOE program
Double validation processes were performed for the applied MOE program by redocking the co-crystallized ligands of the two protein receptor pockets of SARS-CoV-2 used for docking studies (6ZCZ and 6LU7, respectively). The valid performance was confirmed by both processes through obtaining low RMSD values (< 2) which indicates the very close positioning of both the native co-crystallized ligand and the docked one inside the receptor pocket in each case. The RMSD value obtained by the redocked NAG ligand in case of 6ZCZ was 0.94 while that obtained by the redocked N3 ligand in case of 6LU7 was 1.77 as illustrated in Figure 3 (McConkey et al., 2002).

2.1.4. Docking of the tested molecules to the two different viral binding sites
Docking of the previously prepared database composed of our tested nine medicinal plant compounds and the co-crystallized reference standards at different specified viral receptors (3CLpro, and S-protein) was performed. The following methodology was applied: the file of the prepared active site was loaded, and the general docking process was initiated. The program specifications were adjusted so that the docking site as (dummy atoms), the placement methodology is (triangle matcher) and the scoring methodology is (London dG). Rigid receptor as refinement methodology and GBVI/WSA dG as the scoring methodology for selection of the best 20 poses from 100 different poses for each tested compound at each receptor pocket. The MDB file of the nine medicinal plant ligands was loaded and general dock
calculations were run automatically. The obtained poses were studied after completion, and the best ones having the best energy scores, RMSD values, and ligand–enzyme interactions were selected and stored for energy calculations. In the beginning, a validation process was also performed for the target receptor by docking the co-crystallized ligand and low RMSD values between docked and crystal conformations indicate a valid performance (Davis & Baker, 2009; McConkey et al., 2002).

2.2. Molecular dynamics simulations

The best-docking scored models of the most promising lead Sal-B in complex with SARS-CoV-2 3Clpro and Spike protein were chosen as starting coordinates for 100 ns all-atoms molecular dynamic simulation using GROMACS-2019 software package (GNU, General Public License; http://www.gromacs.org) and CHARMM36 force field (da Silva et al., 2020). Each ligand-protein complex was solvated within a cubic box of the transferable intermolecular potential with three-point (TIP3P) water model allowing a minimum of 10 Å marginal distance between protein and each side of the 3D box (Izadi et al., 2014). Under periodic boundary conditions implementation, the protein residues were assigned for their standard ionization states at physiological conditions (pH 7.0) and the whole complexes were neutralized via sufficient numbers of $K^+$ and $Cl^-$ ions added via Monte-Carlo ion-placing method (Ross et al., 2018). The MD simulation was conducted over three stages and a 1000 KJ/mol.nm$^2$ force constant was used for restraining all heavy atoms and preserving original protein folding (Helal, 2020). The first stage involved initial optimization of each system geometry using 5,000 iterations (5 ps) with the steepest descent algorithm. The subsequent

---

**Figure 3.** 2D and 3D representations reference controls (A) NAG, (B) N3 showing the re-docking processes between the native co-crystallized ligands (red) and the docked ones (green) at 3CLpro and S-protein with PDB: 6ZCZ and 6LU7, respectively.
step involved system two-staged equilibration for 100,000 iterations (100 ps) at each stage. The first equilibration stage was preceded under a constant number of particles, Volume, and Temperature (NVT) ensemble guided by the Berendsen temperature coupling method for regulating the temperature within the 3D box (Golo & Sha, 2002). Subsequently, the second equilibration stage was performed under a constant number of particles, Pressure, and Temperature (NPT) ensemble at 1 atm and 303.15 K guided by using the Parrinello-Rahman barostat (Tuble et al., 2004).

Finally, the MD simulations were run for 100 ns under constant pressure (NPT ensemble) and long-range electrostatic interactions were computed using Particle Mesh Ewald (PME) algorithm (Darden et al., 1993). However, the implemented linear constraint LINCS method was used to constrain all covalent bond lengths, including hydrogens, allowing an integration time step size of 2 fs (Hess et al., 1997). The non-bounded interactions, Coulomb and van der Waals interactions were truncated at 10 Å using the Verlet cut-off scheme (Páll & Hess, 2013). Computing comparative data, including root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), solvent accessible surface area (SASA), and radius of gyration (Rg), was performed through analyzing the MD trajectories using the GROMACS built-in tools and Visual Molecular Dynamics 1.9.3 (VMD) package (the University of Illinois at Urbana-Champaign, USA) (Humphrey et al., 1996). All MD simulation data are reported with standard deviation of multiple replicas (triplicates). Moreover, the VMD’s Distance Calculation Tool was utilized to calculate the change in the distance between the specified ligand/protein atoms over the whole simulation period to monitor and investigate the possibility of interactions of ligands with the most important protein residues. Finally, the binding-free energy between the ligand and protein was estimated via the GROMACS “g_mmpbsa” module (Kumari et al., 2014). The Pymol graphical software ver. 2.0.6 (Schrödinger™, NY, USA) was utilized for figure generation of ligand-protein conformational deformation (DeLano, 2020).

### 2.3. Quantum mechanics calculations

Quantum mechanical calculations were performed on Sal-B. The two-dimensional structure of Sal-B was sketched using Chemdraw, where it was then prepared by Gaussiaview software. The structure of Sal-B was gradually optimized at different basis sets integrated with the Becke’s three-parameter hybrid exchange-correlation functional (B3LYP) (Becke, 1993; 1993). The 3-21G, 6-31G, 6-311G, 6-311G*, and 6-31 + G* were the basis sets used. All chemical and physical properties of Cal-B were calculated at B3LYP/6-311 + G* level of theory. While computed properties of Cal-B at its excited state were calculated using the td-DFT method. The potential energy surface (PES) scan of Sal-B structure was performed for different dihedral angles utilizing B3LYP/6-31G model. The most energetically optimized structure, extracted from the four performed PES scans, was then taken and reoptimized at B3LYP/6-311G, B3LYP/6-311G*, and B3LYP/6-31 + G*. For all optimization calculations, no imaginary frequencies were identified for the optimized structures, therefore all of the optimized geometries are true local minima structures. All calculations were performed using GAUSSIAN 09 Revision C.01 (Frisch, 2010) on Swinburne supercomputing facilities.

### 2.4. Physicochemical, ADME, and pharmacokinetic properties prediction

The free Swiss ADME web tool available from the Swiss Institute of Bioinformatics (SIB) was used for the computation of the physicochemical descriptors as well as to predict the ADME parameters, pharmacokinetic properties of the isolated medicinal plant products 1-9. Structures of the studied compounds were converted to SMILES notations, then submitted to the online server for calculation (Daina et al., 2017).

### 3. Results and discussion

#### 3.1. Docking study

By analyzing the docking results of the tested compounds, it was found that most of these compounds showed promising binding scores and modes compared to the co-crystallized inhibitor at the two specified receptors. Interestingly, Sal-B showed the most promising results between tested compounds at 3CLpro and S-protein. Sal-B binding energy score was slightly higher than N3 and clearly higher than umifenovir at 3CLpro and Spike receptors, respectively. So, it was concluded that Sal-B is highly stable at its different binding sites on COVID-19 virus, hence, among tested compounds, Sal-B is expected to exert the most antiviral inhibition on COVID-19. The results of Sal-B binding interactions at 3CLpro and S-protein in comparison with docked references are shown in **Tables 1** and **2**. The detailed binding modes of all tested medicinal plant compounds (1-9) and docked

---

**Table 1.** Receptor binding energies and interactions of the tested Sal-B and N3 inhibitor into the N3 binding site of COVID-19 main protease (3CLpro).

| Drug No. | Tested compound | S* Kcal/mol | RMSDb | Amino acid bond | Distance Å |
|---------|----------------|-------------|-------|----------------|------------|
| 6       | Sal-B          | -9.4339     | 1.9366| GLU 166/H-donor | 2.69       |
| 10      | Reference (N3) | -9.4286     | 1.6312| ASP 187/H-acceptor | 3.10      |

*S: the score of placement of a compound into the binding pocket of protein using London dG scoring function, RMSDb: root-mean-squared-deviation (RMSD) between the heavy atoms of the predicted pose (after refinement) and those of the crystal structure (before refinement).

**Table 2.** Receptor binding energies and interactions of the tested Sal-B and umifenovir inhibitor at COVID-19 Spike receptor-binding domain.

| Drug No. | Tested compound | S* Kcal/mol | RMSDb | Amino acid bond | Distance Å |
|---------|----------------|-------------|-------|----------------|------------|
| 6       | Sal-B          | -7.4614     | 1.6796| ASP 364/H-donor | 3.35       |
| 12      | Reference (Umifenovir) | -5.1880 | 1.6980| ASN 343/pi-H | 3.52       |

---

4: the RMSD_refine: the root-mean-squared-deviation (RMSD) between the heavy atoms of the predicted pose (after refinement) and those of the crystal structure (before refinement).
references, and all of their 2 D, 3 D binding interactions, surface and maps, and 3 D positioning inside the protein pocket are found in supplementary data. However, the 3 D binding interactions and 3 D protein positioning of Sal-B along with at the native ligand at 3CLpro and S-protein are deposited in Table 3.

3.2. Structure activity relationship

Studying the structure-activity relationship of our tested medicinal plant compounds according to their binding affinities to COVID-19 target receptors (3CLpro, and S-protein) showed the following interesting results:

The presence of non-rigid flexible structure (Conformational flexibility) containing dihydrobenzofuran scaffold attached to bi-hydroxy phenyl ring through carboxylic and carboxylate moieties (Sal-B, 6), experienced exclusively the best activity ever at the two specified receptors and better than triterpenoid saponins (Glycyrrhizin, 1) than phenanthridine alkaloids (Lycorine, 2).

Moreover, 7-hydroxy isoflavones (Puerarin, Daidzein, and Daidzin, 3, 4, and 5) showed lower binding affinities than triterpenoid saponins (Glycyrrhizin, 1) at 3CLpro, and Spike.

Table 3. The 3 D view of binding interactions and the 3 D positioning of Sal-B compared to the docked references at COVID-19 3CLpro and Spike target receptors. (Red dashed lines refer to hydrogen bonds. Black dashed ones refer to H-pi bonds).
Among isoflavone compounds, it was noticed that substitution of chromene ring with a sugar moiety as glucose at positions 7 and 8, (Daidzin and Puerarin, 5 and 3) increases their binding affinities at the two specified receptors and better than unsubstituted one (Daidzin, 4). Besides, isoflavones containing 7-O-glucose (Compound 5); showed better binding affinity than 8-C-glucose isoflavones (Compound 3) at Spike and lower binding affinity than 8-C-glucose isoflavones (Compound 3) at 3CLpro.

On the other hand, lipophilic abietane biterpenoids (Dihydrotanshinone I, Tanshinone I, and Tanshinone IIa, 7, 8, and 9) showed different binding affinities for 3CLpro, and Spike. At 3CLpro, based on its lipophilic parameters, Tanshinone IIa, 9 (much lipophilicity owing to partial saturation of one rings of naphthalene moiety substituted by two methyl groups) showed better binding affinities than Dihydrotanshinine I, 7 (substitution of fully unsaturated naphthalene ring by one methyl group and partial saturation of furan ring, (2,3-dihydro furan)) than tanshinone I, 8 (substitution of fully unsaturated naphthalene ring by one methyl group). However at Spike, tanshinone I, 8 showed better binding affinity to these receptors than tanshinone IIa, 9 and dihydrotanshinine I, 7.

The structure activity relationship and order of activity of the studied medicinal plants are illustrated in Figure 4. Besides, the detailed binding affinity arrangements for all tested medicinal plant compounds at each COVID-19 target proteins are declared in the supplementary information.

3.3. Molecular dynamics simulations

Considering it as an efficacious approach for validating the stability of the predicted docked Sal-B acid-3CLpro complex, an all-atom MD simulation study was performed. Adopting such a study would also provide valuable information regarding the dynamic behavior of both Sal-B and 3CLpro protein as well as evaluate the ligand’s key binding interactions with important 3CLpro residues (Karplus & Petsko, 1990).

Therefore, the predicted Sal-B-protein complex was enrolled in a 100 ns all-atom MD simulation for being compared with the potent inhibitor, N3, or reference control, umifenovir, bounded to the 3CLpro or S-protein target, respectively.

3.3.1. MD Simulation analysis of 3CLpro complexes

The stability profile of both Sal-B and N3 in complex with SARS-CoV-2 main protease (3CLpro) was monitored using the GROMACS command line gmx_rmsd. So, their respective RMSD values are estimated throughout the simulation runs, hence provides an inference regarding the deviation extent for a group of atoms (protein, ligand, or even ligand-protein complex) in relation to the respective initial reference structure (Schreiner et al., 2012). Thus, high RMSD values would be correlated to significant instability, being related to changes within the conformation of the investigated molecule. Moreover, ligands depicting high RMSD values would suggest inadequate ligand accommodation within the studied pocket across the adopted MD simulation timeframes (Liu et al., 2017).

Within the presented MD simulation, both investigated protein targets exhibited successful conversion within the designated 100 ns timeline (Figure 5A). The obtained complex RMSD trajectories, in respect of their backbone, showed minimal fluctuations. This inferring great stability for the protein in complex with Sal-B (3.001 ± 0.31 Å) as being similar to that for N3 inhibitor (2.980 ± 0.17 Å). Following the start of the MD simulation runs, the RMSD tones rises a little bit throughout the few initial frames. This was preceded till the RMSDs level off at around 25 ns where the trajectories proceeded around the average value till the end of the MD simulation at 100 ns. The latter presented complex dynamic behaviors that indicate the confinement of both ligands within their own respective protein active site. Moreover, the obtained RMSD trajectories ensure that the adopted equilibration and minimization steps, for conditioning the ligand-protein complexes prior to MD simulation runs to bring smooth valid MD simulations. It worth noting that the 3CLpro- Sal-B complex showed slightly higher RMSD values (maximum 4.091 ± 0.27 Å and minimum 1.328 ± 0.20 Å) as compared to those of the N3-bound 3CLpro complex (maximum 3.17 Å and minimum 1.57 Å) suggesting comparable stability and minimal fluctuation pattern.

For further evaluation of the ligand-protein complex stability, the RMSD trajectories of both protein and ligand were monitored. Both proteins adopted comparable dynamic behavior as that observed with their respective complexes (Figure 5B). A little elevation of the protein RMSD tones, with respect to their C-alpha atoms, was depicted at the first frames of MD simulation where then an equilibrium plateau was achieved till the end of simulation window with final RMSD of 3.117 ± 0.08 Å and 3.355 ± 0.54 Å, for Sal-B and N3-bound proteins, respectively. Such protein behavior is typical for optimum MD runs since all the applied constrains, prior to the simulation, were released. The protein starts to relax till reaching an equilibration state around which the RMSD revolves until reaching the MD simulation end. The above results ensure superior stability of Sal-B-bound protein being further correlated to the ligand’s backbone RMSD trajectories (Figure 5C). Sal-B showed a well-behaved simulation for rapidly reaching its dynamic equilibration. Then, the RMSD tones were maintained below 6 Å (i.e. nearly 1.5-fold the protein RMSDs) with minimal fluctuations till 100 ns (5.019 ± 0.21 Å).

Concerning the N3 ligand, the depicted RMSD trajectories showed strong fluctuations across several intervals within the MD simulation run (5.040 ± 0.55 Å). The latter observation might suggest that the potent 3CLpro inhibitor, N3, exhibited significant flexibility within the binding pocket. However, the depicted low average RMSD value confirms that the ligand was still confined within the protein binding site.

Interestingly, the furnished ligand RMSDs were approximately 1.5-fold higher than the RMSDs of respective bounded protein, particularly at the last 50 ns of the MD simulation run. Exhibiting such dynamic behavior, the RMSD convergence analysis has been highly correlated to successful convergence and retention of Sal-B within the 3CLpro...
Further monitoring and validation of the MD simulation convergence was done through performing the principal component analysis (PCA) that examines the collective dynamic behaviour/motion of the protein out of the MD simulation trajectories. The PCA approach relies on the constructing and diagonalizing covariance matrix from the C-\(\alpha\) atomic coordinates of the investigated protein for capturing strenuous atom motions through the eigenvalues and eigenvectors (David & Jacobs, 2014). Generally, the eigenvectors of the covariance matrix elucidate the overall motion direction of the atoms, whereas, eigenvalues are the representative values of atom-wise contributions of motion. Thus,

**Figure 4.** Structure-activity relationship of tested medicinal plant compounds against SARS-CoV-2 target proteins (3CLpro and S-protein).
both eigenvalues and eigenvectors yield the modes of collective motion and their amplitudes. For constructing and diagonalizing the covariance matrix, the GROMACS "gmx_covar" command script was used, while, "gmx_anaeig" was utilized for visualizing the most dominant modes (eigenvectors 1 and 2) as well as calculating the overlap between the trajectory coordinates and principal components. With the corresponding eigenvalues providing an indication of the dynamic behaviour and degree of fluctuations, lower trace values of the covariance matrix correlates with minimal escalation within the collective protein motion and so denoting MD simulation convergence (Aier et al., 2016; Pandey et al., 2018; Srikumar et al., 2014). Applying PCA on the MD trajectories at the last 40 ns and comparing it with the rest of MD simulation frames would provide fundamental tool to monitor and validate the convergence within the conducted MD simulation. Notably, the average trace value of the covariance matrix at the last 40 ns was depicted lower as compared to that along the first 60 ns MD simulation trajectories. For Sal-B bounded protein an average trace value of 6.157 ± 0.60 nm² and 3.843 ± 1.12 nm² were assigned for the trajectories along first 60 ns and last 40 ns, respectively. This was obvious since the conformational space covered by the 3CLpro protein along the first 60 ns was wider (Figure 6). Further lower values were obtained with the atoms of the N3 bound protein (2.918 ± 0.48 nm² at last 40 ns and 5.371 ± 2.73 nm² at first 60 ns). All these finding ensure the higher stability of the protein atoms at the last 40 ns which in turn conferring a validated convergence of the adopted MD simulation.

For gaining more insights regarding the investigated complex stability, the radii of gyration (Rg) were monitored across the whole MD trajectories using the GROMACS "gmx_gyrate" command script. This stability parameter accounts for global stability of either ligand or protein ternary structure, where Rg is the mass-weighted RMSD for a group of atoms relative to their common mass center (Likić et al., 2005). Therefore, sustained stability/compactness of the investigated molecule would be inferred through depicted low Rg values achieving a plateau around an average value. Within the furnished study, the obtained Rg tones confirm the preferential stability of the Sal-B-3CLpro complex as compared to those of N3 (Figure 7). Steadier Rg trajectories were obtained for the Sal-B complex with lower maximum, average, and minimum values (22.895 ± 0.23 Å, 22.402 ± 0.08 Å, and 21.898 ± 0.02 Å, respectively) suggesting
Figure 6. Projection of protein atoms in phase space along the first two principal eigenvectors (1 and 2). (A) Sal-B bounded 3CLpro protein; (B) N3 bounded 3CLpro protein. The PCA calculations were conducted cross selected MD simulation trajectories for the first 60 ns and last 40 ns, exhibiting an expected significant structural stability and convergence.

Figure 7. Global stability analysis of ligand-3CLpro protein complexes throughout 100 ns all-atom MD simulation. (A) complex Rg; (B) protein Rg; (C) ligand Rg, versus MD simulation time in nanoseconds. Sal-B/3CLpro and N3/3CLpro complexes are illustrated in red and blue color, respectively.

Table 4. The Rg values for investigated ligand-3CLpro complexes across the all-atom MD simulation.

| Reference atom group | Sal-B-3CLpro complex | N3-3CLpro complex |
|----------------------|-----------------------|-------------------|
|                      | Maximum (Å)           | Average (Å)       | Minimum (Å)     | Maximum (Å) | Average (Å) | Minimum (Å) |
| Complex              | 22.895 ± 0.23         | 22.402 ± 0.08     | 21.898 ± 0.02   | 23.984 ± 0.16| 22.364 ± 0.10| 22.489 ± 0.45|
| Protein              | 22.949 ± 0.21         | 22.438 ± 0.08     | 21.937 ± 0.02   | 22.013 ± 0.16| 22.323 ± 0.09| 22.427 ± 0.35|
| Ligand               | 6.186 ± 0.07          | 6.116 ± 0.14      | 4.693 ± 0.17    | 7.006 ± 0.07 | 5.739 ± 0.25 | 4.826 ± 0.03 |
compaction and stability of the ligand within the protein binding. The latter complex Rg findings were highly correlated with those of Sal-B ligand and its bounded protein since minimal fluctuations and low Rg values were observed (Table 4). On the other hand, N3 depicted strong fluctuations across the MD simulation window being most obvious for the ligand Rg trajectories (maximum 7.006 ± 0.07 Å, average 5.739 ± 0.25 Å, and minimum 4.826 ± 0.03 Å), suggesting loss of compactness and intermolecular binding at these fluctuations. All obtained Rg findings came in high agreement with the previous RMSD analysis.

To evaluate the impact of ligand binding on the conformational flexibility of protein residues, relative difference root-mean-square fluctuation (ΔRMSF) was estimated for each ligand-bound protein relative to SARS-CoV-2 3CLpro apo state. The provided analysis parameter, RMSF, provides the mean deviation of each protein residue, with respect to its backbone, from its respective initial position at the minimized reference structure (Benson & Daggett, 2012). Thus, this flexibility parameter could further correlate to the stability of the ligand-protein complex since residue immobility can be evident through the influence of significant ligand interactions with the protein residues. The command line, GROMACS “gmx_rmsf,” was used to estimate the individual backbone RMSF for each investigated protein in terms of amino acids versus MD simulation time in ns. Significant alterations within the structural conformations of protein residues were correlated to the predefined ΔRMSF threshold at 0.30 Å, where ligand limited the residues mobility at values above such cut-off (de Souza et al., 2019). The ΔRMSF analysis is best conducted across frames exhibiting significant structural stability the thing that made analysis at the last 40 ns is highly relevant (Dong et al., 2018).

To the most of interest, both bounded ligands induced significant limited mobility (ΔRMSF > 0.3 Å) for the 3CLpro residues at three distinct residue ranges including; range-I (47-58), range-II (165-169), and range-III (202-278). Except for the residue range-I, the two less mobile residue ranges were at comparable ΔRMSF trajectories across the designated MD simulation window. Within the residue range-I, the ΔRMSF trajectories for the N3-bound protein were depicted at higher values (up to 2.917 ± 0.17 Å) as compared to those for Sal-B (maximum 2.286 ± 0.20 Å). The latter suggests the great impact of residues range-I on the stability of the N3-3CLpro complex since these residues were the most immobile compared to the rest of the core residues. It worth noting that residues within the residue range-III are at distances being greater than 29 Å from the bounded ligands the thing that can infer the capability of the 3CLpro binding site to accommodate large-sized inhibitors without influencing the site general stability. Regarding residues with the highest fluctuations, terminal free residues are of the highest negative ΔRMSF (up to – at di ± 1.20 Å and −4.984 ± 0.48 Å; for N3 and Sal-B, respectively). This behavior is highly reasoned since these terminal residues are most likely to fluctuate at

---

Table 5. Calculated ΔRMSF* trajectories of ligand-3CLpro proteins along the most stabilized structural frames (last 40 ns).

| Binding site | Sub-pocket | Residue | Sal-B-3CLpro protein | N3-3CLpro protein |
|--------------|------------|---------|---------------------|------------------|
| S1’          | His41      | 0.525 ± 0.06 | 0.562 ± 0.02       |
|              | Gly143     | −0.066 ± 0.13 | −0.265 ± 0.21      |
|              | Ser144     | −0.211 ± 0.10 | −0.377 ± 0.04      |
|              | Cys145     | −0.139 ± 0.07 | −0.425 ± 0.002     |
| S1           | Phe140     | −0.138 ± 0.35 | −0.184 ± 0.33      |
|              | Leu141     | 0.089 ± 0.17  | −0.278 ± 0.59      |
|              | Asn142     | 0.331 ± 0.10  | 0.057 ± 0.39       |
|              | His163     | 0.033 ± 0.09  | 0.052 ± 0.01       |
|              | Glu166     | 0.475 ± 0.09  | 0.415 ± 0.14       |
| S2           | Met49      | 2.034 ± 1.00  | 2.917 ± 0.17       |
|              | Tyr54      | −0.663 ± 0.39 | 0.201 ± 0.16       |
|              | His164     | 0.303 ± 0.05  | 0.322 ± 0.003      |
|              | Asp187     | 0.398 ± 0.28  | 0.397 ± 0.14       |
|              | Arg188     | 0.692 ± 0.34  | 0.617 ± 0.14       |
| S3           | Met165     | 0.463 ± 0.07  | 0.493 ± 0.02       |
|              | Leu167     | 0.526 ± 0.13  | 0.457 ± 0.13       |
|              | Glu189     | 0.773 ± 0.59  | 0.925 ± 0.01       |
|              | Thr190     | 0.615 ± 0.63  | 0.787 ± 0.09       |
|              | Gin192     | −0.267 ± 1.28 | 0.272 ± 0.07       |

*Relative difference root-mean-square fluctuation (ΔRMSF) ± SD was estimated for each ligand-bound protein relative to SARS-CoV-2 3CLpro apo state. Residues exhibiting significant immobility (ΔRMSF above 0.3 Å) are only written in bold and representative ΔRMSF value is highlighted.

---

Figure 8. Relative ΔRMSF analysis of ligand-3CLpro protein complexes throughout 100 ns all-atom MD simulation. Protein backbone ΔRMSF trajectories were determined from independent MD simulated SARS-CoV-2 3CLpro apo state against the complexed protein with either Sal-B or N3, which were shown as a function of residue number; 1-to-306. Sal-B/3CLpro and N3/3CLpro complexes are illustrated in red and blue color, respectively.
the highest deviations in comparison to core residues the thing that is typical for a well-behaved MD simulation.

Comparative analysis of the furnished ΔRMSF trajectories permitted more insights regarding ligand interactions with the key residue lining the 3CLpro binding site (Table 5). For both ligand-protein complexes, the S1’ subsite catalytic His41 illustrated significant immobility with ΔRMSF value of 0.525 ± 0.06 Å and 0.562 ± 0.016 Å for Sal-B and N3-bound protein, respectively. Nevertheless, most of the rest S1’ subsite residues, including the other catalytic residue (Cys145), showed regained mobility despite the presence of bound ligand at close proximity. It worth mentioning that residues being vicinal to His41 (Cys44-to-Glu47) exhibited significantly high RMSF trajectories (1.162 ± 0.07 Å and up to 2.433 ± 0.18 Å) for the investigated complexes. Thus, it is evident that His41 has a greater impact on ligand-protein stability rather than that of Cys145. Regarding residue comprising the subsite S1, only Glu166 exhibited significant immobility for both complexes (0.474 ± 0.09 Å and 0.415 ± 0.14 Å for Sal-B and N3, respectively) suggesting the residue’s key role in ligand-pocket anchoring. Moving towards the residues lining both S2 and S3 subsites, it was depicted that several residues including: Met49, His164, Met165, Leu167, Asp187-to-Thr190, exhibited significant immobility/stability within the investigated complexes. The presented ΔRMSF analysis suggested comparable stability for Sal-B in regard to that of N3 within the 3CL-pro binding site (Figure 8). This came in high agreement with furnished findings at the above Rg and RMSD analysis. However, it worth mentioning that more preferential stability for the N3 can be assigned at residues being vicinal to the residue range-I and -III where higher fluctuations were depicted within Sal-B-bound protein showing ΔRMSF values −0.659 ± 1.50 Å to −0.165 ± 2.49 Å and −0.449 ± 1.39 Å to −0.225 ± 0.76 Å, respectively. This might be correlated to less favored Sal-B-protein interaction near such residue range.

Investigating the hydrogen bonding between the 3CLpro residues and investigated ligands, over the 100 ns MD simulation, was considered crucial for understanding the observed conformational changes and stability of ligand-protein complexes (Figure 9). Using the VMD “Hydrogen bonds” tool, it was useful to explore the established ligand-protein hydrogen bond interactions and their relative frequencies (Humphrey et al., 1996). The cut-off values for hydrogen bond (Donor-H…Acceptor) distance and angle were assigned at 3.0 Å and 20°, respectively (Albuquerque et al., 2020; de Souza et al., 2019). Interestingly, several hydrogen bond pairs were depicted including the key 3CLpro residues responsible for Sal-B recognition and anchoring towards the binding site (Figure 9A). The highest timeline was assigned for the catalytic His41 sidechain (61 ± 0.33%) and Glu166 mainchain-OE2 (45 ± 0.01%) indicating the key role of these residues for ligand binding the thing that is in good agreement with ΔRMSF analysis. It worth noting that the Glu166 sidechain-O depicted a greatly fluctuating hydrogen bonding with the Sal-B H3O-atom (maximum-to-minimum distances = 13.743 ± 0.01 to 0.613 ± 0.001 Å). The latter observation infers the superior contribution of Glu166 mainchain over its sidechain for stabilizing the Sal-B-protein complex.

Residues of the S2 subsite, Asp187, and Arg188 mainchain, showed moderate lifetime (31 ± 0.003% and 15 ± 0.11%, respectively) being occurred mainly within the first 40 ns and at respective average distances of 1.662 ± 0.28 Å and 3.789 ± 0.63 Å from the ligand. Other residues, including Gln189 sidechain (3 ± 0.43%) and Ser144 mainchain (2 ± 0.17%), showed steady but distant hydrogen bonding with the Sal-B at average distances of 4.357 ± 0.62 Å and 5.567 ± 0.58 Å, respectively. Even the Sal-B/Phe140-mainchain hydrogen bond pair has shown minimal occupancy (0.1 ± 0.009%), despite showing promising results within the initial docking study. This hydrogen bond was rapidly lost at the initial MD simulation frames since the ligand moves far from this S1 subsite residue being fluctuating around an average distance of 4.094 ± 0.70 Å till the end of the MD run. Despite the limited contribution of S1’ subsite residue for Sal-B hydrophilic binding, one vicinal residue, Thr26, showed significant and almost steady hydrogen bonding (32 ± 1.23%) with Sal-B O13-atom at an average distance of 1.792 ± 1.38 Å.

Regarding the hydrogen bond analysis for the N3-3CLpro complex, a lower number of polar contacts were depicted as compared to Sal-B (Figure 9B). Adopting the above hydrogen bond distance and angle cut-offs has revealed only 16 hydrogen bonds versus 85 for Sal-B across the MD simulation window. Acting as a hydrogen-bond donner, the catalytic His41 sidechain showed moderate occupancy (16 ± 0.74%) with the steadiest tones along with the entire MD simulation window at an average distance of 1.906 ± 0.27 Å from the ligand NH-atom. The S1 subsite residue, Glu166, depicted favored hydrogen bonding through its mainchain-NH which was at the first 65 ns time frame. This hydrogen bond pair was at a comparable frequency (16 ± 0.41%) as that of the His41 residue. Both previous findings confirm the important role of His41 sidechain and Glu166 mainchain for anchoring both proteinomimetic and non-peptide based small ligands within the 3CLpro binding site. Notably, bonding with Glu166 was lost (reaching up at 5.999 ± 1.02 Å) beyond 65 ns the thing that could be correlated to the significant elevation of complex Rg trajectories following the same 65 ns simulation window. Therefore, the compactness of the ligand-protein complex was highly correlated with the mutual interactions between ligand and protein, since Rg elevation is related to a more weakened ligand-protein binding.

Other pocket residues, including Glu189 sidechain and Thr190 sidechain, showed interrupted hydrogen bonding exhibiting several maxima within different time intervals. Nevertheless, Glu189 sidechain hydrogen bonding exhibited the highest frequency (24 ± 0.99%) inferring the significant contribution of Glu189 for N3 anchoring at the S3 subsite. Finally, poor hydrogen bonding was depicted between N3 and Gly143 mainchain (0.9 ± 1.23%) as well as the vicinal Ser46 mainchain (7 ± 0.62%) along the 100 ns MD simulation. Both hydrogen bond pairs were rapidly lost following a few initial MD frames. It was rationalized that losing or weakening of several hydrogen bonding, previously identified relevant through initial docking analysis, could be related to...
significant conformation/orientation changes for N3 ligand within the binding pocket. Based on these regards, a conformational analysis was performed at selected trajectories to illustrate the differential ligand binding and possible structural changes across the 100 ns MD simulation.

Selected frames of each ligand-protein system were extracted and minimized to a gradient of 0.001 Kcal.mol$^{-1}$Å$^{-2}$ using MOE package for further analysis of key structural alterations. For both ligand-3CLpro complex, the comparative conformations analysis was performed at 0 and 100 ns of the MD simulation run. Notably, a significant molecular drift was depicted for Sal-B towards the S1$'$ subsite of the 3CLpro binding pocket (Figure 10A). This molecular drift was mediated by weakened or even losing several initial hydrogen bond interactions, with significant residues of the S1 and S2 pocket, by the end of the MD simulation. Hydrophilic contacts between the ligand’s dihydrobenzofuran scaffold and S2 Glu166/S1 Phe140 mainchains were weakened resulting in almost 2 Å ring shift adopting a new conformation being at closer proximity towards the S1$'$ subsite. The new conformation of the dihydrobenzofuran ring was stabilized via an extensive hydrogen bond network with Gly143 and Ser46 mainchains.

Moving towards the N3-3CLpro complex, a less significant drift was depicted for N3 as compared to Sal-B (Figure 10B). However, the dynamic drift of N3 was similar to Sal-B as being directed towards the S1$'$ subsite. Interestingly, several initial hydrogen bonding between the N3 and key pocket residues were conserved across the MD simulation window. The hydrogen pairs between phenyl ester arm and S1 Glu166 mainchain as well as N3 proline residue and S3 Thr190 were conserved till the end of the MD simulation run. On the other hand, the N3 modified amino acid lost its initial hydrophilic contacts with Ser46 sidechain and S1$'$ Gly143 main chain resulting in a deeper anchoring of the oxazole ring towards the vicinal cleft of the S1$'$ subsite. Same as Sal-B, the stability of N3 at 100 ns, particularly for the modified amino acid, was mediated by the Thr26 polar interaction.

3.3.2. MD Simulation analysis of spike (S)-protein complexes

The stability of the salvianolic-spike complex was confirmed through the obtained findings of RMSD, Rg, and RMSF trajectories for both ligands and targets. Notably, the Sal-B-spike...
exhibited lower overall maxima and steadier backbone RMSD tones along the 100 ns all-atom MD simulation window as compared to umifenovir (Figure 11A). With a starting RMSD value of 1.532 ± 0.46 Å, minimal RMSD fluctuation was depicted, as the MD simulation run proceed, till reaching a final RMSD value of nearly 2.871 ± 0.44 Å at the end of MD run. Except for limited fluctuations around 40 ns and 60 ns, the latter depicted dynamic behavior represents a typical well-behaved MD simulation of stabilized ligand-protein complex. On the other hand, the RMSD for umifenovir-S-protein complex managed to achieve its dynamic equilibration at later MD frames being around 60 ns from the start of the MD simulation run. Moreover, the backbone RMSD values, following the umifenovir-S-protein complex convergence,
were at higher values as compared to those of Sal-B bound complex (average RMSD 3.076 ± 0.15 Å versus 2.722 ± 0.15 Å, respectively). These findings ensure better behavior and superior stability for Sal-B within the binding interface of the spike receptor-binding domain. Further comparative analysis of the protein C-alpha RMSD trajectories illustrated lower maxima and average values for Sal-B -bound protein (3.066 ± 0.39 Å and 2.319 ± 0.17 Å versus 3.723 ± 0.09 Å and 2.604 ± 0.31 Å for N3-bound protein) ensuring a superior stability of the protein complex (Figure 11B). Despite the

---

**Figure 11.** Analysis of RMSD trajectories for the ligand-S-protein complexes throughout 100 ns all-atom MD simulation. (A) complex backbone RMSD; (B) protein C-alpha RMSD; (C) ligand backbone RMSD, versus MD simulation time in nanoseconds. Sal-B/S-protein and umifenovir/S-protein complexes are illustrated in red and blue color, respectively.

**Figure 12.** Projection of protein atoms in phase space along the first two principal eigenvectors (1 and 2). (A) Sal-B bounded S-protein; (B) Umifenovir bounded S-protein. The PCA calculations were conducted cross selected MD simulation trajectories for the first 60 ns and last 40 ns, exhibiting an expected significant structural stability and convergence.
greater stability of the Sal-B/S-protein complex, higher backbone RMSD fluctuations were depicted for the Sal-B ligand across the adopted MD simulation run (Figure 11C). The greatest fluctuations were around the 45 ns and 80 ns inferring that the ligand stability affects the complex backbone RMSD at these times. Nevertheless, the steady complex RMSD trajectory ensures the Sal-B confinement at the initial docked site of S-protein receptor-binding interface. For monitoring and validating the convergence within the conducted MD simulation, PCA calculation was performed on the MD trajectories of the first 60 ns and last 40 ns (Figure 12). The PCA analysis on the Sal-B and umifenovir-bounded S-proteins illustrated lower average trace value of the covariance matrix at the last 40 ns as compared to that along the first 60 ns MD simulation trajectories. Such dynamic behavior confers successful convergence of the bounded proteins. For Sal-B bounded protein an average trace value of 2.577 ± 0.69 nm² and 2.298 ± 0.73 nm² were assigned for the trajectories along first 60 ns and last 40 ns, respectively. Wider conformational space was covered by the S-protein at first 60 ns. Concerning the reference ligand, umifenovir, much higher values were obtained (3.891 ± 1.41 nm² at last 40 ns and 6.898 ± 4.99 nm² at first 60 ns) relative to Sal-B. The latter suggest higher stability of the Sal-B-S-protein complex as compared to umifenovir. Moreover, tighter trace values and concise data scattering were depicted for S-protein as compared to 3CLpro in case of Sal-B binding (2.298 ± 0.73 nm² versus 3.843 ± 1.12 at last 40 ns and 2.577 ± 0.69 nm² versus 6.157 ± 0.60 at initial 60 ns). Thus, preferential binding of Sal-B towards Spike protein over the protease target can be suggested.

Moving towards another stability parameter, findings for the estimated Rg trajectories came in great agreement with those of the RMSD. Both the complex and protein dynamic behavior was similar for Rg and RMSD tones where the umifenovir-S-protein complex depicted a significant rise in Rg values following the 60 ns MD simulation window (Figures 13A and B). The latter might be correlated to the loss of inter-and/or intra-molecular interactions inferring the decreased complex/protein compactness and stability within the last 40 ns time frame. On the other hand, the Sal-B-bound complex showed steady Rg trajectories with minimal fluctuation across the whole MD simulation run reaching final Rg values at 18.499 ± 0.09 Å. It worth mentioning that the ligand mass-weighted deviations from their respective common mass center were of limited fluctuations (Figure 13C). The maximum, minimum, and average Rg values of ligand, protein, and complex atom categories, listed in Table 6, illustrate the significant compactness/rigidity of the Sal-B-S-protein complex which in turn validates its stability.

Investigating the conformational flexibility of protein residues, following ligand binding, showed interesting findings. The ΔRMSF analysis was performed between the complexed SARS-CoV-2 S-protein and its apo state across the frames exhibiting significant structural stability being within the last 40 ns of the MD simulation window (Dong et al., 2018). Using the predefined ΔRMSF threshold (0.30 Å), a more
Table 6. The Rg values for investigated ligand-3CLpro complexes across the all-atom MD simulation.

| Reference atom group | Maximum (Å) | Average (Å) | Minimum (Å) | Maximum (Å) | Average (Å) | Minimum (Å) |
|----------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Complex              | 18.900 ± 0.02 | 18.499 ± 0.09 | 18.108 ± 0.03 | 18.961 ± 0.18 | 18.533 ± 0.12 | 18.131 ± 0.004 |
| Protein              | 18.885 ± 0.14 | 18.405 ± 0.09 | 18.077 ± 0.02 | 19.004 ± 0.19 | 18.546 ± 0.15 | 18.113 ± 0.02 |
| Ligand               | 5.956 ± 0.50  | 5.096 ± 0.12  | 4.647 ± 0.24  | 4.362 ± 0.07  | 3.883 ± 0.08  | 3.481 ± 0.03  |

Table 7. Calculated ΔRMSF trajectories of ligand-S-proteins along the most stabilized structural frames (last 40 ns).

| Residue ranges | Residue | Sal-B- S-protein | Umifenovir- S-protein |
|----------------|--------|------------------|-----------------------|
| Range-I        | Thr333 | 1.262 ± 0.88     | −0.286 ± 0.21         |
| Range-II       | Asp364 | 0.395 ± 0.29     | 0.027 ± 0.03          |
| Range-III      | Val367 | 0.875 ± 0.06     | 0.195 ± 0.05          |
| Range-IV       | Leu368 | −0.056 ± 0.05    | −0.040 ± 0.12         |
| Range-V        | Tyr369 | 0.344 ± 0.15     | −0.135 ± 0.05         |
| Range-II       | Ser383 | 0.315 ± 0.22     | 0.147 ± 0.01          |
| Range-III      | Pro384 | 0.032 ± 0.32     | 0.367 ± 0.09          |
| Range-IV       | Thr385 | 0.020 ± 0.39     | 0.364 ± 0.14          |
| Range-IV       | Lys386 | 0.411 ± 0.27     | 0.124 ± 0.08          |
| Range-V        | Arg408 | 0.535 ± 0.11     | −0.011 ± 0.04         |
| Range-V        | Glu409 | 0.975 ± 0.12     | −0.045 ± 0.05         |
| Range-V        | Ile410 | 0.264 ± 0.09     | 0.098 ± 0.04          |
| Range-V        | Ala411 | 0.365 ± 0.14     | −0.045 ± 0.006        |
| Range-V        | Leu517 | 0.360 ± 0.14     | 0.018 ± 0.07          |
| Range-V        | Leu518 | 0.420 ± 0.14     | 0.122 ± 0.19          |
| Range-V        | His519 | 0.361 ± 0.16     | −0.005 ± 0.20         |

*Relative difference root-mean-square fluctuation (ΔRMSF) was estimated for each ligand-bound protein relative to SARS-CoV-2-S-protein apo state. Residues exhibiting significant immobility (ΔRMSF above 0.30 Å) are only written in bold and representative ΔRMSF value is highlighted.

Figure 14. Relative ΔRMSF analysis of ligand-S-protein complexes throughout 100 ns all-atom MD simulation. Protein backbone ΔRMSF trajectories were determined from independent MD simulated SARS-CoV-2-S-protein apo state against the complexed protein with either Sal-B or N3, which were shown as a function of residue number, 333-to-527. Sal-B-S-protein and umifenovir/S-protein complexes are illustrated in red and blue color, respectively.

Figure 14. Relative ΔRMSF analysis of ligand-S-protein complexes throughout 100 ns all-atom MD simulation. Protein backbone ΔRMSF trajectories were determined from independent MD simulated SARS-CoV-2-S-protein apo state against the complexed protein with either Sal-B or N3, which were shown as a function of residue number, 333-to-527. Sal-B-S-protein and umifenovir/S-protein complexes are illustrated in red and blue color, respectively.

Overall stabilized profile for the S-protein residues were depicted upon the Sal-B binding (Figure 14). Significant residue immobility was assigned to Sal-B bound-S-protein with the residue ranges including; range-I (333-336), range-II (364-369), range-III (383-386), range-IV (408-411), and range-V (517-519), where the ligand limited the residues mobility at values above such cut-off (de Souza et al., 2019). Comparable positive ΔRMSF was depicted for the residues within the ranges-II (364-369) and -III (383-386) as they comprise the ligand-S-spike binding site suggesting relevant stability of these residues upon either Sal-B or umifenovir binding. On the contrary, the residue range-I (333-336) showed limited mobility (maximum ΔRMSF 1.262 ± 0.88 Å for Thr333). Despite being terminal-free residues with expected greater flexibility, these residues managed to adopt limited mobility which could be rationalized as being at close proximity towards the Sal-B binding site. Nevertheless, this was the opposite concerning the same residue range with bounded umifenovir, where highly negative values were obtained (ΔRMSF down to −0.286 ± 0.21 Å for Thr333) inferring great fluctuation for such residue range.

Similar differential residue fluctuation patterns were observed at the residue range-IV (408-411) and -V (517-519), where umifenovir-bound protein residues exhibit negative ΔRMSF values as compared to positive ones upon salvianolic acid binding (Table 7). As residue range-V (517-519) is located at > 20 Å distance far from the Sal-B site, it was suggested that the S-protein binding site is capable to accommodate large-sized inhibitors without influencing the site general stability. Interestingly, the other residue range-IV (408-411) is presented at close proximity towards the S-protein receptor binding interface assigned for hACE-II binding. Thus, it can be speculated that Sal-B binding could mediate conformational alterations hampering the hACE-II binding and impacting the SARS-CoV-II host-entrance. Finally, residues within the range 467-to-496 were assigned to the greatest fluctuations for both bounded S-protein, however, less negative ΔRMSF values were assigned for Sal-B-bound protein. All these observations suggest better stabilization of the Sal-B-S-protein complex as compared umifenovir.

Investigating the hydrogen bonding between the S-protein residues and Sal-B, over the 100 ns MD simulation, allow
an important understanding of the conformational changes and stability of ligand-protein complexes (Figure 15A). Notably, a single hydrogen bond pair between the ligand and Asp364 was assigned with the highest frequency (98 ± 0.67%) suggesting the validity of the proposed initial docking pose obtained through the docking study. The latter was in good agreement with ΔRMSF analysis where significant immobility was depicted for this polar ionizable residue (0.395 ± 0.29 Å). Both the Asp364 residue mainchain (O) and sidechain (OD1 and OD2) contributed within the hydrogen bonding with the terminal dimethoxyphenyl moiety-(H26 and H27) extended from the C3 of the Sal-B benzodihydrofuran scaffold. However, the residue sidechain showed steadier trajectories and higher frequency (97 ± 0.21% versus 2 ± 0.96%) at hydrogen bond distances below the depicted cut-off (< 3.0 Å). This indicates the superiority of the Asp364 sidechain over its mainchain for Sal-B-S-protein complex stability. It worth noting that over certain time intervals; around 12 ns, 20 ns, and 35-65 ns, the hydrogen bonding with either the Asp364 sidechain or mainchain was lost suggesting significant alteration of the Sal-B conformation within the target pocket around these timelines. Following Asp364 residue, minimal hydrogen bonding contributions were assigned for Asn343, Glu340, and Ser372 at occupancy percentages of 17 ± 1.001%, 1 ± 0.99%, and 4 ± 0.75%, respectively. Such humble hydrogen bond frequencies present the minimal contributions of these residues within the Sal-B stability within S-protein pocket. Similarly, limited hydrogen bond frequencies were assigned for Asn343 and Ser373 with the ester linker of umifenovir (2 ± 1.92% and 3 ± 0.23%, respectively) (Figure 15B). The latter polar residues depicted favored short-range hydrogen bonding (< 4.0 Å) only within the initial MD simulation frames (5-20 ns). Following this timeline, the hydrogen bonds were lost where the ligand adopted another favored conformation where the ester linker was kept at an average distance of 5.746 ± 1.90 Å and 8.321 ± 1.59 Å from Asn343 and Ser373, respectively. Nevertheless, both humble hydrogen bondings were of the highest frequency among a limited number of hydrogen bonds being depicted between the S-protein residues and umifenovir across the 100-ns MD simulation window. Therefore, it was suggested that umifenovir was maintained within the NAG-S-protein binding site majorly through hydrophobic interactions rather than through polar ones. In these regards, analysis of the distances between the umifenovir hydrophobic moieties and the C-alpha of relevant surrounding non-polar residues has proceeded. Interestingly, both Val367 and Phe342 depicted the
closest distances and favored hydrophobic contacts with the umifenovir fused benzene ring across the MD simulation run. Despite limited fluctuations, the non-polar residue, Val367, showed the steadiest hydrophobic distance (average 5.594 ± 1.19 Å) trajectories along the 100-ns simulation run. While as, the aromatic residue, Phe342, illustrated steady tones around 5.172 ± 1.38 Å which were only beyond the 45 ns time frames. On the other hand, hydrophobic interactions between either Trp436 or Leu368 and dimethylamino group, the closest hydrophobic scaffold of umifenovir, were assigned insignificant since bond distance trajectories were at far distances with average values more than 6 Å (9.551 ± 3.09 Å and 13.210 ± 2.27 Å, respectively). The latter observations illustrated the superior contribution of hydrophobic interactions for the stability of umifenovir within the S-protein NAG-binding site. It worth mentioning that Trp441

Figure 16. Conformations of ligand-protein complex at SARS-CoV-2 S-protein NAG-binding site through selected trajectories. (A) Sal-B; (B) umifenovir control. Protein is represented in green or red cartoon 3D-representation corresponding to initial (0 ns) and last (100 ns) extracted trajectories, respectively. The key binding residues (lines), ligands (sticks), and hydrophilic interactions (hydrogen bonding; dashed lines), are all presented in colors corresponding to their respective extracted trajectory.
showed favored hydrophobic contacts (average $5.622 \pm 0.43 \text{Å}$) only at the initial frames (5-20 ns) suggesting a comparable contribution of polar by Asn343/Ser373 and non-polar Trp441 for umifenovir-S-protein complex stability only at initial MD simulation frames.

To further analyze the key structural alterations for ligand-protein complex, selected frames of each ligand-protein system were extracted and minimized to a gradient of 0.001 Kcal.mol$^{-1}$A$^{-2}$ using MOE package. For both ligand-S-protein complexes, the comparative conformations analysis was performed at 0 and 100 ns of the MD simulation run. As illustrated, a relevant molecular shift was observed for Sal-B (Figure 16A). Nevertheless, polar interaction with Asp364 and terminal dimethoxyphenyl scaffold was maintained as being represented previously through hydrogen bond analysis (Figure 15A). On the other hand, the initial hydrogen bond pair between the other dimethoxyphenyl moiety of Sal-B and Ala344 was lost. However, the same distal ligand scaffold was further stabilized at the end of MD simulation via a vicinal hydrophilic residue, Asp343, thus maintaining the Sal-B at comparable orientation within the NAG-S-protein pocket relative to the initial docking pose. Regarding the umifenovir-S-protein complex, a more significant orientation shift was depicted for the ligand (Figure 16B). The ligand ended with a more stabilized orientation at close proximity to several hydrophobic pocket residues including, Phe342, Leu368, Val367, and Trp436. This was expected since the previous hydrogen bond analysis depicts limited polar interactions between the umifenovir and target residues. As illustrated in the represented figure, a favoured π-π stacking between Phe342 and the ligand dihydrobenzofuran ring mediated the stability of the ligand-protein complex at the end of MD simulation run. Moreover, non-polar interactions between the ligand aromatic arm and both Leu368 and Val367 were illustrated as relevant for stabilizing the umifenovir ligand at the 100 ns timeline. Based on the above findings, it was clear that polar ligand-protein interactions greatly govern the Sal-B bonding to the S-protein NAG pocket. Nonetheless, the positive control reference, unifenovir, was maintained within the S-protein binding site mainly through hydrophobic contacts.

3.3.3. Binding-free energy calculations

In an attempt to understand the nature of the ligand-protein binding, explore the comparative ligand-binding site affinity, and obtain more information concerning individual ligand/residue contribution, the calculation of the binding-free energy was proceeded (Cavasotto, 2020). The MD-based Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) approach was adopted for the designated binding-free energy calculations, using "g_mmpbsa" tool on GROMACS. The approach accounts for more accurate ligand-protein affinity as compared to the most sophisticated flexible molecular docking technique (Kumari et al., 2014). Generally, MM/PBSA estimates binding-free energy as a contribution of several energy terms through these given Equations (Kumari et al., 2014):

$$
\Delta G_{\text{binding}} = G_{\text{complex}} - (G_{\text{ligand}} + G_{\text{protein}})
$$

$$
G_X = (E_{\text{MM}}) - T S + G_{\text{solvation}}
$$

$$
E_{\text{MM}} = E_{\text{bonded}} + (E_{\text{vdW}} + E_{\text{electrostatic}})
$$

$$
G_{\text{solvation}} = G_{\text{polar}} + \gamma SASA + b
$$

where $\Delta G_{\text{binding}}$ is the binding-free energy correlating to ligand-protein binding where the higher negative energy values infer greater protein-ligand affinity. The energy term; $G_{\text{complex}}$, $G_{\text{protein}}$, and $G_{\text{ligand}}$ are respective total free energies of ligand–protein complex, isolated protein, and isolated ligand in solvent. Vacuum MM potential energy ($E_{\text{MM}}$) together with the entropic contribution to free energy (TS) and free energy of solvation ($G_{\text{solvation}}$) provided the total free energy of protein, ligand, or ligand–protein complex ($E_X$). Terms T and S denote temperature and entropy, respectively, while $E_{\text{MM}}$ was calculated based on molecular mechanics force-field parameters. Using the solvent-accessible surface-area (SASA)-Nonpolar Model, the $G_{\text{solvation}}$ energy term comprises polar and non-polar parts, where the latter was estimated via SASA and fitting constant (b). Finally, $G_{\text{polar}}$ is to be solved from the Poisson-Boltzmann equation.

Typically, the binding-free energy should be estimated from the MD simulation trajectories depicting stabilized protein-ligand systems. Thus, representative frames were extracted and saved from the last 40 ns to be enrolled within the calculation of each energy term. Adopting this specific timeframe was rationalized by the above RMSD and Rs analysis where equilibrated plateau tones were illustrated within the 60-to-100 ns timeframe interval. Notably, the $\Delta G_{\text{binding}}$ of N3-3CLpro complex was estimated at higher negative values as compared to that of Sal-B complex (-76.046 ± 28.005 kJ.mol$^{-1}$ versus -58.590 ± 23.559 kJ.mol$^{-1}$, respectively) (Table 8). Considering the reported superior inhibition activity of N3 ligand against SARS-CoV-2 3CLpro (Jin et al., 2020), the obtained data within the presented study highlights the potential activity of Sal-B against the same target enzyme. Dissecting the furnished $\Delta G_{\text{binding}}$ for both investigated complexes showed a preferential contribution of the hydrophobic an der Waal interactions as compared to that of the electrostatic energy term. However, the significant occupancy of the depicted hydrogen-bond interaction analysis can suggest a somewhat balanced contribution between both energy terms. Interestingly, higher $\Delta G_{\text{electrostatic}}$ contribution was assigned for the Sal-B as compared to N3, suggesting the significant role of the polar sugar moieties for ligand anchoring within 3CLpro binding site. It worth mentioning that presence of sugar moiety in Sal-B could act as a double-bladed influencer on ligand-protein binding. Despite its suggested role in anchoring the ligand through a polar hydrogen bond with protein residues, these polar sugar scaffolds impose higher $\Delta G_{\text{solvation}}$ impacting the Sal-B-protein binding since such a process is a solvent-substitution approach.

Regarding the ligand-S-protein complexes, there was a comparable pattern of energy contribution for the $\Delta G_{\text{electrostatic}}$ and $\Delta G_{\text{van der Waal}}$ interactions within both investigated ligand-S-protein complexes as compared to those of 3CLpro ones (Table 8). The highest energy contribution was
assigned for the van der Waal interactions over the electrostatic ones. However, the $\Delta G_{\text{electrostatic}}$ energy term was much more inferior for the umifenovir as compared to that of Sal-B. The latter highlights the role of the polar sugar moieties for ligand anchoring within the S-protein binding site. Moreover, the low electrostatic contribution for umifenovir-S-protein binding energy came in great agreement with the above bonding analysis where hydrogen bonding

| Energy terms                  | N3-3Clpro (60-100 ns) | Sal-B-3Clpro (60-100 ns) | Umifenovir-S-protein (60-100 ns) | Sal-B-S-protein (60-100 ns) |
|------------------------------|-----------------------|-------------------------|---------------------------------|-----------------------------|
| $\Delta G_{\text{Van der Waal}}$ | $-218.104 \pm 20.751$ | $-180.787 \pm 14.989$  | $-173.541 \pm 10.185$          | $-127.621 \pm 16.863$     |
| $\Delta G_{\text{Electrostatic}}$ | $-52.058 \pm 16.911$  | $-83.655 \pm 16.923$  | $-6.594 \pm 5.191$             | $48.224 \pm 21.238$       |
| $\Delta G_{\text{Solvation (Polar)}}$ | $219.152 \pm 16.009$  | $228.284 \pm 23.370$  | $72.800 \pm 10.303$            | $142.675 \pm 18.777$      |
| $\Delta G_{\text{Solvation (SASA)}}$ | $-22.432 \pm 1.330$   | $-22.432 \pm 1.330$   | $-18.233 \pm 1.148$            | $-15.532 \pm 1.368$       |
| $\Delta G_{\text{Binding}}$       | $-76.046 \pm 28.005$  | $-58.590 \pm 23.559$  | $-125.567 \pm 28.005$          | $-48.702 \pm 26.695$      |

Figure 17. Binding-free energy/residue decomposition illustrating the protein residue contribution at ligand-protein complex $\Delta G_{\text{binding}}$ calculation. Sal-B/3CLpro and N3/3CLpro residue-wise energy contributions are represented in red and blue colors, respectively. Lower panels are expanded versions of three designated residue regions (Ser1-Ala70, His80-Phe150, and Phe150-Tyr209) of the upper panel.
between umifenovir and S-protein residues were limited as well as of minimal frequencies/occupancies. Again, the presence of sugar moiety in Sal-B could act as a double-bladed influencer on the ligand-protein binding as these polar sugar scaffolds impose higher $\Delta G_{\text{solvation}}$ influencing the Sal-B-protein binding since such process is a solvent-substitution approach. This could reason why umifenovir furnished higher $\Delta G_{\text{binding}}$ as compared to that of Sal-B (-125.567 ± 28.005 kJ.mol⁻¹ versus -48.702 ± 26.695 kJ.mol⁻¹, respectively) since the earlier furnished minimal solvation energy. It is worth mentioning that despite the high $\Delta G_{\text{solvation}}$ energy term for Sal-B, binding of Sal-B with either 3CLpro or S-protein was confirmed favoured and stable for the advent of the high electrostatic and van der Waal energy contribution afforded by the Sal-B scaffold.

For identifying the critical residues involved within the binding of investigated ligands with 3CLpro protein, the residue-wise energy contribution to the obtained $\Delta G_{\text{binding}}$ was also estimated using g_mmpbsa (Figure 17) (Kumari et al., 2014). Several key residues, which have participated within the initial docked complexes, showed significant contributions within the calculated $\Delta G_{\text{binding}}$ of the investigated complexes. Regarding Sal-B, residues of S1 and S1’ subsites showed the greatest share within the total ligand binding-free energy. Residues includes His41, Leu141, Asn142, Gly143, and Cys145 exhibited significant residue-binding energy contribution confirming the key role of these residue in small molecule binding to SARS-CoV-2 3CLpro protein. Only Met165 (subsites S2) and Arg188 (subsites S3) illustrated significant residue-wise energy contribution with Met165 with the highest contribution of all residues. Significant energy contribution was further assigned to the S1’ subsite vicinal residue, Thr25-to-Leu27, which came in great agreement with the new adopted orientation of Sal-B arm at the end of the MD simulation (Figure 10A). The above findings infer a preferential anchoring of Sal-B at S1 and S1’ subsites. For the potent inhibitor, N3, differential pattern of residue-wise energy contribution has been depicted. There was a major residue-contribution share for the hydrophobic residues over those of the polar ones. This was expected since the ligand exhibited superior contribution of the $\Delta G_{\text{van der Waal}}$ over that of the $\Delta G_{\text{electrostatic}}$ ones (Table 8). Notably, the highest contribution was for the non-polar residue Phe342, Val367, and Leu368 (-15.685 ± 1.753, −9.132 ± 1.764, and −4.619 ± 0.637 kJ.mol⁻¹, respectively) which was in good agreement with the above hydrophobic bond distance analysis (Figure 16B). Other significant hydrophobic residues include Phe374, Ile434, and Trp366 with energy contribution values ranging from −3.348 ± 2.652 to −8.050 ± 4.554 kJ.mol⁻¹. On the other hand, the negative-charged polar residue, Asp364, showed significant energy contribution within the umifenovir-S-protein complex (−9.132 ± 1.764 kJ.mol⁻¹). The latter infers the important role of Asp364 for anchoring both the Sal-B and umifenovir within the S-protein NAG-binding site.

3.4. Quantum mechanics study

Sal-B is the major component of Danshen (Chinese sage) and studies in the last few decades disclosed its broad pharmacological and therapeutic activities (Jing et al., 2016; Li et al., 2020; Ma et al., 2019; Qiao et al., 2019; Su et al., 2020; Wang et al., 2018; Zhao et al., 2020; Zhu et al., 2018). Sal-B was firstly discovered in 1981 and its configurational structure was initially determined which was later corrected upon performing further geometrical studies using experimental tools (Gong et al., 2015). Therefore, we herein utilize theoretical computations to allocate the most energetically stabilized geometries of Sal-B.

A potential energy surface (PES) scan of Sal-B was performed by varying a dihedral angle rotation from 0° to 360° in 10° stepwise rotation while allowing the rest of the molecule to undergo free iterations. As can be seen in Figure 1, four dihedral angles were selected where two come out from furan moiety (scan-1 and scan-2), another one projected out from the benzene moiety (scan-3), and the last one originated from the catechol moiety (scan-4). Diagrams of all performed PES scans are represented in Figure 19 where a big energy difference between conformational rotamers of Sal-B was noticed in all PES scans except for scan-4.

In scan-1, three troughs were obtained where structures 12, 20, and 31 are more energetically stabilized than the initially optimized geometry of Sal-B (structure 2). Structure 31 was found the most energetically stable form of Sal-B exhibiting an energy gap of approximately 10 kcal/mole with the least stable rotamer of Sal-B. Regarding scan-2, a bigger energy gap (~16 kcal/mole) was detected between the least and most energetically stabilized geometries of Sal-B.
Structure 16 was found more stabilized than structures 2 and 33 as can be seen in Figure 19. Surprisingly, scan-4 showed the biggest observed discrepancies between Sal-B conformers. The energy gap was estimated at ≈35 kcal/mole. Three geometries of Sal-B (Structure 1, 12 & 32) had nearly equal energies which are ≈15 kcal/mole higher than that of structure 37. Structure 37 is also the structure possessing the least energy (the most stable structure) among all other structures obtained from the four PES scans. By looking at a diagram of scan-4, it is obvious that rotamers of Sal-B are somehow equally stabilized, compared to other scans, with an energy gap of only ≈4 kcal/mole as can be seen in Figure 19. The calculated energies in Hartree of Sal-B geometries obtained at the troughs of four scans are also depicted in Table 9. Geometrical coordinates of Sal-B structures obtained from troughs of four scans are listed in the Supplementary Materials. Apart from defining the possible conformations of Sal-B, we also aimed to calculate some of the physicochemical properties while in an excited state. UV-Vis spectrum of Sal-B...
reported in the literature is an unstructured spectrum which can be ascribed to the presence of more than one rotamer of Sal-B in a solution. In our study, the computed UV-Vis spectrum of Sal-B showed two prominent peaks at 332 nm and 297 nm along with absorption shoulder at approximately

Table 9. Computed energies (in Hartree) of Sal-B structure obtained from the troughs of potential energy surface (PES) scan at B3LYP/6-31G level of theory. Values in bold refer to the most stabilized geometry obtained from each PES scan.

| Scan-1 | Scan-2 | Scan-3 | Scan-4 |
|--------|--------|--------|--------|
| Iteration No. | 2 | 2 | 1 | 1 |
| Energy (a.u.) | -2592.695351 | -2592.695364 | -2592.695385 | -2592.695414 |
| Iteration No. | 12 | 16 | 12 | 21 |
| Energy (a.u.) | -2592.696108 | -2592.717882 | -2592.696108 | -2592.696108 |
| Iteration No. | 20 | 33 | 32 | 37 |
| Energy (a.u.) | -2592.696013 | -2592.709558 | -2592.697867 | -2592.697286 |
| Iteration No. | 31 | 33 | 37 | |
| Energy (a.u.) | -2592.706023 | -2592.709558 | -2592.720085 | -2592.696631 |

Table 10. Absorption energy, oscillator strength, and electronic transitions of the main absorption bands of Sal-B obtained using B3LYP/6-311+G* model.

| Absorption energy | Oscillator strength | Major contributions |
|-------------------|--------------------|---------------------|
| 332 nm | 0.2343 | H-2 LUMO (83%) |
| | | H-4 LUMO (8%) |
| 297 nm | 0.2612 | H-4 LUMO (84%) |
| | | H-2 LUMO (8%) |

Figure 19. Computed potential energy surface (PES) scans of Sal-B structure using B3LYP/6-31G model. The flexible scan was performed at four different sites constituting four different dihedral angles (marked in red) as illustrated in Figure 1.

Figure 20. Calculated UV-Vis absorption (excitations) spectrum of Sal-B and the electronic circular dichroism (ECD) of Sal-B using B3LYP/6-311 + G* model.
250 nm, as depicted in Figure 20. The 332 nm band has a higher intensity (oscillator strength) than that of the peak calculated at 297 nm.

The electronic transitions responsible for the absorption peaks of Sal-B at 332 nm and 297 nm are the same but contributing unequally (different ratios) as listed in Table 10. The electronic transition from the highest occupied molecular orbital —2 (H-2) to the lowest unoccupied molecular orbital (LUMO) contributed by 83% to the 332 nm absorption band. While H-4 to LUMO electronic transition contributed only by 8%. The 297 nm peak has an opposite trend where the H-4 to LUMO transition contributed more significantly than H-2 to LUMO transition as can be read from Table 10.

Electron density of the outermost molecular orbitals of Sal-B were also computed at B3LYP/6-311G* level of theory. Representations of the electronic charge distribution of each molecular orbital are depicted in Table 11. Referring to the diagram of HOMO-4 and HOMO-2, it is obvious that the electron density is localized over the benzo-dihydrofuran nucleus and the aliphatic side chain projected from the benzene ring. While the electronic density distribution over
LUMO did not change significantly except for being contracted from the dihydrofuran moiety suggesting that the involvement of dihydrofuran moiety in binding interactions with target proteins. Therefore, it is postulated that the dihydrofuran moiety can act as a hydrogen-bond acceptor group within Sal-B.

The molecular electrostatic potential (MEP) map of Sal-B is also depicted in Table 11. MEP diagram gives a whole representation of electronic density distribution at a molecular level. The red areas represent areas with the most intense electronegativity while the blue region represents the most electron deficient areas. By referring to MEP of Sal-B deposited in Table 11, it is obvious that the molecule is rich in sites of intense electronegativity and electropositivity which suggests its potential to engage with the formation of several hydrogen bonds with the target protein. This is also consistent with our molecular docking studies where Sal-B was found to form six H-bonds which may be the cause for enhancing its calculated binding affinities and energies and make Sal-B superior to other studies NSAIDs.

3.5. Physicochemical, ADME and pharmacokinetic properties prediction

SwissADME online web tool provided by the Swiss Institute of Bioinformatics (SIB) is utilized for calculation of the physicochemical properties and prediction of the pharmacokinetic properties of the isolated medicinal plant products and reference controls 1-11 as illustrated in Table 12. This was carried out to assure that the most active medicinal plant product in molecular modeling studies, Sal-B 6, is a promising candidate in terms of molecular docking efficacy, and pharmacokinetic characteristics. Sal-B exhibits a predicted wlogP value of 2.9 with no blood-brain barrier (BBB) permeability and consequently no predicted CNS adverse effects. However, Sal-B shows low gastrointestinal tract (GIT) absorption with poor water solubility, therefore, a prodrug approach may be utilized to increase its water solubility or it may be administrated by injection. Moreover, the submitted Sal-B is not a P-glycoprotein substrate (Pgp-), whereas N3 is a substrate for P-glycoprotein, so at this point Sal-B is better than N3 as Sal-B not amenable to the efflux mechanism done by this transporter which is used by many tumor cell lines as a drug-resistance mechanism. In contrast to umifenovir, Sal-B did not manifest any inhibition for the metabolizing enzymes (CYP 1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4).

4. Conclusion

Owing to their previously known activity against SARS-CoV, nine traditional Chinese medicinal plant compounds were virtually screened by molecular docking to reveal their ability to bind to the active site of two key SARS-CoV-2 target proteins; 3CLpro, and S-protein. The study confirmed the affinities of screened compounds to these two key SARS-CoV-2 target proteins. Sal-B 6 showed the most promising binding energy scores among other screened compounds as well as significant interactions with the key target residues. The promising affinity of Sal-B to SARS-CoV-2 targets provoked the all-atom MD simulation study. Through 100 ns MD simulation runs, Sal-B illustrated great thermodynamic stability and confinement within the binding sites of both COVID-19 target proteins. Hence, the enhanced binding interactions between Sal-B and target proteins can be attributed to the structural flexibility of Sal-B as evidenced by the quantum mechanics calculations. Based on the above evidence, Sal-B arose as a promising lead compound for future biological investigations aiming to provide a more effective pharmacotherapeutic drugs active against COVID-19 pandemic. Meanwhile, the study highlights the SAR information required for targeting COVID-19 pivotal biological targets. This would be of great benefit in facilitating the future development of new candidates with potential activity against the COVID-19 virus. Moreover, the performed ADME study revealed reasonable pharmacokinetic properties of Sal-B and fortunately with some benefits over reference standards; N3 and umifenovir.

Acknowledgements

The authors acknowledge Swinburne University of Technology, Melbourne, Australia, for using Swinburne’s supercomputing facilities to perform the Quantum mechanics calculations details. The MD simulations within this work were performed at King Abdulaziz University’s High Performance Computing Center (Aziz Supercomputer) (http://hpc.kau.edu.sa”), Jeddah, Saudi Arabia and the authors, therefore, acknowledge with thanks the center for technical support.
Disclosure statement
No potential conflict of interest was reported by the authors.

ORCID
Sameh S. Elhady http://orcid.org/0000-0003-3799-0581

References
Aier, I., Varadwaj, P. K., & Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Scientific Reports*, 6, 1–10. https://doi.org/10.1038/srep34984
Aktas, A., Tuzun, B., Taskin, A. H., Sayin, K., & Ataseven, H. (2020). How do arbidol and its analogs inhibit the SARS-CoV-2 *Bratislava Medical Journal*, 121(10), 705–711. https://doi.org/10.4149/BML_2020_115
Al-Karmalawy, A. A., Alnajjar, R., Dahab, M. M., Metwaly, A. M., & Eissa, I. H. (2021). Molecular docking and dynamics simulations reveal the potential of anti-HCV drugs to inhibit COVID-19 main protease. *Pharmaceutical Sciences*. https://doi.org/10.34172/PS.2021.13
Belouzard, S., Millet, J. K., Licitra, B. N., & Whittaker, G. R. (2012). Dendritic cell responses to the SARS-CoV-2 spike protein. *Journal of Virology*, 86, 705–715. https://doi.org/10.1128/JVI.06377-11
Buhner, S. (2013). *Aier, I., Varadwaj, P. K., & Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. Scientific Reports, 6, 1–10. https://doi.org/10.1038/srep34984*

Aktas, A., Tuzun, B., Taskin, A. H., Sayin, K., & Ataseven, H. (2020). How do arbidol and its analogs inhibit the SARS-CoV-2 *Bratislava Medical Journal*, 121(10), 705–711. https://doi.org/10.4149/BML_2020_115
Al-Karmalawy, A. A., Alnajjar, R., Dahab, M. M., Metwaly, A. M., & Eissa, I. H. (2021). Molecular docking and dynamics simulations reveal the potential of anti-HCV drugs to inhibit COVID-19 main protease. *Pharmaceutical Sciences*. https://doi.org/10.34172/PS.2021.13
Belouzard, S., Millet, J. K., Licitra, B. N., & Whittaker, G. R. (2012). Dendritic cell responses to the SARS-CoV-2 spike protein. *Journal of Virology*, 86, 705–715. https://doi.org/10.1128/JVI.06377-11
Buhner, S. (2013). *Aier, I., Varadwaj, P. K., & Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. Scientific Reports, 6, 1–10. https://doi.org/10.1038/srep34984*
