Plant-derived chemicals as potential inhibitors of SARS-CoV-2 main protease (6LU7), a virtual screening study

Mohsen Sisakht1 | Amir Mahmoodzadeh2 | Maryam Darabian3

1Biochemistry Department, Shiraz University of Medical Sciences, Shiraz, Iran
2Kermanshah University of Medical Sciences Vice Chancellor for Research and Technology, Medical Biology Research Center, Health Technology Institute, Kermanshah, Iran
3Shiraz Medical School, Department of Radiology, Faculty of Paramedical Sciences, Shiraz University of Medical Sciences, Shiraz, Iran

Correspondence
Amir Mahmoodzadeh, Medical Biology Research Center, Kermanshah University of Medical Sciences, P. O. Box 67155-1616, Kermanshah, Iran.
Email: amir.neurochem@gmail.com

SARS-CoV-2 has caused millions of infections and more than 700,000 deaths. Taking the urgent need to find new therapeutics for coronavirus disease 2019 (COVID-19), a dataset of plant-based natural compounds was selected for the screening of antiviral activity. The viral 3-chymotrypsin-like cysteine protease (Mpro, 3CLpro) was selected as the target. Molecular docking was performed on 2,845 phytochemicals to estimate the spatial affinity for the active sites of the enzyme. The ADMET screening was used for the pharmacological and physicochemical properties of the hit compounds. Nelfinavir and Lopinavir were used as control for binding energy comparison. The top 10 hits, based on the binding energy (Kcal/mol), were Ginkgolide M (−11.2), Mezerein (−11), Tubocurarine (−10.9), Gnidicin (−10.4), Glycobismine A (−10.4), Sciadopitysin Z-10.2, Gnididin (−9.2), Glycobismine A (−10.4), Sciadopitysin (−10.2), Gnididin (−9.20, Emetine (−8.7), Vitexin (−8.3), Calophyllolide (−8.3), and 6-(3,3-Dimethylallyl)galangin (−7.9). The binding energy for nelfinavir and lopinavir were −9.1 and −8.4, respectively. Interestingly, some of these natural products were previously shown to possess antiviral properties against various viruses, such as HIV, Zika, and Ebola viruses. Herein, we suggest several phytochemicals as the inhibitors of the main protease of SARS-CoV-2 that could be used in the fight against COVID-19.

KEYWORDS
ADMET, COVID-19, molecular docking, phytochemicals, protease, SARS-CoV-2

1 INTRODUCTION

The new strain of the coronavirus family, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is the etiological agent responsible for the COVID-19 pandemic (WHO, 2020). The number of infected people has been dramatically on the rise, and so far, more than 45 million confirmed cases are reported, and the total death toll surpassed 700,000, October 29, 2020. Due to the spread of the COVID-19 outbreak, and associated global health concerns, and social and economic impacts, there is an urgent need to introduce efficient treatments. Despite worldwide attempts, unfortunately, no specific and effective therapy is available (Khaerunnisa, Kurniawan, Awaluddin, Suhartati, & Soetjipto, 2020). The current treatments are mainly focused on supportive care and symptom relief. Due to the lack of any effective and specific therapies for COVID-19, considerable adverse effects of the available drugs, and an abruptly increasing number of infected people, there is an urgent need for finding new therapeutic approaches against COVID-19. Natural products because of their proven health benefits, less adverse effects could be considered as new agents for decreasing the prognosis of the disease and inhibition of the viral life cycle.

Coronavirus family (CoVs) is single-stranded positive-sense RNA viruses, infecting humans and animals and has been regarded as a global threat to public health (Hui et al., 2020). The most common symptoms are usually fever, dry cough, and tiredness (Adhikari et al., 2020). Several infectious epidemics have occurred by this family of viruses, SARS-CoV, (2000–2004), and the middle east respiratory syndrome coronavirus (MERS-CoV, 2012-present) (Liu et al., 2020).
Whole-genome analysis has shown that SARS-CoV-2 and SARS belong to clade b of the genus Beta-coronavirus, and share 82% sequence homology (Chan et al., 2020; Hasan et al., 2020).

Upon transcription of the SARS-CoV-2, an ~800 kDa polypeptide is produced, subsequently is proteolyzed and generates non-structural proteins. The proteolytic processing is mediated by papain-like protease (PLpro), and 3-chymotrypsin-like protease (3CLpro, also called Mpro); both are cysteine proteases. There is a high structural similarity for this protein among SARS-CoV and SARS-CoV-2, 96% sequence identity for 3CLpro (Anand, Ziebuhr, Wadhwani, Mesters, & Hilgenfeld, 2003; Zhang et al., 2020). The proteolytic activity of 3CLpro is vital for the production of structural and nonstructural proteins, replication, and packaging of the new virus. This protein was successfully crystallized, and very soon being investigated as a potential target in drug design studies (Wu et al., 2020).

The amino acids found in the active site of the enzyme are THR24, THR26, PHE140, ASN142, GLY143, CYS145, IS163, HIS164, GLU166, HIS172 (Khaerunnisa et al., 2020). Earlier efforts to target SARS-CoV resulted in the identification of several 3CLpro inhibitors targeting the catalytic dyad of the protein defined by His41 and Cys145 residues (Paasche et al., 2014). Here, we used a diverse library of phytochemicals, based on the antiviral, other health benefits, and less adverse effect, for finding possible inhibitors of 3CLpro. The molecular docking and Lipinski Rule of Five were calculated. The binding energy of the compound to the active site of the enzyme was compared to lopinavir/ritonavir, since these are currently used in the clinic as one of the possible therapies of COVID-19 (Arabi et al., 2020; Hung et al., 2020).

Several studies have shown the therapeutic value of targeting of 3CLpro in SARS-CoV and MERS-CoV (Liang et al., 2018). 3CLpro not only is involved in the life cycle of the virus but also has interactions with the host’s immune system. It suppresses host’s immune response through deubiquitinating of interferon regulatory factor 3 (IRF3), and inactivation of nuclear factor κ-light-chain-enhancer (NF-κB), in the activated B cells (Baez-Santos, St John, & Mesecar, 2015). So, it seems that inhibitors of 3CLpro may benefit the host through inhibition of viral replication and suppressing its interaction with host’s immune system (Zhang et al., 2020). Several protease inhibitors have been introduced in the treatment of COVID-19. For example, lopinavir, as an inhibitor of HIV virus inhibitor, has shown that completely recover signs of COVID-19 related pneumonia (Wu et al., 2020; Wu, Liu, et al., 2020). The critical role of the enzyme in the management of the viral life cycle, and its uniqueness, no similar enzyme is found in the host cells, has made 3CLpro the most attractive target for the treatment of COVID-19.

The genome of SARS-CoV-2, positive-sense, single-stranded RNA, encodes non-structural proteins (such as 3-chymotrypsin-like protease, papain-like protease, helicase, and RNA-dependent RNA polymerase [RdRp]), structural proteins (such as spike glycoprotein) and accessory proteins (Li & De Clercq, 2020). Based on the previous studies, three distinct targets could be regarded as the potential targets ineffectively combat CoVs, including Angiotensin-converting enzyme II (ACE2) entry receptor, the RdRp and the 3CLpro proteins (Adhikari et al., 2020). Inhibitors against ACE2 is probably associated with severe side effects due to the possible interference with the host’s physiological conditions (Han, Penn-Nicholson, & Cho, 2006). Also, RdRp inhibitors are associated with low potency and specificity (Zumla, Chan, Azhar, Hui, & Yuen, 2016). The 3CLpro is crucial in the proteolytic maturation of the virus and has been considered as a potential key target to halt the viral life cycle (Liang et al., 2018). 3CLpro is the main protease found in SARS-CoV 2, which has been structured and repositioned in PDB (6LU7) and can be accessed by the public in February 2020 (Zhang et al., 2020). The amino acids found in the active site of the enzyme are THR24, THR26, PHE140, ASN142, GLY143, CYS145, IS163, HIS164, GLU166, HIS172 (Khaerunnisa et al., 2020). Earlier efforts to target SARS-CoV resulted in identifying of several 3CLpro inhibitors targeting the catalytic dyad of the protein defined by His41 and Cys145 residues (Paasche et al., 2014).

It has shown that nelfinavir can inhibit CoVs in vivo and in vitro models (IC50 = 0.048 μM) (Hsieh et al., 2010). Fortunately, specific inhibitors against 3CLpro, such as lopinavir/nelfinavir, have led to promising results (Lin, Shen, He, Li, & Guo, 2020). Taking the promising outcomes of previous studies, the availability of X-ray crystal structure of the enzyme and crucial role in the viral life cycle, 3CLpro could be regarded as the most promising drug target in the fight against COVID-19 (Park et al., 2016). So, in this study, we selected lopinavir/nelfinavir, as the inhibitors of 3CLpro and docked them as controls.

Plant-derived compounds, also known as phytochemicals, are secondary metabolites in the plants, involved in plant’s defense system against free radicals, viruses, bacteria, and fungi (Barbieri et al., 2017). These bioactive compounds possess various health benefits, such as antioxidant, antiinflammatory, cytoprotective, anticancer, antimicrobial, and antiviral properties (Barbieri et al., 2017; Chikhale et al., 2020; Islam et al., 2020; Oguntibeju, Aboua, & Omodaniyi, 2020). It has been reported that antiviral properties of phytochemicals are exerted through a wide range of mechanisms, including inhibition of the DNA/RNA metabolism or obstructing viral entry (Ben-Shabat, Yarmolinsky, Porat, & Dahan, 2020; Liu & Du, 2012). It has shown that the plant-based compounds have promising anti-SARS-CoV-2 properties, and their effective antiviral potential and minimum side effects have opened new opportunities (Khaerunnisa et al., 2020; Ubani et al., 2020). For example, several phenolic compounds extracted from Isatis indigotica have shown to have a remarkable inhibitory effect on 3CLpro with IC50 values, ranged from 8.3 to 1,210 μM (Lin et al., 2005).

This study aims at screening a set of phytochemicals to find potent inhibitors of the main protease of SARS-CoV-2, as a possible approach in the fight against COVID-19. Molecular docking was performed using two AutoDock 4 and Vina software on a set of 2,223 phytochemicals, and the spatial affinity for the active sites and interactions with active site residues were studied.

## 2 | MATERIALS AND METHODS

### 2.1 | 3D structures preparation

We used a dataset of compounds covering different groups of phytochemicals, obtained from KEGG “Phytochemical Compounds” (http://www.genome.jp/kegg-bin/get_htext?br08003.keg). The resulting
library was diverse, representing over 2,000 scaffolds. The library included alkaloids (714 compounds), amino acid-related compounds (104), fatty acids-related compounds (37), flavonoids (478), phenylpropanoids (189), polyketides (136), shikimate/acetate-malonate pathway derived compounds (45), terpenoids (101), and others (41), totally 2,845. 3D crystal structure of SARS-CoV-2 3clpro/Mpro (PDB ID: 6LU7) was retrieved from protein data bank (PDB: (https://www.rcsb.org/), in pdb format. The native inhibitor for 3CLpro n-[(5-methyl- isoxazol-3-yl)carbonyl]alanly-l-valyl-1-[(1r,2z)-4-(benzyloxy)-4-oxo- 1-(31) but-2-ethyl-l-leucinamide was docked as a potential noncovalent inhibitor of SARS-CoV 3CLpro (Khaerunnisa et al., 2020). Furthermore, specific inhibitors against 3CLpro, including Lopinavir and Nelfinavir were used for comparison (Lin et al., 2020).

2.2 | Molecular docking

The preparation of 3CLpro structure for the docking process was included water molecules and the co-crystal inhibitor exclusion. MGLTOOLS 1.5.6 (Morris, Huey, & Olson, 2008) was applied for converting the structure to PDBQT format and adding gasteiger partial charges. Self-docking was used for validation of docking protocol. This procedure was performed with modified parameters in such a way to result in RMSD values less than 2 Å. The carbon 19 atom of the ligand with the coordinates $x = -11.993, y = 15.425, and z = 65.951$ was considered as the center of the grid with dimensions of 40, 40, 40 in the active site of the enzyme. The exhaustiveness parameter was considered to be 8. The docking simulations were performed using AutoDock Vina (Trott & Olson, 2010) on a 40 cores system. To validate the result of molecular docking, AutoDock4 was used.

2.3 | Processing AutoDock4

AutoDock4 (AD4) (release 4.2.6) (Morris et al., 2009) was used for the validation molecular docking result of AutoDock Vina. AutoGrid (Morris et al., 2009) program supplied with AD4 was used for the preparation of grid maps. The grid box size was set at 40 Å for $x, y,$ and $z$ dimensions. The spacing between the grid points was 1.0 Å. The grid center was established at $x = -11.993, 15.425,$ and 65.951 Å for $x, y,$ and $z$, respectively. The Lamarckian Genetic Algorithm (LGA) was chosen for searching the best conformers. During the docking process, a maximum of 10 conformers was considered for each ligand. All the docking processes were performed with the default parameters of AD4. Population size was set to 150, the maximum number of evaluations 2,500,000, the maximum number of generations 27,000, maximum number of top individuals that automatically survived 1, gene mutation rate 0.02 and crossover rate 0.8.

2.4 | Swiss ADME

Drug-like properties of 2,845 phytochemicals were calculated using Lipinski’s rule of five, which proposes that molecules with poor permeation and oral absorption have molecular weights ≤500, consensus logP ≤5, equal to five or less than 5 hydrogen-bond donors, and equal to 10 or less than 10 acceptor groups (Gimenez, Santos, Ferrarini, & Fernandes, 2010; Lipinski, Lombardo, Dominy, & Feeney, 1997). Adherence with Lipinski’s rule of five was calculated using SWISSADME prediction (http://www.swissadme.ch/). 2,223 phytochemicals were approved for drug-like properties.

3 | RESULTS AND DISCUSSION

In this study, a comprehensive compounds library, including 2,845 phytochemicals, was obtained from KEGG “Phytochemical Compounds.” The library included a wide range of phytochemicals, such as alkaloids, flavonoids, and phenylpropanoids. A total of 2,845 phytochemical compounds were considered for docking analysis using AutoDock Vina, and validated using AutoDock4. Finally, the best 10 compounds were filtered using binding energy. These compounds included, based on lower binding energy (kcal/mol), Ginkgolide M ($−11.2$), mezerein ($−11$), tubocurarine ($−10.9$), Gnidicin ($−10.4$), Glycobismine A ($−10.2$), Sciadopitysin ($−10.2$), Gnididin ($−10.2$), Sciadopitysin ($−10.4$), Gnidicin ($−9.2$), Glycobismine A ($−10.4$), Sciadopitysin ($−10.2$), Gnididin ($−9.2$), EMETINE ($−8.7$), Vitexin ($−8.3$), Calophyllolide ($−8.3$) and 6-(3,3-Dimethylallyl) galangin ($−7.9$). The binding energy is related to the number of H-bonds and other interactions, such as $π−π$ interactions, formed with the active site pocket of the 3CLpro (Politi, Durdagi, Moutevelli-Minakakis, Kokotos, & Mavromoustakos, 2010). The 10 lead compounds and their binding energy, corresponding 2D structure, the interacting residues, and the hydrogen bonding are presented in Tables 1 and 2. This part of the study indicates remarkable binding affinities for the studied phytochemicals, compared to nelfinavir and lopinavir, −9.1 and −8.4, respectively.

Furthermore, the best 10 phytochemicals obtained by AutoDock Vina and AutoDock 4 with their corresponding docking scores and the interacting residues are presented in Table 3; catalytic dyad residues (His41 and Cys145) are shown in bold. Binding energy distribution for AD4 and Vina is shown in Figure 1. A comparison of the predicted binding energies from both programs is shown in Figure 2, demonstrating a marked correlation between the docking results. As evidenced by both Kendall rank correlation and traditional Pearson correlation, there was a clear association between the predictions from AD4 and Vina, (Pearson, correlation: $r = 0.83$, $p$ value = $< .0001$, and Kendall, correlation: $η = 0.2$, $p$ value = $< .0001$). Figure 3 shows the top 10 compounds obtained from AD4 and Vina in KEGG sub-libraries form.

The interacting of the binding residues of 3CLpro with the top the compounds, native inhibitor, and lopinavir and nelfinavir are depicted in Figure 1 and Table 3. As it is shown in Table 2, Ginkgolide M-3CLpro complex depicted significant interactions with Asn142, Cys145, Glu166, Gly143, His163, and Phe140 residues.

Docking analysis also revealed the possible H-bonds interactions between the active site amino acids and phytochemicals, Table 2. Accordingly, Ginkgolide M interacts with Gly143A, Cys145A,
TABLE 1  Molecular docking of the studied top 10 compounds according

| Target | Compound     | PubChem ID | Hydrophobic contacts (C–C) | Hydrogen bonds | Binding energy |
|--------|--------------|------------|-----------------------------|----------------|---------------|
| 3CLpro | Nelfinavir   | 64143      | 19                          | 2              | −9.1          |
|        | Lopinavir    | 92727      | 47                          | 2              | −8.4          |
|        | 3CLpro_inhibitor | 146025593 | 22                          | 3              | −6.6          |
|        | Ginkgolide M | 46173836   | 7                           | 7              | −11.2         |
|        | Mezerein     | 5281382    | 23                          | 4              | −11           |
|        | Tubocurarine | 6000       | 23                          | 3              | −10.9         |
|        | Gnidicin     | 5281366    | 15                          | 3              | −10.4         |
|        | Glycobismine A | 5462453 | 23                          | 3              | −10.4         |
|        | Emetine      | 10219      | 28                          | 3              | −8.7          |
|        | Vitexin      | 5280441    | 16                          | 4              | −8.3          |
|        | Calophyllolide | 5281392  | 14                          | 3              | −8.3          |
|        | Dauricine    | 73400      | 19                          | 3              | −8.2          |

Note: Binding energy against 3CLpro.

Ser144A, and His163A in the active site of the main protease. Mezerein forms H-bonds with the 3CLpro amino acids His163A, Ser144A, Cys145A, Leu141A, Asn142, and Met49A. Also, Glycobismine A forms H-bonds with the 3CLpro amino acids Ser144A, Cys145A, Gly143A, Leu141A, His163A, and Asn142A. The interactions of the other studied compounds are presented in Table 2.

Ginkgolides (Ginkgolic acid) are terpene lactones, and frequently found in the lipid fraction of the nutshell of the Ginkgo biloba tree (Schocka, Sobczyński, Ochnik, Zwolińska, & Leszek, 2019). Interestingly, it has shown that ginkgolic acid is effectively able to inhibit human immunodeficiency virus-1 (HIV-1) protease activity in a cell-free system and HIV infection in human peripheral blood mononuclear cells (PBMCs) without significant cytotoxicity (Lü et al., 2012). These compounds show negligible cytotoxicity (Liu & Zeng, 2009), and could possibly be used as potent antiviral therapies.

Mezerein, as the second top potent compound in our study, is a toxic diterpene ester found in the sap of Daphne mezereum and related plants. Mezerein is highly liposoluble and can cause vomiting, diarrhea, and burning of the mouth (Nelson, Shih, Balick, & Lampe, 2007). Tubocurarine is also known for its toxic effects. So, apparently these cannot be regarded as therapeutic due to cytotoxic effects. However, gnidicin as the fourth top compound is previously shown to be an antiviral compound. Vidal et al. showed that this compound, as one of the active ingredients of Daphne Gnidium, exhibits strong antiretroviral activity, CXCR4-tropic HIV-1 strain NL4-3 or the CCR5-tropic HIV-1 strain NLAD87, and absence of cytotoxicity (Vidal et al., 2012). Also, there are reports indicating the health benefits of other studied compounds, such as anti-Pneumocystis Carinii activity of Glycobismine A (Queener et al., 1991) and neuroprotection and antiinflammatory effects of Sciadopitysin (Choi, Suh, Rhee, & Kim, 2014). Furthermore, it has shown that emetine, a natural alkaloid, strongly reduced the production of RNA and DNA viruses without generating drug-resistant virus variants (Khandelwal et al., 2017). Similarly, emetine is able to inhibit Zika and Ebola virus infections through two molecular mechanisms; inhibiting viral replication and decreasing viral entry (Yang et al., 2018). Emetin also inhibits HIV-1 replication by interfering with reverse transcriptase activity (Valadão et al., 2015). Vitexin, isolated from Trollius chinensis Bunge, has shown to have antiviral activity against paraminfluenza type 3. The flower of this plant is used for treating upper respiratory infections, pharyngitis, tonsillitis, and bronchitis in Chinese folk medicine (Li, Ma, Yang, Ye, & But, 2002). Furthermore, vitexin exhibits significant antiviral activity (EC50 = 35 ± 2.7 and 18 ± 3.3 μg/ml against HAV-H10 and HSV-1 virus, respectively) (Fahmy et al., 2020). Calophyllolide, isolated from Calophyllum inophyllum, is shown to be a strong anti-HIV-1 compound (Laure, Raharivelomanana, Butaud, Bianchi, & Gaydou, 2008). Galangin is a strong antiviral compound against herpes simplex virus type 1 (HSV-1) and coxsackie B virus type 1 (Cox B1) (Meyer, Afolayan, Taylor, & Erasmus, 1997). Details of the selected compounds with previously proven antiviral activity are summarized in Table 4.

Ginkgo biloba and its derivatives, such as Ginkgolide M are widely used in traditional Chinese medicine. Nowadays, the leaf extracts are being sold as phytotherapy worldwide (Strømgaard & Nakanishi, 2004). Several health benefits have been reported for this phytomedicine, including improvement of memory, neuroprotection, increased blood circulation, as well as beneficial effects to sufferers of Alzheimer’s disease (Mohammad Nabavi et al., 2015; Polich & Gloria, 2001; Zimmermann, Colciaghi, Cattabeni, & Di, 2002).

The same inhibitory effect of the studied compound on the proteases from different sources may rise the question that what is the structural similarities among them. Generally, search of structures in the Protein Data Bank has identified remarkable structural similarity among SARS-CoV2 3CLpro protease and proteases of flaviviruses, noroviruses, and entroviruses; for example, coxsackievirus, hepatitis A, Norwalk, Zika, Dengue, and West Nile viruses (Bafna, Krug, & Montelione, 2020). The exact inhibitory mechanism of HIV protease inhibitors on the main proteases of SARS, MERS, and SARS-CoV-2 is not well understood. The protease of HIV is an aspartic protease, while the main protease in coronaviruses belongs to the
### TABLE 2
Details of the interaction of the compounds and SARS-Cov 2 M<sup>pro</sup> [Colour table can be viewed at wileyonlinelibrary.com]

| Compound (ID)       | Lipinski's rule of five | Interaction with 3CLpro active site |
|---------------------|-------------------------|-------------------------------------|
| 3CLpro_inhibitor    |                         | ![Image](image1.png)                |
|                     | H-bond acceptors 9      |                                     |
|                     | H-bond donor 6          |                                     |
| LogP                | 3.33                    |                                     |
| MW                  | 680.79                  |                                     |
| Violations          | 3                       |                                     |
| Nelfinavir (64143)  |                         | ![Image](image2.png)                |
|                     | H-bond acceptors 5      |                                     |
|                     | H-bond donor 4          |                                     |
| LogP                | 4.41                    |                                     |
| MW                  | 567.78                  |                                     |
| Violations          | 1                       |                                     |
| Lopinavir (92727)   |                         | ![Image](image3.png)                |
|                     | H-bond acceptors 5      |                                     |
|                     | H-bond donor 4          |                                     |
| LogP                | 4.53                    |                                     |
| MW                  | 628.8                   |                                     |
| Violations          | 1                       |                                     |
| Ginkgolide M (46173836) |                     | ![Image](image4.png)                |
|                     | H-bond acceptors 10     |                                     |
|                     | H-bond donor 3          |                                     |
| LogP                | −0.22                   |                                     |
| MW                  | 424.4                   |                                     |
| Violations          | 0                       |                                     |
| Compound (ID)              | Lipinski's rule of five | Interaction with 3CLpro active site |
|---------------------------|-------------------------|-------------------------------------|
| Mezereum (5281382)        | H-bond acceptors: 10    |                                     |
|                           | H-bond donor: 3         |                                     |
|                           | LogP: 3.52              |                                     |
|                           | MW: 654.7               |                                     |
|                           | Violations: 1           |                                     |
| Tubocurarine (6000)       | H-bond acceptors: 7     |                                     |
|                           | H-bond donor: 2         |                                     |
|                           | LogP: 3.55              |                                     |
|                           | MW: 609.73              |                                     |
|                           | Violations: 1           |                                     |
| Gnidicin (5281366)        | H-bond acceptors: 10    |                                     |
|                           | H-bond donor: 3         |                                     |
|                           | LogP: 3.06              |                                     |
|                           | MW: 628.67              |                                     |
|                           | Violations: 1           |                                     |
| Glycobismine A (5462453)  | H-bond acceptors: 6     |                                     |
|                           | H-bond donor: 4         |                                     |
|                           | LogP: 6.03              |                                     |
|                           | MW: 602.68              |                                     |
|                           | Violations: 1           |                                     |
| Compound (ID)            | Lipinski's rule of five | Interaction with 3CLpro active site |
|-------------------------|-------------------------|------------------------------------|
| Sciadopitysin (5281696) | H-bond acceptors 10     | ![Sciadopitysin (5281696) interaction with 3CLpro active site](image1) |
|                         | H-bond donor 3          |                                    |
|                         | LogP 4.76               |                                    |
|                         | MW 580.54               |                                    |
|                         | Violations 1            |                                    |
| Gnididin (5281367)      | H-bond acceptors 10     | ![Gnididin (5281367) interaction with 3CLpro active site](image2) |
|                         | H-bond donor 3          |                                    |
|                         | LogP 3.92               |                                    |
|                         | MW 648.74               |                                    |
|                         | Violations 1            |                                    |
| Emetine (10219)         | H-bond acceptors 6      | ![Emetine (10219) interaction with 3CLpro active site](image3) |
|                         | H-bond donor 1          |                                    |
|                         | LogP 4.24               |                                    |
|                         | MW 480.64               |                                    |
|                         | Violations 0            |                                    |
| Vitexin (5280441)       | H-bond acceptors 10     | ![Vitexin (5280441) interaction with 3CLpro active site](image4) |
|                         | H-bond donor 7          |                                    |
|                         | LogP −0.02              |                                    |
|                         | MW 432.38               |                                    |
|                         | Violations 1            |                                    |
TABLE 2  (Continued)

| Compound (ID) | Lipinski’s rule of five | Interaction with 3CLpro active site |
|---------------|-------------------------|------------------------------------|
| Calophyllolide (5281392) | H-bond acceptors 5 | |
|                | H-bond donor 0         | |
|                | LogP 4.76              | |
|                | MW 416.47              | |
|                | Violations 0           | |
|                | Violations 1           | |

6-(3,3-Dimethylallyl) Galangin (5281945) | H-bond acceptors 5 | |
|                                          | H-bond donor 3     | |
|                                          | LogP 3.59          | |
|                                          | MW 338.35          | |
|                                          | Violations 0       | |
|                                          | H-bond donor 0     | |
|                                          | LogP 3.11          | |
|                                          | MW 338.44          | |
|                                          | Violations 0       | |

Note: The properties of Lipinski's rule of five should be H-bond acceptor <10, H-Bond donor <5, LogP <4.15 and MW <500 Da. Violation from these criteria should be 0 or 1. The 3CLpro inhibitor has three violations and we report it for comparison.

TABLE 3  The best 10 phytochemicals obtained by AutoDock Vina and AutoDock 4 with their corresponding docking scores and the interacting binding residues are presented

| Compound      | ID          | Dock_score | Active site amino acid                               |
|---------------|-------------|------------|------------------------------------------------------|
| Ginkgolide M  | 46173836    | −11.2      | Asn142, Cys145, Glu166, Gly143, His163, Phe140       |
| Mezerein      | 5281382     | −11        | Asn142, Cys145, Glu166, Gly143, His163, His172, Phe140 |
| Tubocurarine  | 6000        | −10.9      | Asn142, Cys145, Glu166, Gly143                        |
| Glycobilamine A| 5462453     | −10.4      | Asn142, Glu166, Gly143, His163, His172, Phe140       |
| Gnidicin      | 5281366     | −10.4      | Asn142, Glu166, Gly143, His163, His172               |
| Sciadopitysin | 5281696     | −10.2      | Asn142, Cys145, Glu166, Gly143                        |
| Gnididin      | 5281367     | −9.2       | Asn142, Cys145, Glu166, Gly143, His163               |
| Emetine       | 10219       | −8.7       | Asn142, Glu166, Gly143                               |
| Vitexin       | 5280441     | −8.3       | Asn142, Cys145, Glu166, Gly143                       |
| Calophyllolide| 5281392     | −8.3       | Asn142, Cys145, Glu166, Gly143                       |

Note: Catalytic dyad residues (HIS41 and CYS145) are shown in bold.

cysteine proteases family. Furthermore, HIV protease inhibitors specifically fit to C2 symmetry in the catalytic site of the HIV protease dimer, but this C2-symmetric pocket is absent in coronavirus proteases (Pandey et al., 2020). As chymotrypsin and other proteins with similar 3D folding are serine proteases, it raised the hypothesis that 3CLpro was originally a chymotrypsin-like serine protease that...
later evolved into a cysteine protease (Shan, Li, & Xu, 2004). That a chymotrypsin-fold protease may be a target for HIV-protease inhibitors is supported by in vitro experiments of independent groups, showing that HIV-protease inhibitors also inhibit chymotrypsin-like components of the mammalian proteasome (Piccinini et al., 2002). Moreover, it has shown that SARS-CoV2 3CLpro protease has a striking three-dimensional with hepatitis C virus (HCV) NS3/4A protease, particularly in the arrangement of key active site residues (Bafna et al., 2020). This could be due to convergent evolution, in which different structural topologies create similar binding pockets. For example, Cus145 and His41 create catalytic dyad in 3Cl pro; however, Cys172 and His41 sidechains in the three-dimensional structure of hepatitis A virus 3Cpro protease.

The possible cytotoxic effects of the studied compounds may also be considered as a limitation in the clinical use of the. We have tried to do a literature review to find reports indicating possible toxic effects of the studied compounds. It has shown that Gnidicin and other active components of Daphne gnidium have high antiviral activity with absence of cytotoxicity at even high concentrations, 50 μM (Vidal et al., 2012). It has shown that glycobismine A is a potent anti-malarial agent that are comparable to or greater than that of chloroquine diphosphate, no cytotoxic effect was found for this compound (Fujioka, Nishiyama, Furukawa, & Kumada, 1989). Sciadopitysin is a strong anti-oxidant, neuroprotective and anti-inflammatory bioflavonoid, and no cytotoxic effect has been reported for it at concentrations up to 10 μM (Cao et al., 2017; Choi et al., 2014). Vitexin is a cytoprotective compound, acting through a wide range of protective mechanism. It has shown that cyto- and neuroprotective properties of vitexin is exerted through decrease of intracellular Ca²⁺, inhibition of inflammation, block the production of free radicals as well as effect on the gene expression level of the cellular antioxidant system (Chen, Zhang, Shan, & Zhao, 2016; Malar, Prasanth, Shafreen, Balamurugan, & Devi, 2018). Finally, it has shown that various concentrations of Calophyllolide (10–1,000 ng/ml) have no significant cytotoxicity on HaCaT and RAW264.7 cell viability after 24 hr treatment (Nguyen et al., 2017).

In a recently published study, Khaerunnisa et al. investigated the possible inhibitory effect of some derivatives of medicinal plants on the 3CLpro of SARS-CoV-2 virus using molecular docking (Khaerunnisa et al., 2020). Based on the docking result, they claimed that Luteolin-7-glucoside (−8.17), demetoxycurcumine (−7.99) were potent compounds interacting the active site of the enzyme. They also reported that two flavonol compounds, kaempferol (−8.58) and quercetin (−8.47) can significantly interact the enzyme’s active site. It has proposed that Hydroxy groups (−OH), ketone groups (==O) and ether groups (−O−) in these compounds play roles amino acid residue interactions at the active site of the 3CLpro. This study revealed a significant aspect of the inhibitory effect of these phytochemicals on the 3CLpro of the virus; however, it was carried out on a limited number of the compounds. In another similar study, released on March 26, 2020, Tahir ul Qamar et al. performed a study by analyzing the 3CLpro sequence, constructing its 3D homology model, and screened it.
against a medicinal plant library containing 32,297 potential antiviral phytochemicals/traditional Chinese medicinal compounds (ul Qamar, Alqahtani, Alamri, & Chen, 2020). They showed that some derivatives of isoflavone (PubChem ID:11610052 and NPACT00105), Myricitrin (5281673), methyl rosmarinate (6479915) and Amaranthin (6123095) might strongly inhibit SARS-CoV-2 3CLpro. The high number of studied phytochemicals, high binding energy (such as $-29.57$ kcal/mol) of the compounds, and effective interaction with catalytic dyad residues (Cys-145 and His-41) were the positive aspects of this study.

### Conclusion

Several phytochemicals have been recently considered as the potential inhibitors of the proteolytic activity of the main protease of SARS-CoV-2.
CoVs. The antiviral mechanism of the mentioned compounds is mainly mediated through inhibition of the main protease, as a proven drug discovery target, of the coronaviruses as well as inhibition of the ion channels (Schwarz et al., 2014). Our effort to target the main protease of the SARS-CoV2 yielded promising results. We performed molecular docking on more than 2000 phytochemicals, and we found 10 novel compounds to serve as inhibitors of 3CLpro protein. The selected natural products showed stronger binding energies than nelfinavir and lopinavir. Interestingly, these compounds were previously shown to possess antiviral properties. We, therefore, anticipate that the insights given in the current study could be regarded valuable towards the exploration and development of new therapies against COVID-19 by efficiently targeting and inhibiting the catalytic function of the main protease of the virus. This study warrants further experimental work for finding useful candidates for COVID-19 drug therapy.

DATA AVAILABILITY STATEMENT
Data available on request from the authors.

ORCID
Amir Mahmoodzadeh https://orcid.org/0000-0002-0523-8152

REFERENCES
Adhikari, S. P., Meng, S., Wu, Y.-J., Mao, Y.-P., Ye, R.-X., Wang, Q.-Z., … Zhou, H. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: A scoping review. Infectious Diseases of Poverty, 9(1), 29.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., & Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. Science, 300(5626), 1763–1767.

Arabi, Y. M., Asiri, A. Y., Assiri, A. M., Jokhdar, H. A. A., Alothman, A., Balkhy, H. H., … Mohammah A. H. (2020). Treatment of Middle East respiratory syndrome with a combination of lopinavir/ritonavir and interferon-β1b (MIRACLE trial): Statistical analysis plan for a recursive two-stage group sequential randomized controlled trial. Trials, 21(1), 1–8.

Baez-Santos, Y. M., St John, S. E., & Mesecar, A. D. (2015). The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds. Antiviral Research, 115, 21–38.

Bafna, K., Krug, R. M., & Montellione, G. T. (2020). Structural similarity of SARS-CoV2 M(pro) and HCV NS3/4A proteases suggests new approaches for identifying existing drugs useful as COVID-19 therapeutics. ChemRxiv [Preprint]. 1, https://doi.org/10.26434/chemrxiv.12153615v1.

Barbieri, R., Coppo, E., Marchese, A., Daglia, M., Sobarzo-Sánchez, E., Nabavi, S. F., & Nabavi, S. M. (2017). Phytochemicals for human disease: An update on plant-derived compounds antibacterial activity. Microbiological Research, 196, 44–68.

Ben-Shabat, S., Yarmolinsky, L., Porat, D., & Dahan, A. (2020). Antiviral effect of phytochemicals from medicinal plants: Applications and drug delivery strategies. Drug Delivery and Translational Research, 10, 1–14.

Cao, J., Lu, Q., Liu, N., Zhang, Y.-x., Wang, J., Zhang, M., … Wan-chun, S. (2020). Sciadopitysin suppresses RANKL-mediated osteoclastogenesis and prevents bone loss in LPS-treated mice. International Immunopharmacology, 2017(49), 109–17.

Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., … Kwok-Yung, Y. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. The Lancet, 395(10223), 514–523.

Chen, L., Zhang, B., Shan, S., & Zhao, X. (2016). Neuroprotective effects of vitexin against isoflavone-induced neurotoxicity by targeting the TRPV1 and NR2B signaling pathways. Molecular Medicine Reports, 14(6), 5607–5613.

Chikhale, R. V., Sinha, S. K., Patil, R. B., Prasad, S. K., Shakyah, A., Gurav, N., … et al. (2020). In-silico investigation of phytochemicals from Asparagus racemosus as plausible antiviral agent in COVID-19. Journal of Biomo-lecular Structure and Dynamics, 1, 1–15.

Choi, E. M., Suh, K. S., Rhee, S. Y., & Kim, Y. S. (2014). Sciadopitysin alleviates methylglyoxal-mediated glycation in osteoblastic MC3T3-E1 cells by enhancing glyoxalase system and mitochondrial biogenesis. Free Radical Research, 48(7), 729–739.

Fahmy, N. M., Al-Sayed, E., Moghannem, S., Azam, F., El-Shazly, M., & Singab, A. N. (2020). Breaking down the barriers to a natural antiviral agent: Antiviral activity and molecular docking of Erythrina speciosa extract, fractions, and the major compound. Chemistry & Biodiversity, 17(2), e1900511.

Fujikoa, H., Nishiyama, Y., Furukawa, H., & Kumada, N. (1989). In vitro and in vivo activities of atalaphilline and related acidine alkaloids against rodent malaria. Antimicrobial Agents and Chemotherapy, 33(1), 6–9.

Gimenez, B., Santos, M., Ferrarini, M., & Fernandez, J. (2010). Evaluation of blockbuster drugs under the rule-of-five. Die Pharmazie-an Interna-tional Journal of Pharmaceutical Sciences, 65(2), 148–152.

Han, D. P., Penn-Nicholson, A., & Cho, M. W. (2006). Identification of poten-tial determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. Virology, 350(1), 15–23.

Hasan, A., Paray, B. A., Hussain, A., Qadir, F. A., Attar, F., Aziz, F. M., … Falahati, M. (2020). A review on the cleavage priming of the spike pro-tein on coronavirus by angiotensin-converting enzyme-2 and furin. Journal of Biomolecular Structure and Dynamics, 22, 1–9.

Hsieh, L. E., Lin, C. N., Su, B. L., Jan, T. R., Chen, C. M., Wang, C. H., … Chueh, L. L. (2010). Synergistic antiviral effect of Galanthus nivalis agglutinin and nelfinavir against feline coronavirus. Antiviral Research, 88(1), 25–30.

Hui, D. S., Azhar, E. I., Madani, T. A., Ntoumi, F., Kock, R., Dar, O., … Eskild, P. (2020). The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-the latest 2019 novel coronavirus outbreak in Wuhan, China. International Journal of Infectious Diseases, 91, 264–266.

Hun, I. F.-N., Lung, K.-C., Tso, E. Y.-K., Liu, R., Chung, T. W.-H., Chu, M.-Y., … Yuen, K. Y. (2020). Triple combination of interferon beta-1b, lopinavir-ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: An open-label, randomised, phase 2 trial. The Lancet, 395(10238), 1695–1704.

Islam, R., Parves, M. R., Paul, A. S., Uddin, N., Rahman, M. S., Mamun, A. A., … Halim, M. A. (2020). A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2. Journal of Biomolecular Structure and Dynamics, 12, 1–12.

Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S., & Soetjipto, S. (2020). Potential inhibitor of COVID-19 Main protease (Mpro) from several medicinal plant compounds by molecular docking study. Preprints, 1, 1–14. https://doi.org/10.20944/preprints202003.0226.v1.

Khandelwal, N., Chander, Y., Rawat, K. D., Riyesh, T., Nishanth, C., Sharma, S., … Kumar, N. (2017). Emetine inhibits replication of RNA and DNA viruses without generating drug-resistant virus variants. Antiviral Research, 144, 196–204.

Laure, F., Raharivelomanana, P., Butaud, J.-F., Bianchini, J.-P., & Gaydou, E. M. (2008). Screening of anti-HIV-1 inophyllums by HPLC agglutinin and nelfinavir against feline coronavirus. Antiviral Research, 78(3), 256–262.

Mamun, A. A., Halim, M. A. (2020). A molecular modeling approach for identifying existing drugs useful as COVID-19 therapeutics. ChemRxiv [Preprint]. 1, https://doi.org/10.26434/chemrxiv.48(7), 729–739.
potential drugs by computational methods. Acta Pharmaceutica Sinica B, 10, 766–788.

Yang, S., Xu, M., Lee, E. M., Gorshkov, K., Shiryaev, S. A., He, S., ... Zheng, W. (2018). Emetine inhibits Zika and Ebola virus infections through two molecular mechanisms: Inhibiting viral replication and decreasing viral entry. Cell Discovery, 4(1), 31.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., ... Rolf, H. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Science, 368(6489), 409–412. https://doi.org/10.1126/science.abb3405

Zimmermann, M., Colciaghi, F., Cattabeni, F., & Di, M. L. (2002). Ginkgo biloba extract: From molecular mechanisms to the treatment of Alzheimer's disease. Cellular and Molecular Biology (Noisy-le-Grand, France), 48(6), 613–623.

Zumla, A., Chan, J. F., Azhar, E. I., Hui, D. S., & Yuen, K.-Y. (2016). Coronavirus—Drug discovery and therapeutic options. Nature Reviews Drug Discovery, 15(5), 327–347.

How to cite this article: Sisakht M, Mahmoodzadeh A, Darabian M. Plant-derived chemicals as potential inhibitors of SARS-CoV-2 main protease (6LU7), a virtual screening study. Phytotherapy Research. 2021;35:3262–3274. https://doi.org/10.1002/ptr.7041