Discovery of putative inhibitors against main drivers of SARS-CoV-2 infection: Insight from quantum mechanical evaluation and molecular modeling

Toheeb A. Balogun1*, Onyeka S. Chukwudozie2, Uchechukwu C. Ogbodo3, Idris O. Junaid4, Olugbodi A. Sunday5, Oluwasegun M. Ige6, Abdullahi T. Aborode7, Abiola D. Akintayo8, Emmanuel A. Oluwarotimi9, Isaac O. Oluwafemi10, Oluwatosin A. Saibu11, Prosper Chuckwuemaka12, Damilola A. Omoboyowa13, Abdullahi O. Alausa14, Nkechi H. Atasie15, Ayooluwa Ilesanmi16, Gbenga Dairo17, Zainab A. Tiamiyu18, Gaber E. Batiha19, Afrah Fahad Alkhuriji20, Wafa Abdullah I. Al-Megrin21, Michel De Waard22,23

1Department of Biochemistry, Adekunle Ajasin University, Akungba-Akoko, Nigeria, 2Department of Biological Sciences, University of California, San Diego, San Diego, CA, United States, 3Department of Applied Biochemistry, Nnamdi Azikiwe University, Awka, Nigeria, 4Department of Chemistry and Chemical Biology, Stevens Institute of Technology, Hoboken, NJ, United States, 5Department of Environmental Toxicology, Universitat Duisburg-Essen, Essen, Germany, 6Department of Marine Biological Resources, Ghent University, Ghent, Belgium, 7Department of Chemistry, Mississippi State University, Starkville, MS, United States, 8Department of Chemistry, University of Texas at Dallas, Richardson, TX, United States, 9Department of Chemistry, Missouri University of Science and Technology, Rolla, MO, United States, 10Department of Chemistry, Adekunle Ajasin University, Akungba-Akoko, Nigeria, 11Department of Biotechnology, Federal University of Technology University Akure, Akure, Nigeria, 12Department of Biochemistry, Ladoke Akintola University, Ogbomoso, Nigeria, 13Clinical Pharmacy Department, Nigeria Correctional Service, Enugu Custodial Centre, Enugu, Nigeria, 14Department of Chemistry, Mississipi University for Women Columbus, Columbus, United States, 15Department of Biological Sciences, Western Illinois University, Macomb, IL, United States, 16Department of Biochemistry and Molecular Biology, Federal University Dutse-ma, Dutse-ma, Nigeria, 17Department of Pharmacology and Therapeutics, Faculty of Veterinary Medicine, Damanhour University, Damanhour, Egypt, 18Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia, 19Department of Biology, College of Science, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, 20Smartox Biotechnology, Saint-Égrève, France, 21L’Institut du Thorax, INSERM, CNRS, Université de Nantes, Nantes, France, 22LabEx Ion Channels, Science and Therapeutics, Université de Nice Sophia-Antipolis, Valbonne, France, 23Institut de Neurophysiopathologie (INP), Faculté des Sciences Médicales et Paramédicales, Aix-Marseille Université, CNRS UMR 7051, Marseille, France

SARS-CoV-2 triggered a worldwide medical crisis, affecting the world’s social, emotional, physical, and economic equilibrium. However, treatment choices and targets for finding a solution to COVID-19’s threat are becoming limited. A viable approach to combating the threat of COVID-19 is by unraveling newer pharmacological and therapeutic targets pertinent in the viral survival and adaptive mechanisms within the host biological milieu which in turn...
provides the opportunity to discover promising inhibitors against COVID-19. Therefore, using high-throughput virtual screening, manually curated compounds library from some medicinal plants were screened against four main drivers of SARS-CoV-2 (spike glycoprotein, PLpro, 3CLpro, and RdRp). In addition, molecular docking, Prime MM/GBSA (molecular mechanics/generalized Born surface area) analysis, molecular dynamics (MD) simulation, and drug-likeness screening were performed to identify potential phytodrugs candidates for COVID-19 treatment. In support of these approaches, we used a series of computational modeling approaches to develop therapeutic agents against COVID-19. Out of the screened compounds against the selected SARS-CoV-2 therapeutic targets, only compounds with no violations of Lipinski’s rule of five and high binding affinity were considered as potential anti-COVID-19 drugs. However, lonchocarpol A, diplacol, and broussonol E (lead compounds) were recorded as the best compounds that satisfied this requirement, and they demonstrated their highest binding affinity against 3CLpro. Therefore, the 3CLpro target and the three lead compounds were selected for further analysis. Through protein–ligand mapping and interaction profiling, the three lead compounds formed essential interactions such as hydrogen bonds and hydrophobic interactions with amino acid residues at the binding pocket of 3CLpro. The key amino acid residues at the 3CLpro active site participating in the hydrophobic and polar inter/intra molecular interaction were TYR54, PRO52, CY544, MET49, MET165, CYS145, HIS41, THR26, THR25, GLN189, and THR190. The compounds demonstrated stable protein–ligand complexes in the active site of the target (3CLpro) over a 100 ns simulation period with stable protein–ligand trajectories. Drug-likeness screening shows that the compounds are druggable molecules, and the toxicity descriptors established that the compounds demonstrated a good biosafety profile. Furthermore, the compounds were chemically reactive with promising molecular electron potential properties. Collectively, we propose that the discovered lead compounds may open the way for establishing phytodrugs to manage COVID-19 pandemics and new chemical libraries to prevent COVID-19 entry into the host based on the findings of this computational investigation.

KEYWORDS
SARS-CoV-2, COVID-19, 3CLpro, PLpro, RNA-dependent RNA polymerase (RdRp), molecular modeling, glycoprotein

1 Introduction
With the emergence of different variations of the coronavirus disease 2019 (COVID-19), such as alpha, delta, and omicron, COVID-19 remains a global challenge to health and the economy due to the unexpected emergence of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). COVID-19 is an agile respiratory disease caused by a novel coronavirus first reported in Wuhan, China, in December 2019 and declared a global pandemic by the World Health Organization (WHO) in March 2020 (Chan et al., 2020). Coronaviruses are positive sensed, linear single-stranded RNA viruses which comprise nucleoproteins (N), envelope proteins (E), matrix proteins (M), spike proteins (S), and many non-structural proteins (Luk et al., 2019). The length of the RNA genomes ranges from 26–32 kb, which contains 12 open reading frames. A structural analysis of the SARS-CoV-2 genomes using biophysical and modeling techniques reveals that two polyproteins divided into 15 or 16 non-structural proteins make up the first two-thirds of the coronavirus genome (Xu et al., 2020). A phylogenetic analysis shows that the remaining ORFs contains the genetic makeup of four important structural proteins: envelop (E), membrane (M), nucleocapsid (N), and spike (S) proteins. These proteins played a significant role in the virus survival and replication and have received intense interest from several investigators to develop inhibitors of SARS-CoV-2 (Adedeji et al., 2012; Yu et al., 2012; Jia et al., 2019). Notably, COVID-19 has been associated with an effect on multiple vital organs and the central nervous system and can result in respiratory problems with fatal consequences (Seah and
Agrawal, 2020). Previous reports indicated that SAR-CoV-2 therapeutic targets include receptor binding of glycosylated spike (S) protein, which mediates host cell receptor recognition and host cell entry and induces host immune responses, and non-structural proteins such as RNA-dependent RNA polymerase (RdRp), CoV main protease [Mopar; also known as 3-chymotrypsin–like protease (3CLpro)], and papain-like protease (PLpro) (Liu et al., 2020). Theoretically, drugs competing with RBD for receptor binding sites can inhibit viral entry and replication. Furthermore, the non-structural proteins (PLpro, 3CLpro, and RdRp) play a vital role in proteolysis and viral polyprotein processing. Thus, NSPs had recently emerged as an essential therapeutic biomarker for the design of COVID-19 drug candidates (Chen et al., 2020).

The therapeutic management of COVID-19 involves two target selection approaches. One of the methods is boosting the human immune system or human cells using an attenuated vaccine, whereas the second technique involves inhibiting molecular targets by small molecule inhibitors. Regarding the vaccine, whereas the second technique involves inhibiting human immune system or human cells using an attenuated target selection approaches. One of the methods is boosting the design of COVID-19 drug candidates (Chen et al., 2020).

Scientists have harnessed several strategies for developing novel drugs against COVID-19 (Zumla et al., 2016). The first strategy was to screen existing broad-spectrum anti-viral drugs such as ribavirin and cyclophillin. This approach is advantageous because the pharmacokinetic profile of the anti-viral drugs and their associated side effects have been clearly stated. One of the disadvantages of broad-spectrum anti-viral drugs is their non-specificity, which might, in turn, result in low potency against coronavirus (Chan et al., 2013; de Wilde et al., 2014). The second techniques involve screening molecular databases such as the ZINC database to identify potential anti-coronavirus compounds via high-throughput virtual screening. This approach has been used to develop biologically active compounds, lopinavir/ritonavir, as anti-HIV agents (de Wilde et al., 2014). The third approach is based on analyzing genomic datasets to develop new targeted drug candidates from scratch for precision medicine (Dyall et al., 2014). Therapeutic agents developed against coronavirus via the third strategy exhibits promising pharmacological potential. However, this approach’s long-term process and expenses are major limiting factors.

Herbal medicine has been an alternative medicine since time immemorial in managing various diseases and may be an important source of anti-coronavirus agents (Ling, 2020). An earlier systematic study in 2003 shows that patients infected with SARS-CoV-2 treated with traditional Chinese medicine (TCM) were reported to have short time hospitalization, reduced steroid side effects, and improvements from the viral symptoms (World Health Organization, 2004). Therefore, a significant amount of research has focused on developing therapeutic agents against coronavirus from TCM, ethnobotanical herbs, and dietary supplements (Dudani and Saraogi, 2020; Vellingiri et al., 2020). In vivo, in vitro, and in silico studies have revealed the antiviral potential of numerous bioactive compounds against coronavirus. Some phytocompounds and sources include glycyrrhizin isolated from Glycyrrhiza glabra L. (licorice); tetra-O-galloyl-β-D-glucose (TGG) and luteolin, isolated from Rhus chinensis Mill. and Veronica linariifolia Pall. ex Link; and aurantiamide acetate derived from Artemisia annua L. plant. Several plants such as Sanguisorba officinalis L., Stephania tetrandra S. Moore, and Strobilanthes cusia (Nees) Kuntze have been reported for their antiviral potential toward RNA and protein synthesis of the coronavirus (Wu et al., 2004; Wang et al., 2007; Kim et al., 2010). In this study, manually curated compounds library from some medicinal plants were screened against four main targetable drivers of SARS-CoV-2 (spike glycoprotein, PLpro, 3CLpro, and RdRp) using high-throughput virtual screening. Only compounds with no violations of the Lipinski’s rule of five with at least a binding energy of ~5.0 kcal/mol against the four targets were considered potential drug candidates with therapeutic effects against COVID-19. Herein, lonchocarpol A, diplocal, and broussonol E (lead compounds) were recorded as the best compounds that satisfied this requirement and demonstrated their highest binding affinity against 3CLpro. Therefore, the 3CLpro target and the three lead compounds were selected for further analysis, such as molecular dynamics simulation and quantum mechanical evaluation. Overall, this study serves as a benchmark for developing the discovered three lead compounds as COVID-19 therapeutic agents.

2 Materials and methods

2.1 Quantum mechanical calculation

2.1.1 Molecular docking studies

Theoretical approaches used to compute compounds’ chemical and biological activities have been well documented. The quantum chemical (QM) calculation was executed using the MOPAC 2016 software program (Stewart, 1990). PM7 semi-empirical Hamiltonian incorporating an implicit COSMO solvation model was used to perform the calculation (Klamt and Schuurmann, 1993; Dutra et al., 2013). Notably, geometric pre-optimization of the top four inhibitors was carried out using the molecular mechanics force field (MMFF94) integrated with the Avogadro v 1.2.0 software program (Hanwell et al., 2012), coupled with chemical structure protonation at a pH of 7.4. The pre-optimized geometry serves as a query for QM calculation (Chukwuemeka et al., 2021). A
Broyden–Fletcher–Goldfarb–Shanno (BFGS) geometry optimizer was used for structure minimization and optimization at the semi-empirical theory level. The keywords “DIPOLE” and “MULLIK” were used to compute the dipole moments and Mulliken atomic charges, respectively. The time-dependent Hartree Fock (TDHF) was used to calculate the molecular polarizabilities using the "POLAR" keyword incorporated in MOPAC 2016. A visualization tool (Jmol software program) was used to visualize the charge distribution diagram of frontier molecular orbitals (FMOs) and molecular electrostatic potential (MEP) of the docked compounds (Hanson, 2010). A quantum chemical calculation via density functional theory (DFT) was used to investigate the physicochemical properties of lead phytochemicals with the best conformer distribution. All the quantum chemical reactivity descriptors were computed from the energies of the highest occupied and lowest unoccupied molecular orbitals (EHOMO–LUMO). The descriptors include the following:

- Energy band gaps (Eg).
- Ionization energy (I).
- Electron affinity (A).
- Chemical hardness (η).
- Chemical softness (δ).
- Chemical potential (μ).
- Electronegativity (χ).

2.1.2 Protein and ligand preparation

The crystal structures of the molecular targets RBD of spike glycoprotein (PDB: 6MJO), 3CLPro (PDB: 6M2N), PLPro (7CJM), and RdRp (7D4F) were obtained from the protein data bank (https://www.rcsb.org/). They were prepared using Schrödinger’s protein preparation wizard (Sastry et al., 2013). Hydrogen bond optimizations, water removal, protein structure correction, and ultimately protein energy minimization using the OPLS-2005 force field were carried out during the preparation. The position of the co-crystallized ligands for each target was used to define the protein binding pocket for receptor grid generation. Subsequently, the 3D structure of 1,000 compounds consisting of the substructure of the co-crystallized ligand of the targets was downloaded from the PubChem database (https://pubchem.ncbi.nlm.nih.gov). The structures of the ligands were cleaned, and their geometries were subjected to structural optimization using the default specifications of the LigPrep module incorporated in the Schrödinger suite and utilized for hypothesis generation (Balogun et al., 2021a). The prepared proteins and fully optimized geometry of ligands were used as input for molecular docking.

2.1.3 Molecular docking

High-throughput virtual screening (HTVS) of the prepared phytochemicals library was performed using the HTVS module of Maestro integrated into the Schrödinger suite (Friesner et al., 2004). The HTVS module used the 3D crystallographic structure of the therapeutic targets and fitted the ligands based on their structural conformations. During the virtual screening process, an energy score of −5.0 kcal/mol was set as the threshold to identify potential anti-coronavirus agents. The hits generated, 105 compounds out of 1,000 screened phytochemicals, were subjected to molecular docking by considering the flexibility of the protein using the SP (standard precision) model. The compounds were subjected to XP (extra precision) docking using the GLIDE XP module incorporated in Maestro to achieve further accurate results based on binding affinity and pose. The structural and energy information between the protein–ligand complexes was considered for energetic computation and further stability studies. The 2D interaction profile of protein–ligand complexes was generated using the Ligand Interaction Diagram (LID) in Maestro (Friesner et al., 2004; Balogun et al., 2021b). The reproducibility and reliability of the docking procedure were validated by superimposing and re-docking the co-crystalized ligand structures into the target active site, which generated an RMSD value of 0.76 Å (normal range: 0–2 Å). This confirms the reliability of the docking protocol.

2.2 Binding free energy and contribution energies calculation using MM-GBSA

The XP-screened compounds were further subjected to a Prime MM/GBSA analysis, where their binding energies were computed to investigate the inhibitory potential of the docked compounds against the targets. Based on the number of energy parameters generated by the Prime algorithm, free energy parameters were used to gain mechanistic insight into the biological activity of the compounds. Nonetheless, the ligand strain energy, Coulomb energy, and van der Waals energy were also assessed in filtering the final hit compounds (Genheden and Ryde, 2015; Schrödinger, 2020). The binding free energy and essential amino acid interactions between the protein–ligand complexes were computed using the following equations:

\[
\Delta G_{\text{bind}} = \Delta E + \Delta G_{\text{olv}} + \Delta G_{\text{SA}}.
\]

\[
\Delta E = E(\text{complex}) - E(\text{protein}) - E(\text{ligand}),
\]

where \(E(\text{complex}), E(\text{protein}), \) and \(E(\text{ligand})\) are the minimized energies of the protein–inhibitor complex, protein, and inhibitor, respectively.

\[
\Delta G_{\text{olv}} = G_{\text{olv}}(\text{complex}) - G_{\text{olv}}(\text{protein}) - G_{\text{olv}}(\text{ligand}),
\]

where \(G_{\text{olv}}(\text{complex}), G_{\text{olv}}(\text{protein}), \) and \(G_{\text{olv}}(\text{ligand})\) are the salvation free energies of the complex, protein, and inhibitor, respectively.

\[
\Delta G_{\text{SA}} = G_{\text{SA}}(\text{complex}) - G_{\text{SA}}(\text{protein}) - G_{\text{SA}}(\text{ligand}),
\]

where \(G_{\text{SA}}(\text{complex}), G_{\text{SA}}(\text{protein}), \) and \(G_{\text{SA}}(\text{ligand})\) are the surface area energies for the complex, protein, and inhibitor, respectively.
2.3 Drug likeness and ADMET evaluation

To compute the lead compounds’ physicochemical parameters and pharmacokinetic models, the compounds inputted structures were transformed into their canonical simplified molecular input line entry system (SMILES) form. Therefore, the curated ligand database’s SMILES were uploaded to the admetSAR web server (http://lmmd.ecust.edu.cn/) (Cheng et al., 2012) and SwissADME (Daina et al., 2017). Drug-likeness is a method for determining if a therapeutic agent is appropriate for orally active medications. Lipinski’s rule of five principles are used to compute in silico predictions based on parameters such as molecular weight, hydrogen bond donor, and hydrogen bond acceptor (Lipinski et al., 2001).

2.4 Molecular dynamics simulation

Molecular dynamics (MD) simulations were conducted to predict the protein’s dynamic motion and stability at the atomic scale with the bounded protein. The DESMOND module integrated into the Schrodinger suite was used to generate protein–ligand topologies and trajectories. The protein–ligand complexes were performed for 100 ns with an OPLS3 force field, using the DESMOND version of Schrödinger (2018). The solvation box was designed as the shape of the rhombic dodecahedron type and solvated using the TIP3P (transferable intermolecular potential 3 point) and an orthorhombic box (10 Å × 10 Å × 10 Å buffer) water model. Na+ and Cl− ions in 0.15 mM concentration were added to neutralize the charge of the systems during simulation. The system was performed for 10 ns after all the pre-processing phases. The V-rescale and Parrinello–Rahman methods were used for temperature and pressure coupling, respectively (Parrinello and Rahman, 1981). Leonard-Jones potential and the particle mesh Ewald (PME) method were used to handle van der Waals and long-range electrostatic interactions, respectively (Darden et al., 1993). The complexes underwent a final MD simulation production run for 100 ns. Root-mean-square deviations (RMSD) were computed using the MD trajectory to estimate the variations in protein conformation during the various simulations period, and root-mean-square fluctuation (RMSF) as well as the total number of intermolecular contacts were used for each protein–ligand complex to gain insights into the compound’s inhibitory potential (Pearson, 1986).

3 Results and discussion

3.1 Frontier molecular orbital analysis and global reactivity descriptors

FMOs such as HOMO and LUMO are essential in demystifying the chemical reactivity at the atomic level and are crucial descriptors for rationalizing various chemical reactions. The reactivity descriptors calculated for lonchocarpol A, broussonol A, diplacol, and dexamethasone are shown in Table 1. HOMO energy denotes the potential of a molecule to easily donate an electron, which also corresponds to the ionization potential of a molecule. In contrast, the electron-withdrawing potential of compounds is referred to as the LUMO energy, which signifies the first empty innermost orbital unfilled by an electron and correlates with a molecule’s electron affinity. The band gap energy is the difference between the HOMO and the LUMO energy and provides information about the compound’s chemical stability at the molecular level. Band gap
energy also describes the chemical reactivity of a molecule, deciphering the movement of electrons from the ground state to its excitation state. Furthermore, other parameters (such as chemical hardness, softness, electronegativity, or polarizability) that provide information about compounds’ ionic structure and the electronic configuration can be easily computed via HOMO–LUMO energy (Pearson, 1986; Sylaja et al., 2017). For example, a lower energy gap between two frontier molecular orbitals means lower kinetic stability and higher polarizability and reactivity of a molecule, which indicates the softness of the molecule and vice versa.

Lonchocarpol A has the second highest HOMO energy value (EHOMO = −8.77 eV), denoting the valence electron density distribution for lochocarpol A is more available to be donated, suggesting that lonchocarpol A is the most reactive compound after dexamethasone. Similarly, broussonol E and diplacol recorded a HOMO energy value of −8.647 and −8.763 eV, respectively. Clearly, lonchocarpol A, broussonol E, and diplacol demonstrated an intermolecular charge transfer as they excited from the ground state (S0) to the first excitation state. Interestingly, the LUMO energy is in the following order: lochocarpol A < dexamethasone < diplacol < broussonol E. The LUMO energy suggests that lonchocarpol A and dexamethasone are more susceptible to accepting electronic density because a lower energy molecular orbital will describe the additional electron. The chemical reactivity of a compound is measured using the HOMO–LUMO energy gap (ΔEgap), which represents a lower energy difference (lower energy gap) (Figure 1).

Broussonol E had the lowest energy value of −7.530 eV compared to dexamethasone (−8.275 eV), which implies more chemical reactivity to broussonol E. Lonchocarpol A and
dexamethasone showed the same energy gap (−8.275 eV) which is in consistent with their LUMO values. This suggests that both the compounds may share similar chemical reactivity properties and mechanisms of action toward the targets. To fully gain insights into the reactivity and chemical species of the top four compounds with drug-likeness properties, the following parameters were evaluated from the HOMO and LUMO energy: ionization potential, electron affinity, chemical hardness (η), chemical softness (ζ), electronic chemical potential (µ), electrophilicity index (ω), and electronegativity (χ). The expression for the reactivity parameters, as mentioned earlier, has been described according to Koopman’s theory (Phillips, 1961) and can be calculated by the following mathematical statements:

\[
\text{Energy Gap } \Delta E = \text{HOMO}_c - \text{LUMO}_c; \\
\text{Ionization Potential } I = -\text{E}_{\text{HOMO}}; \\
\text{Electron affinity } A = -\text{E}_{\text{LUMO}}; \\
\text{Chemical hardness } \eta = \frac{1}{2} \left( \frac{\partial^2 E}{\partial N^2} \right) V = \frac{1}{2} \left( \frac{\partial N}{\partial V} \right) V = (I - A)/2; \\
\text{Chemical potential } \mu = \left( \frac{\partial E}{\partial N} \right) V = -(I + A)/2; \\
\text{Electronegativity } \chi = -\mu = -\left( \frac{\partial E}{\partial N} \right) V = (I + A)/2; \\
\text{Softness } \zeta = \frac{1}{\eta}; \\
\text{Electrophilicity index } \omega = \frac{\mu^2}{2\eta}
\]

Ionization energy helps to determine the amount of free energy required to remove an electron from an atom or molecule. Furthermore, electron affinity represents the amount of energy liberated when an atom or molecule is attached to a neutral atom or molecule. Lower ionization potential indicates lower stability or higher reactivity of the compound and its contribution toward analyzing inhibitory potential. Contrarily, electron affinity depicts the high electron-withdrawing ability of a compound. Table 1 shows that lonchocarpol A and dexamethasone had the highest chemical stability and electron-withdrawing potential compared to broussonol E and diplacol. This observation is consistent with the gap energy between the HOMO and LUMO FMOs (Figure 2). The softness and hardness properties of compounds contribute to their chemical stability. Although a higher hardness value means a more stable chemical entity, compound’s stability decreases with softness. Pearson’s HSAB theory proposed that a favorable interaction between two compounds occurs when both are hard and soft (Pearson, 1990; Arjunan and Mohan, 2009). It is evident from Table 1 that lonchocarpol A and dexamethasone have chemical hardness values of 4.138 eV, indicating they are the most stable compound, followed by diplacol (3.931 eV) and broussonol E (3.765). The chemical softness of the compounds shows that there is only a subtle difference between the compounds denoting their chemical stability. The ability of a compound to not decompose spontaneously into an element, which denotes its stability, is determined by a higher negative chemical potential. All the compounds demonstrated chemical stability due to their negative value of chemical potential. Electronegativity and electrophilicity are another important set of reactivity descriptors. Lonchocarpol A and dexamethasone have the same electrophilicity index (2.595 eV) and electronegativity values (4.639 eV), which implies their susceptibility to accept electron density and classifies them as promising electrophilic compounds. Broussonol E was recorded as the most electrophilic molecule with an electrophilicity index value of 3.165 eV. Therefore, Table 1 provides appropriate information regarding the chemical reactivity and stability of the studied compounds.

### 3.2 Molecular electrostatic potential

MEPs have proved useful in determining the relative polarity of compounds as well as providing essential information on molecular charge distribution patterns. As a result, studying the MEP of the examined compounds may provide insight into their electrophilic and nucleophilic cores (Figure 2). It is worth mentioning that molecular electrostatic potential data can be classified using traditional color codes. The electron-rich centers are indicated by a red color scheme, which symbolizes the highest negative electrostatic potential. On the other hand, a blue hue region denotes electron-deficient areas (i.e., the most positive electrostatic potential). The light blue, yellow, and green moieties represent a molecule’s region of slightly electron-deficient cores, marginally electron-rich areas, and zero electrostatic potential.
TABLE 2 Calculated Mulliken atomic charges of the top four compounds.

| Atom (lonchocarpol A) | Mulliken charge (lonchocarpol A) | Atom (broussonol E) | Mulliken charge (broussonol E) | Atom (diplacol) | Mulliken charge (diplacol) | Atom (dexamethasone) | Mulliken charge (dexamethasone) |
|-----------------------|----------------------------------|---------------------|---------------------------------|-----------------|-----------------------------|---------------------|-------------------------------|
| O                     | 0.4443                           | O                   | 0.34135                         | O               | 0.4335                      | F                   | 0.22161                       |
| O                     | 0.5236                           | O                   | 0.5126                          | O               | 0.57712                     | O                   | 0.57211                       |
| O                     | 0.4988                            | O                   | 0.48847                         | O               | 0.5101                      | O                   | 0.56953                       |
| O                     | 0.5915                            | O                   | 0.4938                         | O               | 0.57763                     | O                   | 0.47631                       |
| O                     | 0.5106                            | O                   | 0.59099                         | O               | 0.49375                     | O                   | 0.5496                       |
| C                     | 0.207278                          | O                   | 0.53436                         | O               | 0.47351                     | O                   | 0.50928                       |
| C                     | 0.6191B                           | O                   | 0.48846                         | O               | 0.47101                     | C                   | 0.11952                       |
| C                     | 0.55362                           | C                   | 0.68688B                        | C               | 0.111928                    | C                   | 0.17951                       |
| C                     | 0.73661B                          | C                   | 0.45146                         | C               | 0.09131                     | C                   | 0.19425                       |
| C                     | 0.37697                           | C                   | 0.553329                         | C               | 0.73622B                    | C                   | 0.16731                       |
| C                     | 0.702363B                         | C                   | 0.602767B                       | C               | 0.601256                     | C                   | 0.12679                       |
| C                     | 0.45972                           | C                   | 0.219836                        | C               | 0.597263                     | C                   | 0.40641                       |
| C                     | 0.597616                          | C                   | 0.012863                        | C               | 0.003862                     | C                   | 0.51737                       |
| C                     | 0.560552                          | C                   | 0.21543                         | C               | 0.4738                      | C                   | 0.07233                       |
| C                     | 0.20439                           | C                   | 0.532411                        | C               | 0.611783B                   | C                   | 0.11045                       |
| C                     | 0.21769                           | C                   | 0.620079B                       | C               | 0.6357B                     | C                   | 0.00023                       |
| C                     | 0.2394                            | C                   | 0.29245                         | C               | 0.23452                     | C                   | 0.32459                       |
| C                     | 0.00819                           | C                   | 0.60901B                        | C               | 0.56147                     | C                   | 0.62434B                      |
| C                     | 0.01567                           | C                   | 0.06649                         | C               | 0.30019                     | C                   | 0.40779                       |
| C                     | 0.32883                           | C                   | 0.22206                         | C               | 0.14207                     | C                   | 0.17763                       |
| C                     | 0.35013                           | C                   | 0.3768                         | C               | 0.36688                     | C                   | 0.40718                       |
| C                     | 0.35526                           | C                   | 0.52838                         | C               | 0.123515                    | C                   | 0.60478B                      |
| C                     | 0.44732                           | C                   | 0.164187                        | C               | 0.39687                     | C                   | 0.60459B                      |
| C                     | 0.158178                          | C                   | 0.23667                         | C               | 0.215686                    | C                   | 0.11383                       |
| C                     | 0.371538                          | C                   | 0.37461                         | C               | 0.34313                     | C                   | 0.5243                       |
| C                     | 0.137273                          | C                   | 0.34269                         | C               | 0.29078                     | C                   | 0.23229                       |
| C                     | 0.64618B                          | C                   | 0.210516                        | C               | 0.227416                    | C                   | 0.43236                       |
| C                     | 0.65128B                          | C                   | 0.149934                        | C               | 0.65057                     | C                   | 0.621716B                     |
| C                     | 0.644744B                         | C                   | 0.60605                        | C               | 0.33623                     | H                   | 0.174307                      |
| C                     | 0.65366B                          | C                   | 0.65899                         | C               | 0.161681                    | H                   | 0.199263                      |
| H                     | 0.177918                          | C                   | 0.48924                         | C               | 0.65057                     | H                   | 0.62742                       |
| H                     | 0.265334                          | C                   | 0.65697                         | C               | 0.64253                     | H                   | 0.206263                      |
| H                     | 0.257829                          | H                   | 0.208693                        | H               | 0.209879                    | H                   | 0.217255                      |
| H                     | 0.180081                          | H                   | 0.147876                        | H               | 0.229331                    | H                   | 0.223055                      |
| H                     | 0.186761                          | H                   | 0.290018                        | H               | 0.371701                    | H                   | 0.155531                      |
| H                     | 0.170022                          | H                   | 0.224598                        | H               | 0.172366                    | H                   | 0.20782                       |
| H                     | 0.214067                          | H                   | 0.192595                        | H               | 0.212728                    | H                   | 0.184112                      |
| H                     | 0.193514                          | H                   | 0.204748                        | H               | 0.248627                    | H                   | 0.185519                      |
| H                     | 0.204486                          | H                   | 0.264931                        | H               | 0.222989                    | H                   | 0.21967                       |
| H                     | 0.201087                          | H                   | 0.190502                        | H               | 0.210694                    | H                   | 0.193173                      |
| H                     | 0.217387                          | H                   | 0.21727                         | H               | 0.371795                    | H                   | 0.211806                      |
| H                     | 0.413892                          | H                   | 0.41358                         | H               | 0.196626                    | H                   | 0.200817                      |
| H                     | 0.37606                           | H                   | 0.376127                        | H               | 0.185069                    | H                   | 0.208458                      |
| H                     | 0.242263                          | H                   | 0.37573                         | H               | 0.219815                    | H                   | 0.191525                      |
| H                     | 0.226444                          | H                   | 0.214546                        | H               | 0.197055                    | H                   | 0.211997                      |
| H                     | 0.204763                          | H                   | 0.210442                        | H               | 0.176395                    | H                   | 0.217508                      |
| H                     | 0.204729                          | H                   | 0.208894                        | H               | 0.41009                     | H                   | 0.209423                      |
| H                     | 0.202827                          | H                   | 0.205874                        | H               | 0.220548                    | H                   | 0.221053                      |
| H                     | 0.220196                          | H                   | 0.230009                        | H               | 0.200566                    | H                   | 0.211325                      |
| H                     | 0.203416                          | H                   | 0.209537                        | H               | 0.214513                    | H                   | 0.362716                      |

(Continued on following page)
potions, respectively. We can deduce that a molecule’s potential declines in the following order based on the color scheme: blue > light blue > green > yellow > red. Figure 2 represents molecular electrostatic potential maps of lonchocarpol A, broussonol E, diplacol, and dexamethasone. There is a clear maximum concentration of electrons located at the alkyl groups and oxygen atoms of lonchocarpol A attached to the diphenyl groups. In contrast, the region of most positive electron potential of lonchocarpol A is located at the hydrogen atoms of the methyl group of the phenyl group. Dexamethasone recorded the highest negative electrostatic potential, including multiple hydroxyl group points. All the compounds have been reported for their biological and chemical properties. Therefore, MEP provides detailed insights into the molecular charge distribution clusters in the studied compound.

3.3 Mulliken population analysis

Table 2 shows the atomic charge distribution of lonchocarpol A, broussonol E, diplacol, and dexamethasone computed via the Mulliken population analysis using the PM7-based semi-empirical Hamiltonian calculations.

Because atomic charges affect compounds’ molecular and electrical characteristics, estimating each compound’s partial atomic charges is critical for understanding the charge distribution. Calculating the atomic charges of any small molecule ligand can be used to calculate the adsorptive centers. Table 2 shows that the examined structures’ oxygen and carbon atoms have electron-rich chemical species (i.e., they have the most negative electronic charges), which may be due to their molecular relaxation. However, the predominant positive charge regions were observed to be covered by carbon atoms, despite the fact that some carbon atoms in the investigated compounds possessed negative atomic charges. Table 2 shows that the atoms O7, C9, and C27–30 of lonchocarpol A have the most negative atomic charges, whereas the C11 and C16 of broussonol E have the highest negative atomic charges in lonchocarpol A. In terms of broussonol E, C10 and C16 were observed for negatively charged atoms, whereas C11 and C15 occupied positive regions. Dexamethasone’s ionic structure established negatively charged electrostatic contacts with C18 and C22–23 while demonstrating C27 as the only positive electrostatic atoms. Overall, it can be deduced that there are variations between the atoms of the studied compounds occupying positive and negative regions. This is also supported by the difference between the compound’s inhibitory potential and their chemical stability.

3.4 Non-linear optics analysis

NLO materials have played an important role in contemporary technologies, providing various industrial and medicinal benefits, some of which have been detailed in prior studies (Muhammad et al., 2021). The most prominent quality of analyzing NLO properties from a fascinating perspective on chemical methodologies and applications is their tendency to provide considerable insights into how small changes in molecular structures might alter NLO responses. Tables 3, 4 present and summarize various NLO responses and their components for lonchocarpol A, broussonol E, diplacol, and dexamethasone estimated using the PM7 semi-empirical Hamiltonian calculations in MOPAC 2016. The dipole moment ($\mu$) gives information on a bond’s or molecule’s ionic character state. Ionic property is generally associated with molecules with a higher dipole
Dipole moments are crucial in forecasting a molecule’s structure and reactivity. The dipole moments for the studied compounds, lonchocarpol A, broussonol E, diplacol, and dexamethasone, were 2.663, 4.122, 5.209, and 5.334, respectively. The computation of polarizability ($\alpha_0$) and hyperpolarizability ($\beta_0$ and $\gamma_0$) in molecular systems are

| Parameters                  | Lonchocarpol | Broussonol E | Diplacol | Dexamethasone (control) |
|-----------------------------|--------------|--------------|----------|-------------------------|
| Dipole moment (Debye)       | $\mu_x$      | 0.936        | 5.075    | 1.003                   |
|                             | $\mu_y$      | 4.004        | 1.089    | 1.943                   |
|                             | $\mu_z$      | -0.284       | 0.444    | 4.865                   |
|                             | $\mu_0$      | 4.122        | 5.209    | 5.334                   |
| Polarizability (a.u.)       | $\alpha_{xx}$| 423.2262     | 422.1966 | 474.0150                |
|                             | $\alpha_{xy}$| -32.1733     | 422.1966 | 20.7672                 |
|                             | $\alpha_{yy}$| 321.3188     | 415.5835 | 337.6663                |
|                             | $\alpha_{yz}$| 5.8622       | 3.8562   | 11.7997                 |
|                             | $\alpha_{zz}$| 12.2187      | 38.3700  | 32.3409                 |
|                             | $\alpha_0$   | 32.31294     | 356.1023 | 346.88497               |
| Hyperpolarizability (a.u.)  | $\beta_{xxx}$| -828.1153    | -1,758.3536 | -857.13155              |
|                             | $\beta_{xxy}$| 1,274.9156   | 1,770.9625 | 1,087.95116              |
|                             | $\beta_{xyy}$| 408.19522    | -939.4038 | 392.33997               |
|                             | $\beta_{yyz}$| -205.66502   | 198.04271 | -145.10306              |
|                             | $\beta_{zzz}$| 218.15245    | 622.87825 | -23.37222               |
|                             | $\beta_{xxz}$| -56.15189    | 19.08646  | 31.58709                |
|                             | $\beta_{xxy}$| 42.70298     | -90.27506 | 22.36855                |
|                             | $\beta_{xyy}$| 67.05907     | 86.15574  | 8.42037                 |
|                             | $\beta_{yyz}$| -12.20923    | -74.49353 | 9.07706                 |
|                             | $\beta_0$    | 678.2379     | 1,946.6357| 633.4487                |
|                             | $\gamma_{xxxx}$| 136.082.88516| 437.948.46962 | 122.946.68736 |
|                             | $\gamma_{xxy}$| 40.946.5150  | 86.024.86024 | 18.258.69096 |
|                             | $\gamma_{xyy}$| 6.622.48729  | 6.562.67994 | 3.978.16272 |
|                             | $\gamma_{yyz}$| 49.373.55773 | 208.421.66382 | 48.042.21099 |
|                             | $\gamma_{zzz}$| 364.8.39107  | 18.233.01061 | 11.992.55351 |
|                             | $\gamma_{xzz}$| 4,755.42360  | 11,210.83548 | 2,029.91003 |
|                             | $\gamma_{yzz}$| 61,448.83940 | 205,675.48909 | 53,862.59062 |
|                             | $\gamma_0$    | 61,448.83940 | 205,675.48909 | 53,862.59062 |

Standard value for urea ($\mu = 1.3732$ Debye; $\beta_0 = 0.3728 \times 10^{-30}$ esu); esu, electrostatic unit. (For $\alpha$, 1 a.u is equal to $0.1482 \times 10^{-24}$ esu. Similarly, for $\beta$, 1 a.u is equal to $8.6393 \times 10^{-33}$ esu).

| Compound               | Dipole moment (Debye) | Static polarizability ($\alpha_0 \times 10^{-23}$ esu) | Static first-order hyperpolarizability ($\beta_0 \times 10^{-30}$ esu) | Static second-order hyperpolarizability ($\gamma_0 \times 10^{-39}$ esu) |
|------------------------|-----------------------|----------------------------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------------|
| Lonchocarpol A         | 2.663                 | 5.1226                                                   | 0.0586                                                              | 30.949.9528                                                          |
| Broussonol E           | 4.122                 | 5.2774                                                   | 0.0017                                                              | 103.592.6267                                                         |
| Diplacol               | 5.209                 | 5.1654                                                   | 0.0590                                                              | 32.680.4170                                                         |
| Dexamethasone (control)| 5.334                 | 4.1202                                                   | 0.7207                                                              | 7.952.0242                                                          |

moment value. Furthermore, dipole moments are crucial in forecasting a molecule’s structure and reactivity. The dipole moments for the studied compounds, lonchocarpol A, broussonol E, diplacol, and dexamethasone, were 2.663, 4.122, 5.209, and 5.334, respectively. The computation of polarizability ($\alpha_0$) and hyperpolarizability ($\beta_0$ and $\gamma_0$) in molecular systems are
TABLE 5  Binding energy of compounds library against the SARS-CoV-2 therapeutic target.

| S/ NS | Compounds | Spike glycoprotein RBD (6M2N) (kcal/mol) | Compounds | 3CLpro (6M2N) (kcal/mol) | Compounds | PLpro (7CJM) (kcal/mol) | Compounds | RdRp (7D4F) (kcal/mol) |
|-------|-----------|---------------------------------------|-----------|------------------------|-----------|------------------------|-----------|----------------------|
| 1     | Rutin     | −10.941                               | Nicotiflor | −11.442                | Aucubin   | −8.767                 | Acteosides | −10.632              |
| 2     | Delphinidin 3-O-beta-D-sambubioside | −10.709 | Schafloside            | −11.389 | Ruthin                | −8.698     | Cynaroside            | −9.193     |
| 3     | Hesperidin | −10.627                               | Acteoside  | −11.291                | Nicotiflor | −8.685                 | Hydroxycitric acid | −9.087     |
| 4     | Acteoside  | −10.033                               | Mallophenol B | −11.226                | Mallophenol B | −8.106              | Rutin         | −8.704     |
| 5     | Kuoromanin| −9.902                                | Kolaflavanone | −10.496                | Hesperidin | −7.671                 | Schaftoside | −8.347     |
| 6     | Pelargonidin 3-glucose          | −9.684     | Aucubin               | −10.295 | Cynaroside            | −7.537     | Bergetin              | −8.127     |
| 7     | Lauroside E | −9.599                               | Tanariflavanone C | −10.278                | Kuoromanin | −7.186                 | Kuoromanin | −7.687     |
| 8     | Nicotiflor | −9.447                                | (+)-Gallocatechin gallate | −10.334                | Schaftoside | −7.114                 | Mallophenol B | −7.680     |
| 9     | Diplacol   | −8.733                                | Delphinidin 3-O-beta-D-sambubioside | −10.035                | Pelargonidin 3-glucose | −6.811    | Lauroside E            | −7.641     |
| 10    | Myricetin  | −8.725                                | Rutin      | −8.987                 | Nymphaeol B | −6.687                 | Hydroxycitric acid | −7.629     |
| 11    | Nymphaeol B | −8.291                               | Luteolin   | −8.866                 | (+)-Gallocatechin gallate | −6.663    | Pelargonidin 3-glucose | −7.489     |
| 12    | Schaftoside| −8.251                                | Nymphaeol C | −8.748                 | Macaranone A | −6.548                | (+)-Gallocatechin gallate | −8.937     |
| 13    | (+)-Gallocatechin gallate        | −8.289     | Macaruzza A           | −8.723     | Myricetin            | −6.531     | Delphinidin 3-O-beta-D-sambubioside | −7.333     |
| 14    | Macaruzza A | −8.191                               | Isovitexin | −8.710                 | Bergetin   | −6.378                 | Gallic acid | −7.122     |
| 15    | Tanariflavanone D                | −8.160     | Lonchocarpol A*       | −8.644*     | Quercetin           | −6.377     | Lonchocarpol A         | −7.017     |
| 16    | Chlorogenic acid                 | −8.131     | Diplacol*             | −8.576*     | Isovitexin          | −6.277     | Nicotiflor            | −6.831     |
| 17    | Isovitexin                        | −8.096     | Tometosolano D        | −8.470     | Lonchocarpol A       | −6.161     | Aucubin               | −6.752     |
| 18    | Cynaroside                       | −7.991     | Isolviolflavonol      | −8.451     | Lauroside E         | −6.117     | Tanariflavanone D      | −6.739     |
| 19    | Quercetin                        | −7.996     | Fisetin               | −8.459     | Acteoside           | −6.115     | Broussonol E          | −6.342     |
| 20    | Broussonol E                     | −7.904     | Denticulalflavonol    | −8.253     | Nymphaeol A         | −6.047     | Macaranoside F         | −6.255     |
| 21    | Macaruzza A                      | −7.884     | Glepidotin A          | −8.231     | Diplacol            | −5.972     | Chlorogenic acid       | −6.089     |
| 22    | Aucubin                          | −7.818     | Myricetin             | −8.187     | Catalpol            | −5.733     | Protopicatech acid      | −6.019     |
| 23    | Broussonol E                     | −7.490     | Catalpol              | −8.164     | Tometosolano D      | −5.692     | Diplacol               | −5.997     |
| 24    | Mallophenol B                    | −7.425     | Macaranin             | −8.134     | Broussonol E        | −5.674     | Cianidanol            | −5.992     |
| 25    | Luteolin                         | −7.414     | Cynaroside            | −8.066     | Macaruzza A         | −5.631     | Tometosolano D         | −5.875     |
| 26    | Lonchocarpol A*                  | −7.331     | Broussonol E*         | −8.069*     | Isolviolflavonol    | −5.602     | Isovitexin            | −5.871     |
| 27    | Dexamethasone                     | −5.641     | Dexamethasone*        | −5.302*     | Dexamethasone       | −3.939     | Dexamethasone         | −2.946     |

*Selected compounds and their binding energy against 3CLpro for further molecular dynamics analysis.

Note that $E_0$ describes the energy of the unperturbed molecules; $Fi$ represents the field at the origin; and $\mu_0$, $\alpha_0$, and $\gamma_{0ij}$ correlate to the dipole moment, static polarizability, first-order hyperpolarizability, and second order hyperpolarizability, respectively. The total dipole moment $\mu$, static mean polarizability $\alpha_0$, mean first-order hyperpolarizability $\beta_0$, and second order hyperpolarizability $\gamma_0$ can be estimated by the following equations:

$$
E = E_0 - \sum_{i=1}^{N} \frac{\phi_i}{2} \sum_{j=1}^{N} \frac{\phi_j}{2} \sum_{k=1}^{N} \frac{\beta_{ijk}}{24} \sum_{l=1}^{N} \frac{\beta_{ijkl}}{24} \ldots
$$

$$
\text{Dipole moment } \mu = \sqrt{\mu_x^2 + \mu_y^2 + \mu_z^2};
$$

$$
\text{Static mean polarizability } \alpha_0 = \left( \alpha_{xx} + \alpha_{yy} + \alpha_{zz} \right) / 3;
$$
TABLE 6 Molecular interaction profiling and docking score of top four compounds.

| S/ N | Interacting amino acid residues | Lead compounds against 3CLPro | Docking score (kcal/mol) | H-bond | Hydrophobic | Polar | Charged (negative) | Charged (positive) | Glycine |
|------|--------------------------------|------------------------------|--------------------------|--------|-------------|------|-------------------|-------------------|---------|
| 1    | Lonchocarpol A                  | TYR54, PRO52, MET49, CYS44, VAL42, LEU27, CYS45, VAL186, ALA191, LEU167, PRO168, MET165 | 8.644 | ARG188 | HIS41, ASN142, GLN189, THR190, GLN192, HIS164 | ASP48, ASP187, GLU166 | ARG188 | GLY143 |
| 2    | Diplacol                        | TYR54, PRO52, CYS44, MET49, MET165, CYS145, LEU27 | 8.576 | ARG188 | GLN189, HIS41, HIS164, ASN142, THR26, THR25, THR24 | ASP48, GLU166, ASP187 | ARG188 | GLY143 |
| 3    | Broussonol E                    | CYS145, CYS44, MET49, PRO52, TYR54, MET165, LEU167 | 8.644 | ARG188 | THR26, THR25, THR24, HIS164, GLN192, THR190, GLN189, HIS41, ASN142 | ASP48, ASP187, GLU166 | ARG188 | GLY143 |
| 4    | Dexamethasone                   | PRO168, LEU167, MET165, MET49, CYS44, PRO52, TYR54 | 5.302 | ARG188 | HIS164, GLN189, HIS41, ASN142 | GLU166, ASP48, ASP187 | ARG188 | — |

Static first order hyperpolarizability $\beta = \sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2};$ (16)

where $\beta_x = \frac{3}{5}(\beta_{xxx} + \beta_{xxy} + \beta_{xyx}),$ (17)

$\beta_y = \frac{3}{5}(\beta_{xyy} + \beta_{yyx} + \beta_{yx y}),$ (18)

$\beta_z = \frac{3}{5}(\beta_{zzz} + \beta_{xyz} + \beta_{zyz}),$ (19)

$\beta_{tot} = \gamma(\beta_{xxx} + \beta_{xxy} + \beta_{xyx}) \cdot (\beta_{yxy} + \beta_{yyx} + \beta_{yx y}) \cdot (\beta_{zxx} + \beta_{zzz} + \beta_{xyz}),$ (20)

$\gamma = \frac{1}{3}\{\gamma_{xxx} + \gamma_{yxy} + \gamma_{zxx} + 2(\gamma_{xxy} + \gamma_{xyx} + \gamma_{xyy})\}. $ (21)

Notably, any compound with a higher value of first-order hyperpolarizability denotes an NLO active compound and vice versa. Table 4 shows that the hyperpolarizability value of dexamethasone is (0.7207 × 10−30), which is 10 times higher than that of lonchocarpol A (0.0586 × 10−30), broussonol E (0.0017 × 10−30), and diplacol (0.0590 × 10−30). Collectively, this study proposed that dexamethasone is the most suitable compound for NLO-based technology.

3.5 Molecular docking and binding site analysis

3.5.1 Inhibitory potential of promising phytodrugs against SARS-CoV-2 spike glycoprotein, 3CLpro, PLpro, and RdRp

The 3CLpro, also referred to as NSP5, mediates the maturation of Nsps, which is vital in the virus’s lifecycle. The structural analysis and catalytic mechanism of 3Clpro using biophysical techniques have been widely investigated (Pillaiyar et al., 2016). Therefore, 3CLpro remained an important therapeutic target for developing potential anti-coronavirus drug candidates. Peptide inhibitors and small molecules are inhibitors targeting the SARS-CoV-2 3CLpro. From the molecular docking result, various molecular interactions, including hydrogen bonding, hydrophobic, polar, and pi–pi interactions, were observed and analyzed while ranking the compounds based on their binding poses. Although, nicotiflorin, schaftoside, acetoside, and mallophenol demonstrated average binding energy of −11.20 kcal/mol (Table 5). They were eliminated from further studies because of their undruggable properties. Interestingly, lonchocarpol A, broussonol E, diplacol, and dexamethasone (reference compound) were selected for further analysis because of drug-like properties, molecular interactions, and high binding energy.

Lonchocarpol A is a flavone obtained from Lonchocarpus and Erythrina species and has been reported for its biological activities, including anticancer, insecticidal, and antibacterial activity, amongst others. Interestingly, lonchocarpol A has also been synthesized using various synthetic methods and have also received significant interest as a compound with numerous therapeutic benefits (Salvatore et al., 1998). Lonchocarpol A has a binding affinity of −8.644 kcal/mol and hydrogen bond interactions with ARG188 based on its side hydroxyl group. All significant interactions exhibited by the compound were mainly due to its alkyl side group and phenyl ring.

The alkyl groups present in the phenyl moiety interact with hydrophobic amino acids TYR54, PRO52, MET49, and CYS44,
VAL42, LEU27, CYS145, VAL186, ALA191, PRO168, and MET165, and polar amino acids HIS41, ASN142, GLN189, THR190, GLN192, and HIS164 (Table 6). The other notable interactions were pi–pi/charge interactions between the aromatic ring of lonchocarpol A with ASP48, ASP187, GLU166, and ARG188 (Figure 3). The second selected lead compound, diplacol, showed a similar hydrogen bond with amino acid ARG188 as in lonchocarpol A; however, the two dihydroxylphenyl and the alky group of the compounds were responsible for its hydrophobic interactions with various amino acid residues (TYR54, PRO52, CYS44, MET49, MET165, CYS145, and LEU27) at the 3CLpro active site. The polar protein–ligand interactions exhibited by diplacol follow the same pattern as lonchocarpol A with subtle differences. Diplacol established polar interactions with the following amino acid residues: GLN189, HIS41, HIS164, ASN142,
THR26, THR25, and THR24; these residues were present at the binding pocket of 3CLpro. Broussonol E has a binding energy of −8.069 kcal/mol and shows key biomolecular interactions with certain amino acids such as THR26, THR25, THR24, HIS164, GLN192, THR190, GLN189, HIS41, and ASN142 within the 3CLpro active site. Broussonol E demonstrated other interactions, such as polar and pi-pi interactions. The reference compound (dexamethasone) has the least binding energy with similar inter- and intramolecular interactions with the lead compounds. Notably, dexamethasone interacted with PRO52, TYR54, MET165, MET49, and CYS44 amino acid residues, which were also recorded in the lead compounds interactions. Therefore, the top three compounds were proposed to have a similar mechanism of action as dexamethasone since they share the same amino acid interactions with the targets.

Spike is coronavirus’s major structural protein, which assembles as a trimer into a unique corolla structure on the virus’s surface. The spike protein mediates the virus interaction with the host cell by binding to the host angiotensin-converting enzyme (ACE-2). Certain host cell proteases, such as TMPRSS2, cleave the spike protein into two subunits, S1 and S2, which play

---

### TABLE 7 Pharmacokinetics profile of top four compounds.

| Models                        | Lonchocarpol A | Diplacol | Broussonol E | Dexamethasone |
|-------------------------------|----------------|----------|--------------|---------------|
| Blood-brain barrier           | BBB−           | BBB+     | BBB−         | BBB±          |
| Caco-2 permeability           | Caco-2⁺        | Caco-2−  | Caco-2−      | Caco-2⁺       |
| P-glycoprotein (substrate)    | Non-substrate  | Non-substrate | Non-substrate | Non-substrate |
| P-glycoprotein (inhibitor)    | Inhibitor      | Inhibitor | Inhibitor    | Non-inhibitor |
| LogS (aqueous solubility)     | −3.925         | −4.285   | −3.567       | −3.703        |
| Renal organic cation transporter 2 (OCT2) | Non-inhibitor | Non-inhibitor | Non-inhibitor | Inhibitor |
| Metabolism                    |                |          |              |               |
| CYP450 2C9 (substrate)        | Non-substrate  | Non-substrate | Non-substrate | Non-substrate |
| CYP450 2C9 (inhibition)       | Inhibitor      | Inhibitor | Inhibitor    | Non-inhibitor |
| CYP450 2D6 (substrate)        | Non-substrate  | Non-substrate | Non-substrate | Non-substrate |
| CYP450 2D6 (inhibition)       | Non-inhibitor  | Non-inhibitor | Non-inhibitor | Non-inhibitor |
| CYP450 3A4 (substrate)        | Substrate      | Non-substrate | Non-substrate | Substrate     |
| CYP450 3A4 (inhibition)       | Non-inhibitor  | Inhibitor | Inhibitor    | Non-inhibitor |
| CYP450 1A2                    | Inhibitor      | Inhibitor | Inhibitor    | Non-inhibitor |
| CYP450 2C19                   | Inhibitor      | Inhibitor | Inhibitor    | Non-inhibitor |
| Toxicity                      |                |          |              |               |
| Ames toxicity                 | Non-toxic      | Toxic    | Non-toxic    | Non-toxic     |
| hERG inhibition               | Inhibitor      | Inhibitor | Inhibitor    | Inhibitor     |
| Carcinogenicity               | Non-carcinogenic | Non-carcinogenic | Non-carcinogenic | Non-carcinogenic |
| Acute oral toxicity           | III            | III      | III          | III           |
| Rat LD50                      | 2.705          | 2.591    | 2.129        | 2.189         |
a crucial role in receptor recognition and cell membrane fusion (Millet and Whittaker, 2015). Therefore, blocking the coronavirus entry into the cell by targeting the spike glycoprotein has significantly been harnessed in the development of therapeutic agents against coronavirus. From the virtual screening results, rutin, delphinidin 3-O-beta-D-sambubioside, and hesperidin show the highest binding energy of $-10.941$, $-10.709$, and $-10.627$ kcal/mol, respectively. Unfortunately, these compounds failed the toxicity assessment and were eliminated from a further study. However, the leads compounds, namely, diplacol, broussonol E, and lonchocarpol A, have a binding affinity of $-8.733$, $-7.490$, and $-7.331$ kcal/mol against spike glycoprotein, respectively. The protein–ligand contacts show that the lead compounds established some essential hydrogen and hydrophobic interactions. It can be deduced that lonchocarpol has the lowest binding energy against spike glycoprotein when compared to the remaining two lead compounds (diplacol and broussonol E).

Papain-like proteinase (Plpro) plays a role in the cleavage of the N-terminus of the replicase polyprotein to produce non-structural proteins, including Nsp1, Nsp2, and Nsp3, which are involved in the virus replication (Li et al., 2016). Thus, based on the vital role played by Plpro in virus replication and infection, it has received intense consideration as a therapeutic target for coronavirus inhibitors. There have been no FDA-approved inhibitors of Plpro. Aucubin was recorded with the highest binding energy against Plpro, with $-8.767$ kcal/mol. Aucubin’s high binding energy may be attributed to its imidazole ring. Nicotiflorin optimally occupied the binding pocket of the target (Plpro), which may be attributed to its ring system. The presence of multiple hydroxyl groups at the nicotiflorin structures establishes intermolecular hydrogen bonds. Several docked compounds, including rutin, diplacol, hesperadin, and kuromanin, showed high binding energy against Plpro while demonstrating pi–pi and hydrophobic interactions with amino acid residues at the active site of Plpro. The lead compounds lonchocarpol A, diplacol, and broussonol E have a binding energy of $-6.161$, $-5.972$, and $-5.674$ kcal/mol, respectively, when docked into the binding pocket of Plpro. The binding energy against Plpro follows a similar pattern to that of 3CLpro.

However, the lead compounds’ structural poses, binding energy, and molecular interactions against PLpro are relatively low when compared to that 3CLpro.

RNA-dependent RNA polymerase (RdRp: NSP 12) is a conserved protein in coronavirus with a primary function in the coronavirus replication/transcription complex. Targeting Nsp12RdRp has been well documented for its little to no side effects on the host cell (Ruan et al., 2020). However, there has been no specific RdRp inhibitor till present. The crystal structure of RdRp was downloaded and refined for the protein–ligand docking process. Molecular docking results of RdRp following the extra-precision approach show the antiviral potential of the docked compounds.

Interestingly, acetosides and cynarosides demonstrated the highest binding energy of $-10.632$ and $-9.193$ kcal/mol, respectively. However, they were not selected for further analysis because of their rule of five violations. The lead compounds (lonchocarpol A, diplacol, and broussonol E) and the reference compound (dexamethasone) had a very low binding energy against the RdRp target; this may be attributed to the configuration and nature of the RdRp active site and its amino acid residues. The molecular docking results predicted that the lead compounds could stop the viral replication of SARS-CoV-2 through their inhibitory potential. Lonchocarpol A exhibited a docking score of $-7.017$ kcal/mol when docked into the active site of RdRp. Broussonol E has a binding energy of $-6.342$ kcal/mol, followed by diplacol with a binding energy of $-5.997$ kcal/mol. Although, other screened small molecules such as kuromanin, lauroside E, gallic acid, chlorogenic acid, and isovitexin demonstrated relatively high binding energy of $-7.687$, $-7.641$, $-7.122$, $-6.089$, and $-5.871$ kcal/mol, respectively.

### 3.6 MM-GBSA binding energy of top inhibitors

Molecular mechanics generalized Born surface area (MM-GBSA) has been widely explored as an advanced computational approach to analyze binding energy with an improved algorithm and solvation model. Compared to docking, post-scoring compounds using MM-GBSA have been demonstrated to show a relatively lower binding energy than the docking process.
correlate better to their reported binding affinity of docked complexes (Greenidge et al., 2013; Tripathi et al., 2013). The MM-GBSA method is more accurate in estimating the free binding energies of protein–ligand complexes than docking scores. A post-docking MM/GBSA analysis of the docked complexes was $-55.562$, $-49.137$, $-46.628$, and $-39.605$ kcal/mol for lonchocarpol A, broussonol E, diplacol, and dexamethasone, respectively, as shown in Figure 4. The post-simulation MM/GBSA, which further validates the binding affinity of the compounds, shows similar binding energy to the post-docking analysis.

3.7 Drug likeness and toxicity descriptors prediction

Pharmacokinetic properties of the top four lead potential antiviral flavonoids were predicted, studied, and tabulated, as shown in Table 7. It is clear that except (-) lonchocarpol A and broussonol E, none penetrated the blood–brain barrier. Under the adsorption and distribution, the Caco-2 permeability of the lead compounds shows that lonchocarpol A and dexamethasone showed positive ions of Caco-2 permeability whereas diplacol and broussonol E showed negative ion of Caco-2 permeability.

The action of the four lead compounds on the P-glycoprotein (substrate) showed that lonchocarpol A, diplacol, and broussonol E are non-substrate. In contrast, only dexamethasone showed the level of a substrate to the glycoprotein. Lonchocarpol A, diplacol, and broussonol E are promising glycoprotein inhibitors from COVID-19, whereas only dexamethasone shows its non-inhibiting property. About the LogS (aqueous solubility), broussonol E has the least solubility with $-3.567$, followed by dexamethasone with $-3.703$, which is greater than broussonol E; lonchocarpol A has a solubility value in the aqueous range of $-3.925$, and the highest solubility value out of the four lead compounds in the aqueous state is diplacol with a value of $-4.285$. All the compound complexes exhibit non-inhibitor on renal organic cation transporter 2 (OCT2), except dexamethasone, which shows inhibiting properties.

For the metabolism, the CYP450 2C9 (substrate) and CYP450 2D6 (substrate) showed that all the four lead compounds are non-substrate in nature, and CYP450 2D6 (inhibition) showed the lead compounds as non-inhibitors. CYP450 2C9 (inhibition), CYP450 1A2, and CYP450 2C19 showed that three lead compounds are natural
inhibitors, whereas only dexamethasone was non-inhibitor in nature.

For the Ames toxicity, all three lead compounds are non-toxic, whereas only the diplacol is toxic. In the analysis of hERG inhibition and carcinogenicity, all lead compounds exhibit inhibiting properties and non-carcinogenic ability. The Rat LD50 is higher on lonchocarpol A with a value point of 2.705 and lower on broussonol with a value of 2.129. Thus, natural phytocompounds are not naturally occurring and reported negligible toxicity when tested in vitro; hence, it could be a promising drug candidate and can be tested in vitro then in vivo. Lethal doses (LD50) of all the natural compounds were higher when compared to chemical drugs, which denotes that even at a higher dosage natural compounds are less toxic than chemically synthesized drugs. Thus, chemical drugs are toxic from the pharmacokinetic predictions compared to natural compounds; natural compounds have shown potential against several diseases with the least side effects (Benfenati et al., 2009). The drug-likeness properties (Table 7) of the compounds show they are druggable compounds with no violations of Lipinski’s assessment and Verber’s rules (Table 8). Some classes of compounds have been reported for their false-positive results during virtual screening. These compounds are referred to as Pan-assay interference compounds (PAINS). Chemical compounds in this category have been found to target numerous biological targets rather than a specific target (Baell, 2016). Catechols, quinones, curcumin, and toxoflavin are common examples of PAINS (Baell and Walters, 2014). PAINS screening of our compounds was carried out using the SwissADME web server (Daina et al., 2017), and the results shows that diplacol and broussonol E are PAINS compounds because of their substructural motifs with catechols. However, lochocarpol A and the reference compound (dexamethasone) will produce specific molecular interactions because they do not belong to the PAINS class.

3.8 Molecular dynamics simulation of the complexes

Molecular dynamics (MD) simulation is an essential tool that helps in the study of macromolecules such as nucleosomes, ribosomes, membrane proteins, organic solids, and proteins–ligand complexes and has evolved rapidly over the last 4 decades because of advances in force fields, thanks to
the development of quantum physics and computational chemistry (Mekni phyto-drugs et al., 2019). The simulation is widely used to analyze the structure-to-function relationship of protein and protein–ligand complexes. The current generation molecular dynamics mimic the actual biological systems with a potential simulation period of up to 100 ns for each complex and their behavior in the order of nanoseconds with appropriate system configurations using high-speed supercomputers. It takes thousands to several million steps and involves intra- and interatomic interactions simulated simultaneously, for which supercomputers play a vital role in attaining so. It is essential to study the simulation in the order of shortest duration, preferably femtoseconds, because biomolecules’ structural and functional properties concern nano- and microseconds (Liang et al., 2007).

After the chemical profiling, the association of compound complexes was examined, and the dynamic stability of screened compounds was studied using MD simulation at 100 ns in terms of root-mean-square deviation, root-mean-square fluctuations, and molecular contacts (Figures 5–7). This was achieved with the aid of the Desmond module integrated into the Schrodinger suite. Analyzing the molecular dynamics simulation at the atomic level, all the compounds were relatively stable through the MD simulation period.

Lonchocarpol A-3CLpro complexes were stable within 0–50 ns (Figure 5). However, fluctuations were observed from 55 to 65 ns before the ligand retained its stability. The RMSF analysis explicitly shows that some amino acid residues (PHE3, ARG4, GLY138, and GLU255) contributed to the ligand fluctuations. Clearly, broussonol E was found to be stable when in a complex with the protein backbone, with subtle fluctuation recorded at 25–35 ns and 65–75 ns. The RMSF analysis of broussonol E shows its residue index and established molecular contacts largely dominated by water bridges and hydrophobic interactions. Diplacol demonstrated a varying degree of fluctuations between 0 and 20 ns. Interestingly, it was found to be very stable from 25 to 100 ns. However, a slight increase in the RMSD value of the ligand was observed toward the end (90–100 ns) of the simulation period, as shown in the trajectory. The reference compound (dexamethasone) peaked at RMSD 2.0 Å at 20 ns and established essential interaction profiling such as hydrogen and water bridges. Our results, herein, suggest that the binding of the compounds may prompt conformational alterations. In consistent with this, the analysis of
MM-GBSA with trajectory against MM-BGSA without the trajectory residue number showed that the compound complexes showed higher oscillations in backbone residues when compared to other complexes in the systems, as shown in the figure below. This is consistent with the docking results of the four lead compounds that showed the highest binding free energy to other compounds with the low or least binding free energy of −7.3 and −8.1 kcal/mol.

The establishment and immovability of H-bonds were inspected over the simulation period. H-bond features are essential in drug design and discovery due to their irreplaceable role in drug specificity, metabolism, and absorption (Kurczab, 2017). Figure 7 shows that the four lead compounds could establish at least one hydrogen bond with amino acid residues. Thus, the stability of complexes was maintained by H-bond formation with active site residues. According to the docking and MD simulation analyses, the four lead compounds showed good affinity toward COVID-19 compared to the other compounds. However, lonchocarpol A showed a high docking score (−8.644 kcal/mol) and formed pi-stacking interactions with the essential amino acids of the COVID-19 binding domain. The MD simulation of the complexes in the study was very helpful in analyzing the conformational stability and dynamics of the protein and protein–ligand complexes at different nanosecond time intervals, fluctuations, and their deviations from the reference structure on COVID-19.

4 Conclusion

This study used an integrated computational approach such as molecular docking, molecular dynamics simulation, and semi-empirical Hamiltonian mechanics to discover flavonoids that could serve as potential therapeutic agent against SARS-CoV-2 therapeutic targets 3CLpro, PLpro, spike glycoprotein, and RdRp. Following the virtual screening, three lead compounds (lonchocarpol A, diplacol, and broussonol E) were identified as novel inhibitors of 3CLpro among the selected targets of SARS-CoV-2 by evaluating binding energy and interaction poses, drug likeness, toxicity, and dynamic stability in comparison to a reference compound (dexamethasone). Molecular docking shows that the compounds have high binding energy, resulting in strong molecular complex formation with the molecular SARS-CoV-2 targets. An atomistic study of the protein–ligand interaction via the dynamics simulations shows that deviation in dynamic stability falls within an acceptable range. Therefore, the docked complexes can be considered stable based on intermolecular interactions. The semi-empirical Hamiltonian mechanics elucidated the lead compound polarizability and the high chemical reactivity toward the target receptors. As a result, we propose that the hit compound could serve as a benchmark for developing phytodrugs against COVID-19. However, we recommend further experiments via in vitro pharmacological inhibition and neutralization studies to be carried out to validate our claim to develop these compounds as inhibitors of the SARS-CoV-2 therapeutic targets.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding authors.

Author contributions

TB conceptualized and designed the study. TB, OC, and ATA wrote the original draft of the article. UO, IJ, GD, NA, AI, and OS performed compound library curation and high-throughput virtual screening. TB, EO, OI, AOA, ZT, GD, NA, AI, and IO performed molecular docking and pharmacokinetics study. TB, OC, ADA, and OS performed molecular dynamics simulation and post-simulation analysis. TB and PC performed quantum chemical calculation. UO, IJ, ADA, and EO assisted with data collection and editing. DO, GB, AFA, WA-M, MDW, and J-MS supervised the project. All authors read and approved the final manuscript.

Funding

This study was supported by the researchers of King Saud University, Riyadh, Saudi Arabia Supporting Project number (RSP- 2021/97). MDW thanks the French Agence Nationale de la Recherche and the Region Pays de la Loire for financial support on Covid-19 research (ANR Flash COVID 19 call –name: CoV-E-TARGET – grant number: 2020 07132). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R39), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflict of interest

MDW was employed by Smartox Biotechnology.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
A new molecular dynamics method. syndrome coronavirus infection: A retrospective cohort study. Ribavirin and interferon alfa-2a for severe Middle East respiratory syndrome coronavirus infection: A retrospective cohort study. 

A new molecular dynamics method. 

Absolute electronegativity and hardness correlated with molecular orbital theory. 

Hard and soft acids and bases-the evolution of a chemical concept. 

Generalized koppmans' theorem. 

Pillaiyar, T., Manickam, M., Namasiyam, V., Hayashi, Y., and Jung, S. H. (2016). An overview of severe acute respiratory syndrome coronavirus (SARSCoV) 3CL protease inhibitors: Peptidomimetics and small molecule chemotherapy. 

Ruan, Z., Liu, C., Gao, Y., He, Z., Huang, X., Jia, X., et al. (2020). SARS-CoV-2 and SARS-CoV. Virtual screening of potential inhibitors targeting RNA-dependent RNA polymerase activity (NSP12). 

Salvatore, M. J., King, A. B., Graham, A. C., Onishi, H. R., Bartizal, K. F., Abruzzo, G. K., et al. (1998). Antibacterial activity of lonicarpocaul.  

Sastry, G. M., Adhigirey, M., Day, T., Annambahmu, R., and Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. 

Schrödinger (2018). Schrödinger release 2018-4: Maestro schrödinger. New York, NY: LLC.

Schrödinger (2020). Schrödinger release 2020-2: Prime, schrödinger. New York, NY: LLC.

Seah, I., and Agrawal, R. (2020). Can the coronavirus disease 2019 (COVID-19) affect the eyes? A review of coronaviruses and ocular implications in humans and animals. 

Sylaja, B., Gunasekaran, S., and Srinivasan, S. (2017). Vibrational, NLO, NBO, NMR, frontier molecular orbital and molecular docking studies of diazepam. 

Tripathi, S. K., Muttineni, R., and Singh, S. K. (2013). Extra precision docking, free energy calculation and molecular dynamics simulation studies of CDK2 inhibitors. 

Vellingiri, B., Jayaramayya, K., Iyer, M., Narayanasamy, A., Govindasamy, V., Giridharan, B., et al. (2020). COVID-19: A promising cure for the global panic. 

Wang, S. Q., Du, Q. S., Zhao, K., Li, A. X., Wei, D. Q., and Chou, K. C. (2007). Virtual screening for finding natural inhibitor against cathepsin-L for SARS therapy. 

World Health Organization (2004). SARS: Clinical trials on treatment using a combination of traditional Chinese medicine and western medicine. Beijing: World Health Organization.

Wu, C. Y., Jan, J. T., Ma, S. H., Kuo, C. J., Juan, H. F., Cheng, Y. S. E., et al. (2004). Small molecules targeting severe acute respiratory syndrome human coronavirus. 

Xu, C., Ke, Z., Liu, C., Wang, Z., Liu, D., Zhang, L., et al. (2020). Systemic in silico screening in drug discovery for Coronavirus Disease (COVID-19) with an online interactive web server. 

Yu, M.-S., Lee, J., Lee, J. M., Kim, Y., Shin, Y.-W., Jee, J.-G., et al. (2012). Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, Np31. 

Zumla, A., Chan, J. F., Azhar, E. I., Hui, D. S., and Yuen, K. Y. (2016). Corona viruses-drug discovery and therapeutic options. 

and nonlinear optical properties: A quantum chemical approach. 

Quantum mechanical study of the structure and spectroscopic (FTIR, FT-Raman, first-order hyperpolarizability and NBO analysis of 1, 2-benzoxazol-3-ylmethane sulfonamide. 

Raman), Quantum mechanical study of the structure and spectroscopic (FTIR, FT-