The inhibitory effect of some natural bioactive compounds against SARS-CoV-2 main protease: insights from molecular docking analysis and molecular dynamic simulation

Doaa A. Abdelrheema, Shimaa A. Ahmeda, H. R. Abd El-Mageedb, Hussein S. Mohamedc, Aziz A. Rahmand, Khaled N. M. Elsayede, and Sayed A. Ahmeda

aDepartment of Chemistry, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt; bMicro-Analysis and Environmental Research and Community Services Center, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt; cResearch Institute of Medicinal and Aromatic Plants (RIMAP), Beni-Suef University, Beni-Suef, Egypt; dDepartment of Pharmacy, University of Rajshahi, Rajshahi, Bangladesh; eDepartment of Botany, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt

ABSTRACT
This work aimed at evaluating the inhibitory effect of ten natural bioactive compounds (1–10) as potential inhibitors of SARS-CoV-2-3CL main protease (PDB ID: 6LU7) and SARS-CoV main proteases (PDB IDs: 2GTB and 3TNT) by molecular docking analysis. The inhibitory effect of all studied compounds was studied with compared to some proposed antiviral drugs which currently used in COVID-19 treatment such as chloroquine, hydroxychloroquine, azithromycin, remdesivir, baloxvir, lopinavir, and favipiravir. Homology modeling and sequence alignment was computed to evaluate the similarity between the SARS-CoV-2-3CL main protease and other SARS-CoV receptors. ADMET properties of all studied compounds were computed and reported. Also, molecular dynamic (MD) simulation was performed on the compound which has the highest binding affinity inside 6LU7 obtained from molecular docking analysis to study it is stability inside receptor in explicit water solvent. Based on molecular docking analysis, we found that caulerpin has the highest binding affinity inside all studied receptors compared to other bioactive compounds and studied drugs. Our homology modeling and sequence alignment showed that SARS-CoV main protease (PDB ID: 3TNT) shares high similarity with 3CLpro (96.00%). Also, ADMET properties confirmed that caulerpin obeys Lipinski’s rule and passes ADMET property, which make it a promising compound to act as a new safe natural drug against SARS-CoV-2-3CL main protease. Finally, MD simulation confirmed that the complex formed between caulerpin and 3CLpro is stable in water explicit and had no major effect on the flexibility of the protein throughout the simulations and provided a suitable basis for our study. Also, binding free energy between caulerpin and 6LU7 confirmed the efficacy of the caulerpin molecule against SARS-CoV-2 main protease. So, this study suggested that caulerpin could be used as a potential candidate in COVID-19 treatment.

ARTICLE HISTORY
Received 4 May 2020
Accepted 21 August 2020

KEYWORDS
COVID-19 virus protease; natural products; caulerpin; molecular docking; MD simulation

Introduction
Coronaviruses (CoV) are a great family of viruses varying from the common cold virus which is the lowest serious disease in this family to the more serious diseases such as MERS, CWID, and SARS. Also, their structure has a characteristic RNA genome. However, CoV are more familiar with animals; seven of them can induce the human respiratory system.[1] From December 2019, a new coronavirus called COVID-19 that was first appeared in Wuhan, Hubei Province of China, and now, is spreading rapidly across China and other parts of the world. COVID-19 has become a serious threat to the world public health, causing 244,792 deaths from 3,484,640 cases as of 3 May 2020 from the entire world. The current COVID-19 outbreak is caused by SARS-CoV-2 that has been known as the seventh member of the family of coronaviruses.[1] SARS-CoV-2 was established to have higher sequence homology toward SARS-CoV than that of MERS-CoV according to the whole genome sequence alignment analysis in different studies.[2]

There are four structural proteins characterize coronavirus genome, spike glycoprotein (S), matrix glycoprotein (M), nucleocapsid protein (N), and small envelope protein (E) were reported.[3] Additionally, 3CLpro is also known as Nsp5 of SARS-CoV-2 required for the maturation of coronaviruses. 3CLpro is first automatically cleaved from poly-proteins to produce mature enzymes, and then, further cleaves downstream Nsps at 11 sites to release Nsp4–Nsp16.[3] 3CLpro directly mediates the maturation of Nsps. So, 3CLpro is critical for the viral life cycle, making it an attractive target of anti-coronavirus drug evolution.[4,5] There are different studies that had been performed on the main proteases of SARS-Coronavirus (2GTB) to develop antiviral treatments of COVID-19 virus because it shares 96% similarity with 3CLpro of COVID-19.[6]
Caulerpa racemosa macroalgae and it was isolated from ancient times.[9,10] Caulerpin, the low toxic bisindole alkaloid medicinal plants have attracted significant attention because like antimicrobial, anticancer, anti-inflammatory, antiviral, pharmacological properties and effective biological activities that about 80% of people use natural compounds in the term use. The World Health Organization (WHO) evaluated different side effects and ineffectiveness in some cases for long-the only effective treatment. Chemical drugs are shown dif-ferent side effects and ineffectiveness in some cases for long-term use. The World Health Organization (WHO) evaluated that about 80% of people use natural compounds in the treatment field.[7] Actually, natural compounds have various and effective biological activities such as antimicrobial, anticancer, anti-inflammatory, and anti-diabetic.[8] Medicinal plants are an important source of molecules with various pharmacological properties and effective biological activities like antimicrobial, anticancer, anti-inflammatory, antiviral, and antidiabetic. So, bioactive compounds isolated from medicinal plants have attracted significant attention because there are safe and effective against different pathogenesis from ancient times.[9,10] Caulerpin, the low toxic bisindole alkaloid is a more common compound of the genus Caulerpa of green macroalgae and it was isolated from Caulerpa racemosa,[11] red alga Chondria armata,[12] and brown alga Sargassum platycarpum.[13] Caulerpin has various pharmacological properties and effective biological activities such as antitumor, anti-diabetic, anticancer, anti-larvicidal, anticorrosion, antitu-bercular, antimicrobial, spasmylocytic, antinoiceptive, plant growth regulatory activity, and anti-inflammatory.[13,14] Caulerpin exhibited antiviral activities against Chikungunya virus[15] and herpes simplex virus type 1.[16] Also, caulcrerin showed good inhibition activity against Alzheimer’s disease.[17] Due to the low-toxicity and biological activity of all studied compounds especially caulcrerin in different biological applications as summarized in Table 1. We expected that they could be used to develop formal inhibitors against COVID-19.

In this study, our homology modeling and sequence alignment of 2019-nCoV main protease (PDB ID: 3TNT) shares 96% similarity with 3CLpro of COVID-19. Also, homology modeling and sequence alignment of 2019-nCoV main protease show that SARS-CoV main protease (PDB ID: 2Gtb) shares 96.76% similarity with 3CLpro of COVID-19 as described in this study.[6] So, we study the inhibitory effect of some bioactive compounds obtained from natural sources against SARS-CoV-2-3CLpro and SARS-CoV main proteases (PDB IDs: 2Gtb and 3TNT). Due to their high similarity of 2Gtb and 3TNT with 3CLpro, inhibition of these receptors could be effective in 3CLpro treatment. Based on molecular docking analysis, we found that caulcrerin has the highest binding affinity against 6LU7, 3TNT, and 2Gtb receptors compared to other bioactive compounds and the proposed drugs. The pharmacogenetic and toxicity properties of all studied compounds also are computed and confirmed that most of all studied bioactive compounds especially, caulcrerin obey

| Compounds        | Sources                        | Species name                                      | Biological activity                          |
|------------------|--------------------------------|--------------------------------------------------|---------------------------------------------|
| Oleic acid (1)   | Terrestrial plant              | Zaleya decandra[8]                               | Antimicrobial,[22] anticancer,[23] and antiviral activities[24] |
|                  | Red macroalgae                 | Ceramium virgatum[16]                             |                                             |
|                  | Green macroalgae               | Ulva intestinalis[18]                             |                                             |
|                  | Brown macroalgae               | Fucus sp.[19]                                    |                                             |
|                  | Microalgae                     | Chlorella vulgaris[20]                            |                                             |
|                  | Endophytic fungi form plant    | Torreya Grandis[21]                              |                                             |
|                  | Vegetable oils                 | Hazelnu[22]                                      |                                             |
| Saringosterols (2,3) | Terrestrial plant             | Styrchons spinosa[25]                             | Antitrypanosomal[25]                        |
|                  | Red macroalgae                 | Acantophora spicifera[26]                        |                                             |
|                  | Green macroalgae               | Cladophora fasciculans[27]                       |                                             |
|                  | Brown macroalgae               | Sargassum muticum[28]                            |                                             |
| β-sitosterol (4)  | Terrestrial plant              | Synadenium glaucescens[31]                       | Anti-obesity,[26] a novel selective LXRβ agonist,[29] and anticaner activities[30] |
|                  | Red macroalgae                 | Eucheuma cottonii[32]                            |                                             |
|                  | Green macroalgae               | Ulva fasciata[33]                                |                                             |
|                  | Brown macroalgae               | Sargassum glaucescens[34]                        |                                             |
|                  | Microalgae                     | Nannochloropsis[35]                              |                                             |
| Glycoglycerolipids (5,9) | Terrestrial plant              | Soybean[39]                                      | Accumulation inhibition,[39] antitumor, antiviral, anti-inflammatory and anticaner activities[30,43] |
|                  | Red macroalgae                 | Exophyllum wenti[40]                             |                                             |
|                  | Brown macroalgae Cyanobacteria | Sargassum honeri[41] Phormidium sp.[42]          | Antibiotic effect against E. coli and Bacillus subtilis[44] |
| Kjellmanianone(6) | Brown macroalgae              | Sargassum noahouensi[8]                          |                                             |
| Lolirolide (7)    | Terrestrial plant              | Cancora decussata[46]                            | Antioxidant and a cell protective effect on a monkey kidney fibroblast cell line,[47] and anticaner, antibacterial and antifungal activities[46] |
|                  | Brown macroalgae               | Sargassum noahouensi[44]                         |                                             |
| Hexadecanoic acid (8) | Terrestrial plant             | Canthium parviflorum[48]                        | Antimicrobial, anti-inflammatory, antioxidant, hypcholeosteroemia, pesticide, hemolytic and S-Aldue reducator inhibitor,[50] and antiirritiv activity[53] |
| Caulerpin (10)    | Green macroalgae               | Caulerpa racemosa[11]                            | Antitumor, anti-diabetic, anticancer, anti- larvicidal, anticorrosion, anti-herpes, antitubercular, antimicrobial, cytotoxic, antiviral, spasmylocytic, antinoiceptive, anti-Alzheimer’s disease, plant growth regulatory activity and anti-inflammatory activities[13,17] |
|                  | Red macroalgae                 | Chondria armata[41]                              |                                             |
|                  | Brown macroalgae               | Sargassum platycarpum[13]                        |                                             |
Lipinski’s rule and pass ADME property. Finally, molecular dynamic (MD) simulation confirmed that the complex formed between caulerpin and 6LU7 is stable in water explicit and had no major effect on the flexibility of the protein throughout the simulations and provided a suitable basis for our study. Based on this work, we believe that caulerpin could be used to develop an antiviral molecule against COVID-19. Thus, this compound could be a potential candidate for *in vitro* and *in vivo* antiviral studies followed by clinical COVID-19 treatment. It is worth mentioning that there is not any work in literature discusses the inhibition effect of all studied compounds against COVID-19. Also, most of all related work in this field studied the inhibition effect against only COVID-19 receptors, but in this work we study the inhibition activity of all studied compounds against COVID-19 main protease and two SARS-CoV main proteases receptors. Finally, we also build homology modeling and sequence alignment on SARS-CoV-2-3CL main protease to show its similarity with other SARS-CoV receptors. While, most of all related work in this field did not study it.

**Material and methods**

*Structures and biological activity of all studied bioactive compounds*

In our study, we choose ten public bioactive compounds that already are abundant and isolated from different species of plants or marine algae to study the inhibitory effect of them against main proteases of CoV-2 (3CLpro) with (PDB ID: 6LU7) and main proteases of SARS-CoV (Mpro) with (PDB IDs: 2GTB and 3TNT). According to literature survey, these compounds showed a broad spectrum of biological activities like antimicrobial, antioxidant, anticancer, antiviral, and anti-inflammatory that making them an attractive target to evaluate their potential to become potential candidate inhibitors against SARS-CoV-2-3CLpro. All studied bioactive compounds in this work are shown in Figure 1. Also, their sources, species names and biological activities of all studied compounds are illustrated in Table 1.

**Preparation of the SARS-Cov-2-3CLpro structure and inhibitors compounds**

The crystal structures of SARS-CoV-2-3CLpro (PDB code: 6LU7) and main proteases of SARS-Coronavirus (Mpro) with (PDB IDs: 2GTB and 3TNT) were downloaded from the Protein Data Bank (www.pdb.org), and any heteroatoms and water molecules were removed before molecular docking studies. The 3-dimensional (3D) structures of all studied compounds and some proposed antiviral drugs (bioactive compounds, chloroquine, hydroxychloroquine, azithromycin, remdesivir, baloxvir, lopinavir, and favipiravir) were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/), in .sdf format. PubChem is a chemical substance and biological activities repository consisting of three databases, including substance, compound, and bioassay databases. Determination of the amino acids in the active site of a protein was determined using the Biovia Discovery Studio software[52] to analyze the Grid box and docking evaluation results.

**Homology modeling and sequence alignment of 2019-nCoV main protease**

Homology modeling is performed for the 2019-nCoV main protease. The crystal structure of SARS-CoV-2-3CLpro was used as a template (PDB ID: 6LU7). Therefore, sequence similarity searches were performed by using BLASTp analysis which showed suitable templates for the homology modeling. Target sequences of the SARS-CoV-2-3CLpro were created according to the build homology models.
Implicit in the Discovery Studio software as described in this study.\textsuperscript{[53]}

**Molecular docking**

All Ligands in this study (bioactive compounds, chloroquine, hydroxychloroquine, azithromycin, remdesivir, baloxvir, lopinavir, and favipiravir) were optimized before docking by Avogadro version 1.2 with Force Field type MMFF94 and saved in .pdb format. Discovery Studio used for protein optimization, by removing water and other atoms to prepare protein for the docking process. Molecular docking between ligands and 6LU7 was carried out using Auto Dock Tools (ADT) graphical user interface supported by MGL tools. Polar hydrogen was added and atomic charges were computed by Kollman and Gasteiger method. We defined a grid size with 60 Å × 60 Å × 60 Å for two receptors and the Lamarckian Genetic Algorithm (LGA) was assigned to carry out the molecular docking process, as described in this study.\textsuperscript{[54]} The output of the Auto Dock was further analyzed with discovery studio program and Pymol version 1.7.4.5 software package.

**Analysis of drug likeness and ADMET properties of all studied bioactive compounds**

The drug likeness prediction of all studied bioactive compounds was carried out by Lipinski filter (http://www.scfbio-
iitd.res.in/software/drugdesign/lipinski.jsp), according to which an orally active drug should comply with a minimum of four of the five laid down criteria for drug likeness namely: molecular mass, cLogP, hydrogen donor and acceptor, and molar refractive index.[55] Furthermore, the pharmacokinetic properties like absorption, distribution, metabolism, excretion, and the toxicity of all studied compounds were predicted utilizing admetSAR database (http://lmmd.ecust.edu.cn/admetsar1/predict).[56]

**Table 2.** Protein target structures of 6LU7 and 3TNT and active site amino acids.

| PDB ID | Macromolecule | Active site |
|--------|---------------|-------------|
| 6LU7   | GLN 107, ARG 107, GLN 127, THR 169, VAL125, LYS 5, ARG 105, GLN 127 THR 111, GLY 143, SER 144, CYS 145, GLN 192, GLU 166 LYS 137, ARG 298, GLN 110, LYS 137, GLU 288 |
| 3TNT   | VAL125, TYR126, GLY127, PHE140, GLN 107, ARG 107, GLN 127, THR 169, VAL125, LYS 5, ARG 105, GLN 127 THR 111, GLY 143, SER 144, CYS 145, GLN 192, GLU 166 LYS 137, ARG 298, GLN 110, LYS 137, GLU 288 ARG188, GLN189, GLN192, ALA198, LYS236, TYR237, GLN273 |

**MD simulation**

The structure of the highest binding complex obtained from molecular docking study between compound 10 (caulerpin) and 6LU7 was prepared for MD simulation using slandered dynamic cascade implicit in discovery studio. The MD simulation of SARS-CoV-2-3CLpro- inhibitor complex was carried out at 30 ns using CHARMm force field for all atoms in the complex. The simulation started by solvating the complex in triclinic box using TIP3P water model. The
counter ions were added to the neutralized the system. In which caulerpin–6LU7 complex with solvated in 6506 water molecules and neutralized by 20 sodium and 17 chlorides as counter ions. Periodic boundary conditions were used. Throughout the simulation, the studied system is maintained at a temperature of 300 K with constant pressure. Energy minimization was done for 50,000 steps. The trajectories were collected for every nanosecond to get insights into the interactions at the atomistic level. All MD protocol was carried out according to this study. The complexes result from MD simulation was analyzed for root mean square deviation (RMSD), root mean square fluctuations (RMSF) and hydrogen bonds analysis.

**MM-PBSA binding free energy calculations**

The binding free energy of protein and ligand complexes can be calculated by combining the molecular Mechanic/Poisson–Boltzmann Surface Area (MM-PBSA) with MD. The MD scripts were extracted to perform MM-PBSA calculations. The binding free energy provides an overview of the biomolecular interactions between protein and ligand. The binding energy constitutes of potential energy, polar and non-polar solvation energies. The MM-PBSA binding free energies were calculated by utilizing nanoscale molecular dynamics (NAMD) and visual molecular dynamics (VMD) software programs according to this study.

**Results and discussions**

**Homology modeling and sequence alignment of 2019-nCoV main protease**

It is noticed that SARS-CoV and SARS-CoV-2-3CLpro share remarkable 96.00% sequence alignment among all other human coronaviruses as shown in Figure 2, based on our homology modeling and sequence alignment of 2019-nCoV main protease. The crystal structure of SARS-CoV-2-3CLpro (PDB ID: 6LU7) is highly similar to its SARS-CoV sister (PDB ID: 3TNT) at high resolution with 1.59 Å. Multiple additional sequencing studies have been performed for SARS-CoV-2, including a landmark preprint, which suggested renaming 2019-nCoV to SARS-CoV-2 on based on results similar to ours. The highly conserved region of the structural elements was found, the least PDF total energy of 6LU7 and 3TNT are 2728.98 and 2881.43, respectively, that is, a reliable statistical potential to assess the quality of homology models in protein structure prediction. Also, DOPE score of 6LU7 and 3TNT are 70754.34 and 70644.22, respectively. SARS-CoV-2 and SARS-Cov-3CLpro have nine α-helices and 13 β-strands which make up three distinct domains, i.e., domain I, domain II, and domain III. All CoV proteases family consist of three domain, in which Domains I (residues 8–101) and II (residues 102–184) have one antiparallel β-barrel, that resemble the trypsin-like serine proteases structure. Domain III (residues 201–306) consists of 5 α-helices (25–99), that are connected by a long loop (residues185–200) with domain II as reported in this study. There are only four mutated residues (Phe 13, Asn 65, Ala 94, and Val 35) between SARS-CoV-2 and SARS-CoV-2-3CLpro as shown in Figure 3. The Ramachandran plot built also, by discovery studio shows that 100% of the residues in the allowed regions, 97.0% in the most favored region as shown in Figure 4. Additionally, 88.2% of the residues have averaged 3D–1D score ≤ 0.3 based on the Verify 3D software, while the overall quality factor of ERRAT is 96.0%. Our homology modeling and sequence alignment of 2019-nCoV main protease is the best agreement with this study, which confirmed that the structure of SARS-CoV (PDB Code 3TNT) and SARS-CoV-2-3CLpro (PDB Code 6LU7) share remarkable 96.75% sequence alignment. Due to the high sequence alignment between 3TNT and 2GTB with 6LU7, 3TNT, and 2GTB could be a potential drug target for 2019-nCoV and the inhibition of 3TNT and 2GTB protease would help to restrict the viral maturation thereby decreasing the SARS-CoV-2 infection in humans.

**Molecular docking**

The amino acids found in the active site pockets of 6LU7, 3TNT, and 2GTB and SARS-CoV-2, were elucidated from

---

**Table 3. Molecular docking analysis of studied compounds (1–10) and some antiviral drugs against 6LU7.**

| Ligand          | Binding energy (ΔG) k.cal/mol | Inhibition efficiency | Inhibition constant (µM) | Intermolecular energy | VDW-H bond desolvation energy |
|-----------------|-------------------------------|-----------------------|--------------------------|-----------------------|-----------------------------|
| 1               | –7.24                         | –0.26                 | 940                      | –11.03                | –16.32                      |
| 2               | –7.47                         | –0.19                 | 820                      | –11.21                | –15.32                      |
| 3               | –7.55                         | –0.32                 | 765                      | –11.43                | –14.43                      |
| 4               | –8.02                         | –0.25                 | 523                      | –12.54                | –15.55                      |
| 5               | –8.03                         | –0.29                 | 420                      | –13.42                | –14.43                      |
| 6               | –9.22                         | –0.18                 | 160                      | –15.65                | –14.43                      |
| 7               | –9.02                         | –0.24                 | 220                      | –14.34                | –13.54                      |
| 8               | –8.10                         | –0.40                 | 427                      | –12.13                | –11.31                      |
| 9               | –9.25                         | –0.28                 | 315                      | –14.65                | –12.88                      |
| 10              | –9.28                         | –0.20                 | 189                      | –16.32                | –15.43                      |
| Chloroquine     | –8.95                         | –0.20                 | 525                      | –16.43                | –15.22                      |
| Hydroxychloroquine | –9.23                | –0.23                 | 418                      | –15.32                | –14.43                      |
| Azithromycin    | –8.55                         | –0.22                 | 330                      | –15.54                | –14.13                      |
| Remdesivir      | –9.24                         | –0.24                 | 380                      | –16.32                | –15.43                      |
| Baloxavir       | –9.18                         | –0.18                 | 128                      | –14.65                | –13.22                      |
| Lopinavir       | –9.26                         | –0.29                 | 320                      | –13.62                | –12.86                      |
| Favipiravir     | –9.23                         | –0.27                 | 515                      | –17.43                | –16.25                      |
PDB files by discovery studio and summarized in Table 2. Table 3 displays the molecular docking analysis of all studied bioactive compounds and some proposed antiviral drugs against 6LU7, including binding energy, ligand efficiency, inhibition constant, intermolecular energy, and van der Waals (VDW)-H Bond desolvation energy. Figure 5 is shown the best docking poses of the studied bioactive compounds (1–10) inside 6LU7. Also, Figure 6 displays docking analysis visualization of 6LU7 binding with all studied compounds in which the yellow dots show H-bonds formed between bioactive compounds and 6LU7 residues. The Number of H-bonds, H-bonding Residues and H-bonding distance produced from docking for all studied bioactive compounds against 6LU7 shown in Table 4. Chloroquine, hydroxychloroquine, azithromycin, remdesivir, baloxvir, lopinavir, and favipiravir also docked inside 6LU7 using the same protocol applied on all studied compounds. An in silico analysis study showed that only compound 10 can inhibit 6LU7 which have the higher binding energies and inhibition constants and higher Number of H-bonds with 6LU7 amino acid residues with compared to chloroquine, hydroxychloroquine, azithromycin, remdesivir, baloxvir, lopinavir, and favipiravir inhibitors as shown in Tables 3 and 4. Compound 10 exhibits the highest binding energy (−9.28 k.cal/mol) with compared to other compounds and formed three hydrogen bond interactions with LYS 137, GLU 288, and LYS 5 of 6LU7 amino acids.
acid residues with distance 2.25, 2.27, and 2.38 Å, respectively. Also, all studied bioactive compounds are docked inside 3TNT and 2GTB and the obtained binding energy are summarized in Table 5. Also, compound 10 shows the highest binding energy compared to all studied bioactive compounds and antiviral drugs. So, compound 10 may act as a potential inhibitor of SARS-CoV-2-3CLpro and SARS-CoV-2.

**Analysis of drug likeness and ADMET properties of all studied bioactive compounds**

Lipinski’s rule of five is commonly utilized in development and drug design to expect oral bioavailability of drug molecules. Lipinski’s rule was established based on five rules to compute the ability of the compound to act as an orally active drug was calculated and shown in Table 6. So, orally active drugs must have no more than one violation of the following standards: (i) octanol/water partition coefficient (log P) which measured the lipophilicity of a molecule must be not greater than five. (ii) A molecular weight (MW) less than 500 Da. (iii) not more than five hydrogen bond donors (nON). (iv) not more than 10 hydrogen bond acceptors (nOHN). The topological polar surface area (TPSA) is measured the bioavailability of the drug molecule. TPSA is closely related to the hydrogen bonding potential of a compound. TPSA of studied compounds was noticed in the

![Docking analysis visualization of 6LU7 binding with all studied compounds generated using Pymol software. The yellow dots show H-bonds.](image)
Table 4. The Number of H-bonds, H-bonding Residues, and H-bonding distance produced from docking for all studied bioactive compounds against 6LU7.

| Compounds | N. of H-bonds | H-bonding residues | H-bonding distance (Å) |
|-----------|---------------|--------------------|------------------------|
| 1         | 2             | GLN 107            | 2.24                   |
|           |               | ARG 107            | 2.09                   |
| 2         | 2             | GLN 127            | 2.32                   |
|           |               | THR 169            | 2.33                   |
| 3         | 2             | VAL125             | 2.24                   |
|           |               | LYS 5              | 2.35                   |
| 4         | 1             | ARG 105            | 2.41                   |
| 5         | 2             | GLN 127            | 2.17                   |
|           |               | THR 111            | 2.36                   |
| 6         | 3             | GLY 143            | 2.35                   |
|           |               | SER 144            | 2.39                   |
|           |               | CYS 145            | 2.29                   |
| 7         | 3             | GLN 192            | 2.27                   |
|           |               | GLU 166            | 2.30                   |
|           |               | CYS 145            | 2.33                   |
| 8         | 1             | LYS 137            | 2.35                   |
| 9         | 3             | THR 111            | 2.30                   |
|           |               | ARG 298            | 2.41                   |
|           |               | GLN 110            | 2.33                   |
| 10        | 3             | LYS 137            | 2.25                   |
|           |               | GLU 288            | 2.27                   |
|           |               | LYS 5              | 2.38                   |

Table 5. Molecular docking analysis of studied compounds (1–10) and some antiviral drugs against 3TNT and 2GTB.

| Ligand | Binding energy (ΔG) kcal/mol | Binding energy (ΔG) kcal/mol |
|--------|------------------------------|-------------------------------|
| 1      | −6.14                        | −7.24                         |
| 2      | −7.22                        | −7.37                         |
| 3      | −7.25                        | −7.44                         |
| 4      | −7.88                        | −7.65                         |
| 5      | −7.82                        | −7.64                         |
| 6      | −8.13                        | −8.09                         |
| 7      | −8.32                        | −8.14                         |
| 8      | −8.44                        | −8.37                         |
| 9      | −9.55                        | −9.44                         |
| 10     | −9.62                        | −9.77                         |

Chloroquine | −9.12 | −9.32 |
Hydroxychloroquine | −9.34 | −9.38 |
Azithromycin | −9.24 | −9.42 |
Remdesivir | −9.53 | −9.67 |
Baloxvir | −9.45 | −9.51 |
Lopinavir | −9.56 | −9.56 |
Favipiravir | −9.58 | −9.73 |

Table 6. Prediction of molecular properties descriptors of all studied bioactive compounds.

| Bioactive compounds | Mass | Hydrogen bond donor | Hydrogen bond acceptors | cLogP | Molar refractivity | TPSA  |
|---------------------|------|---------------------|-------------------------|-------|-------------------|-------|
| 1                   | 282.00 | 1                   | 2                       | 6.10  | 87.08             | 32.41   |
| 2                   | 428.00 | 2                   | 2                       | 7.11  | 135.68            | 150.55   |
| 3                   | 428.00 | 2                   | 2                       | 7.16  | 138.78            | 21.26   |
| 4                   | 368.00 | 0                   | 1                       | 0     | 0                 | 35.49   |
| 5                   | 750.00 | 4                   | 10                      | 9.20  | 207.42            | 38.42   |
| 6                   | 186.00 | 1                   | 5                       | −0.60 | 41.68             | 46.00   |
| 7                   | 196.00 | 1                   | 3                       | −1.40 | 51.60             | 150.88  |
| 8                   | 256.00 | 1                   | 2                       | 5.55  | 77.94             | 36.19   |
| 9                   | 712.00 | 4                   | 10                      | 3.39  | 187.92            | 25.87   |
| 10                  | 398.00 | 2                   | 4                       | 3.94  | 114.13            | 36.76   |

Table 7. Prediction of ADMET descriptors of all studied bioactive compounds.

| Compounds | BBB | HIA | Caco-2 permeability | CYP inhibitory promiscuity | AMES toxicity | Carcinogenicity | Rat acute toxicity LD50 mol/kg |
|-----------|-----|-----|---------------------|----------------------------|---------------|----------------|-----------------------------|
| 1         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 1.95                        |
| 2         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.29                        |
| 3         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.38                        |
| 4         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.60                        |
| 5         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.17                        |
| 6         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.18                        |
| 7         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.21                        |
| 8         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.75                        |
| 9         | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.70                        |
| 10        | BBB+ | HIA+ | Caco2+               | Low                        | Non-toxic     | Non-carcinogenic  | 2.23                        |
range of 25.87–153.50 Å and is well below the limit of 160 Å. It can be predicted that all studied bioactive compounds obeyed Lipinski’s rule of five and are likely to be orally active except compounds 5 and 9 as shown in Table 6. The database supports ADMET profiles which involve some features to study the ability of the studied compounds to act as drug leads such as Blood–brain barrier (BBB) penetration, human intestinal absorption (HIA), Caco-2 cell permeability, CYP inhibitory promiscuity, AMES toxicity, carcinogenicity, and rat acute toxicity LD50 are calculated and displayed in Table 7. As shown in Table 6, all studied compounds may cross blood brain barrier (BBB) and absorb in the human intestine (HIA) along are permeable for Cacoe2 cells, whereas, compound 5 showed a negative result for BBB, HIA, and Cacoe2 cell permeability. Cytochrome P450 (CYP) is a group of isozymes containing the metabolism of drugs, steroids, fatty acids, bile acids, and carcinogens. The results indicate that these studied compounds are non-substrate and non-inhibitor of CYP enzymes. In terms of AMES toxicity, all studied compounds were observed to be non-toxic. Carcinogenicity model indicated non-Carcinogenic nature of all studied compounds. Rat Acute Toxicity LD50 of all studied compounds was found between 1.95 and 2.75 mol/kg. The finding strongly provides the ability of most of all studied compounds to act as a drug, except compound 5 as shown in Table 7.

**MD simulation**

To confirm the docking results and get more insight into the stability of the ligand–protein complex, MD simulations
were carried out for the compound which has the highest binding affinity (compound 10) inside 6LU7 complex in the solvated states at 30 ns as shown in Figure 7. The results of MD simulations have been examined based on RMSD, RMSF, and the number of hydrogen bonds as a function of time.

**Root mean square deviation**

To examine the change in the protein dynamics and the conformational stability of the protein–ligand complexes, the protein complexed with compound 10 was subjected to 30 ns MD simulations. Standard dynamics cascade module implicit in DS software was employed to measure the RMSD and RMSF.

The RMSD measures the direct changes in the protein from the initial coordinates. The RMSD values of the protein backbone in complex with the potential inhibitors were computed concerning the initial structure as a frame reference (0 to 10 ns). The RMSD values steadily increased from 0 to 5 ns, and reached equilibration after that throughout the simulation period. The RMSD values for the studied complex showed oscillations between 3 to 5 ns indicating that the studied compound was adapting another conformation within the binding pocket as shown in Figure 8. The average RMSD values for the last 1 ns for the 6LU7 and caulerpin + 6LU7 complex, were 1.13 ± 0.07 and 1. 22 ± 0.05, respectively. Lower RMSD value of the caulerpin in the complex with 6LU7 indicates its stability and provided a suitable basis for our study.

**Root mean square fluctuation**

RMSF was measured based on the backbone atom of each amino acid residues and the plot of RMSF was used to depict the fluctuations at the residue level. RMSF plot as shown in Figure 9 of solvated 6LU7 and 6LU7 in complex with caulerpin during 30 ns MDs simulations exhibited a similar trend of residue fluctuation profile for both free receptors (6LU7) and the complexes with a low average RMSF. This trend in the RMSF plot for the complex indicates that binding caulerpin to the receptor was stable and had no major effect on the flexibility of the protein throughout the simulations. To explore more insights on the local protein flexibility, the time average of RMSF values of the 300 amino acids of 6LU7 in the absence and presence of the caulerpin over simulation period were calculated. The RMSF values for the three complexes suggested that the following residues (ALA206, VAL204, and LEU205) showed less fluctuation in the presence and presence of the caulerpin. The average RMSF values of these residues were 0.33 ± 0.06, 0.40 ± 0.07, and 0.43 ± 0.08 Å for 6LU7, and 0.37 ± 0.07, 0.48 ± 0.08, and 0.52 ± 0.09 Å for caulerpin–6LU7 complex.

**Hydrogen bond analysis and MM-PBSA binding free energy**

Hydrogen bonds are one of the essential elements answerable for the molecular interactions in biological systems. Hydrogen bonds provide the basis for molecular recognition and selectivity by imparting directionality and explicitness to molecular interactions. The protein–ligand interactions were guided by the changes in the secondary structures, which in turn were regulated by the hydrogen bonds. MD simulation provided different conformations in which a protein could be found in actual biological conditions. Each conformation of a protein is supposed to have its interaction pattern with the ligand. We calculated the number of hydrogen bonds formed during the complete run of MD simulations for selected complexes, as presented in Figure 10. As shown in Figure 10, in the caulerpin–6LU7 complex, the most number of conformations formed up to 10 hydrogen bonds during the simulation. A very few conformations showed less than 5 and greater than 10 hydrogen bonds. The bioactive caulerpin molecule was able to maintain strong interaction with the binding pocket of 6LU7 throughout the simulation period. We measured binding free energy at last 10 ns from MD simulation. The final binding energy is a cumulative...
sum of van der Wall, electrostatic, polar solvation, and SASA energy. Except for the polar solvation energy, all other forms of energy contributed favorably to the interaction between caulerpin molecule and 6LU7. The caulerpin molecule showed the binding free energy equal to $-256.44 \pm 27.55$ kJ/mol, SASA equal to $-28.58 \pm 4.32$ kJ/mol, $\Delta E_{\text{polar solvation}}$ equal to $-283.58 \pm 44.53$ kJ/mol, $\Delta E_{\text{Electrostatic}}$ equal to $-127.76 \pm 40.50$ kJ/mol, and $\Delta E_{\text{Van der Waal}}$ equal to $-384.44 \pm 34.40$ kJ/mol inside 6LU7. This confirmed the efficacy of the caulerpin molecule against SARS-CoV-2 main protease.

**Conclusion**

Recently, the rapidly spreading outbreak of COVID-19 has challenged the healthcare sector of the world. Numerous antiviral or other conventional drugs are being examined against COVID-19, but not yet positive results have come. To contribute to this fight against COVID-19, virtual screening-based molecular docking and MD simulation were investigated to identify novel and potential to inhibit SARS-CoV-2-3CLpro main protease. Based on our molecular docking analysis we found that among all studied compounds, caulerpin has the highest binding affinity against all studied receptors 6LU7, 3TNT, and 2GTB with compared to some proposed antiviral drug currently used in COVID-19 treatment. The ADMET and molecular descriptors properties strongly provide that most of all studied compounds obey Lipinski’s rule of five and exhibit significant biological activities especially, caulerpin compound. MD simulation confirmed the stability of caulerpin-6LU7 receptor in the presence of explicit water solvent. RMSD of the receptor–ligands complex has maintained the stability at around 1.8 Å and RMSD of caulerpin complexed with the protein are in the favorable range of within 1.2 Å and has remained stable during the simulations. The backbone atoms of the complex and free receptor show similar RMSF, indicating the stability of the caulerpin inside 6LU7 receptor. Also, the most number of conformations formed up to 10 hydrogen bonds during the simulation in caulerpin–6LU7 complex. A very few conformations showed less than 5 and greater than 10 hydrogen bonds. The bioactive caulerpin molecule was able to maintain strong interaction with the binding pocket of 6LU7 throughout the simulation period. Finally, the caulerpin molecule showed the binding free energy inside 6LU7 equal to $-256.44 \pm 27.55$ kJ/mol which confirmed the efficacy of this molecule against SARS-CoV-2 main protease. Hence, caulerpin is highly effective against SARS-CoV-2 main protease and can be explored further for drug repurposing against the successful inhibition of COVID-19.

**ORCID**

Hussein S. Mohamed [https://orcid.org/0000-0001-6524-9649](https://orcid.org/0000-0001-6524-9649)

**References**

[1] Li, Y.; Zhang, J.; Wang, N.; Li, H.; Shi, Y.; Guo, G.; Zou, Q. Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening. *bioRxiv* 2020.

[2] Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* 2020, 382, 727–733.

[3] Xu, X.; Chen, P.; Wang, J.; Feng, J.; Zhou, H.; Li, X.; Zhong, W.; Hao, P. Evolution of the Novel Coronavirus from the Ongoing Wuhan Outbreak and Modeling of Its Spike Protein for Risk of Human Transmission. *Sci. China Life Sci.* 2020, 63, 457–460. DOI: 10.1007/s11427-020-1637-5.

[4] Xia, B.; Kang, X. Activation and Maturation of SARS-CoV Main Protease. *Protein Cell* 2011, 2, 282–290. DOI: 10.1007/s13238-011-1034-1.

[5] Lu, I.-L.; Mahindroo, N.; Liang, P.-H.; Peng, Y.-H.; Kuo, C.-J.; Tsai, K.-C.; Hsieh, H.-P.; Chao, Y.-S.; Wu, S.-Y. Structure-Based Drug Design and Structural Biology Study of Novel Nonpeptide Inhibitors of Severe Acute Respiratory Syndrome Coronavirus Main Protease. *J. Med. Chem.* 2006, 49, 5154–5161. DOI: 10.1021/jm060207o.

[6] Zhang, H.; Saravanan, K. M.; Yang, Y.; Hossain, M. T.; Li, J.; Ren, X.; Wei, Y. Deep Learning Based Drug Screening for Novel Coronavirus 2019-nCoV. Preprints. 2020.

[7] Baell, J. B. Feeling Nature’s PAINS: Natural Products, Natural Product Drugs, and Pan Assay Interference Compounds (PAINS). *J. Nat. Prod.* 2016, 79, 616–628. DOI: 10.1021/acs.jnatprod.5b00947.

[8] Malavizhi, D.; Anusooriya, P.; Meenakshi, P.; Sundaram, S.; Oirere, E.; Gopalakrishnan, V. K. Isolation, Structural Characterization of Oleic Acid from *Zaleya decandra* Root.
