In silico screening of natural antivirals as potential inhibitors of SARS-CoV-2 virus

Ta Thi Thu Hang, Do Thi Hong Khanh, Bui Thanh Tung*
Department of Pharmacology and Clinical Pharmacy, VNU University of Medicine and Pharmacy, Vietnam National University Hanoi, Office 506, Y1 Building, 144 Xuan Thuy, Cau Giay, Hanoi 10000, Viet Nam

Submitted October 24, 2021; Revised January 4, 2022; Accepted January 17, 2022

Abstract

Coronavirus infectious disease 2019 (COVID-19) is an infectious disease of the human respiratory tract caused by the SARS-CoV-2 virus. Spike protein is a class I glycoprotein trimeric TM involved in viral entry and infection. Four major targets to inhibit the SARS-CoV-2 virus are spike protein, angiotensin-converting enzyme 2 (ACE2), main protease and the enzyme RNA-dependent RNA polymerase (RdRp). In this study, we evaluated the inhibitory potential of natural antiviral compounds against spike protein, ACE2, main protease, RdRp targets by molecular docking and molecular dynamics simulations. Lipinski Rule of Five was used to evaluate the drug-like properties of these compounds. The pkCSM tool was used to assess the pharmacokinetic parameters of prospective substances. Based on the ChemFaces database, we have collected 273 natural antiviral compounds. The results showed that the 7/273 compounds with the most potential to inhibit SARS-CoV-2 were: hinokiflavone, sotetsuflavone, mulberroside C, daphnoretin, morellic acid, digitoxin, and hypericin. Among them, sotetsuflavone is the most potent compound that inhibits four targets, with drug-like properties, good intestinal absorption, and low toxicity. The molecular dynamics simulation results of the complexes are also relatively stable. As a result, in vitro and in vivo test should be carried out to verify the potential for COVID-19 treatment of this compound.

Keywords. Antivirals, phytochemicals, SARS-CoV-2, spike protein, ACE2, RdRp, main protease, molecular docking, molecular dynamics, in silico.

1. INTRODUCTION

Coronavirus infectious disease 2019 (COVID-19) is an infectious disease of the human respiratory tract caused by the SARS-CoV-2 virus. This disease first appeared in Wuhan city, Hubei province, China, in December 2019.[1] As of November 19th, 2021, 256072650 cases of COVID-19 have been recorded and territory to 5132202 deaths, accounting for 2 % (WHO). Currently, these numbers are still increasing. COVID-19 can be transmitted from person to person or from animal to humans through direct contact through droplets or air. The clinical presentation of the cases included fever, cough, dyspnea, chest pain, and pulmonary infiltrates, similar to those of SARS and MERS patients. One of the concerns is that the symptoms of the disease are often very diverse, even can manifest differently in each patient. People with poor resistance or some reduced physiological functions such as the elderly and people with chronic diseases, easily develop the severe disease if infected. In addition, pregnant women and infants are also susceptible to the disease, mainly due to poor immunity.[2]

SARS-CoV-2 has a 29.9 kb positive-sense RNA genome, consisting of 14 open reading frames (ORFs) and coding for a total of 27 proteins.[3] From the two open reading frames ORF1a and ORF1b, the first two-thirds of the genome encode nonstructural proteins. Structural proteins are encoded in the genome’s final portion. Spike protein (S), membrane protein (M), envelope protein (E), and nucleocapsid protein (N) are the four structural proteins in CoVs that are conserved.[4] Covering the surface of the virus is a large amount of glycosylated protein S that is glycosylated and bound to the host cell's angiotensin II-converting enzyme (ACE II).[5] Spike protein is a class I glycoprotein trimeric TM involved in viral entry and infection. Spike protein has two main functions: receptor binding and virus fusion.[6] ACE2 is a transmembrane protein belonging to a class of functional type 2 receptors expressed in alveolar epithelial cells. The ACE2 receptor is the primary target for viral access and adhesion to host cells.[7]

Main protease (Mpro) is an important enzyme for viral replication, specifically the cleavage of translated polyproteins from viral RNA to produce
functional proteins.\textsuperscript{[8]} RNA-dependent RNA polymerase (RdRp), an important enzyme, plays a role in RNA replication and transcription in host cells.\textsuperscript{[9]} RdRp is a viral enzyme with no host cell homolog, and selective inhibitors of RdRp have little effect on cells other than the target cell. Therefore, it can be developed to improve efficiency and cause fewer unwanted effects. This is a safer and more effective treatment for COVID-19 than inhibitors of other molecular targets.\textsuperscript{[10]}

Remdesivir is the first antiviral drug approved by FDA to treat COVID-19.\textsuperscript{[11]} The mechanism of remdesivir is to inhibit the RNA-dependent RNA polymerase of the SARS-CoV-2 virus.\textsuperscript{[12]} In addition to remdesivir, other pharmacological antiviral drugs have also been extensively studied for their potential to treat COVID-19. Besides medicinal chemical compounds, natural compounds are of increasing interest to scientists due to their few side effects and great health benefits. Therefore, our research goal is to discover potential natural antiviral compounds that inhibit four important targets: spike protein, ACE2, main protease, and RdRp of SARS-CoV-2 virus by in silico approach.

2. MATERIALS AND METHODS

2.1. Preparation of ligands

We collected 273 natural antiviral compounds from ChemFaces with the target of inhibiting the SARS-CoV-2 virus. The ligand structures of these 273 compounds and the positive controls (remdesivir, artemisinin) were downloaded from the PubChem database in SDF format. Then, we were optimized the structures of these compounds using the MMFF94 force field, Conjugate Gradients method, and saved them in PDB format. For this process, we were used Avogadro software. Finally, we were converted the PDB format to PDBQT format using Autodock tools.

2.2. Preparation of protein structures

The 3D structures of the angiotensin-converting enzyme 2 (PDB ID: 1R4L), spike protein (PDB ID: 6VSB), RNA dependent RNA polymerase (PDB ID: 6M71), and main protease (PDB ID: 6W63) of the SARS-CoV-2 virus were retrieved from the Protein Data Bank RCSB. The structure of angiotensin-converting enzyme 2 (ACE2) (ID: 1R4L) contains the co-crystallized ligand XXX. The crystal structure of the main protease (ID: 6W63) is a complex with the non-covalent inhibitor X77. It is important to note that the majority of Mpro inhibitors reported are covalent inhibitors.\textsuperscript{[13-15]} These co-crystallized ligands were used to evaluate and optimize the docking model. In preparation for docking, all water and co-crystals were removed from the protein molecule using Discovery Studio Visualizer 4.0 software. Hydrogen molecules will be added using Autodock Vina software before regenerating the active site using MGL Autodock tools 1.5.6 software. The active sites of spike protein, main protease, RdRp, and ACE2 were located in grid boxes of corresponding sizes (50 Å × 50 Å × 50 Å); (20 Å × 20 Å × 20 Å); (30 × 30 Å × 30 Å); (20 × 20 Å × 20 Å). These proteins were saved in pdbqt format to prepare for the docking program.

2.3. Molecular docking study

The ligands were docked to the active site of the four targets using Autodock vina software.\textsuperscript{[16]} The ligand-protein interaction energy is calculated by the scoring function of Autodock vina. The molecular interactions between compounds which have good free binding energies and molecular targets were viewed using Discovery Studio Visualizer 2020.

2.4. Validation of docking protocol

To evaluate the docking results, the co-crystal ligand was redocked to the active site of the target. The process was performed successfully if the root-mean-square deviation (RMSD) value was less than or equal to 1.5 Å.

2.5. Lipinski’s rule of five

Lipinski’s rule of five provides a method to compare drug-like and non-drug-like molecules.\textsuperscript{[17]} Lipinski’s rule of five is popularly used to evaluate the potential molecular to become a therapeutic drug. This rule acts as a filter to screen promising compounds with a particular pharmacological. We used the online tool to evaluate Lipinski’s rule of five.\textsuperscript{[18]} The chemical structures were downloaded from the PubChem database and set at pH 7.0.

2.6. Prediction of ADMET by computational analysis

The physicochemical efficiency of the compounds was analyzed by in silico ADMET profiling. ADMET profile involves five parameters: absorption, distribution, metabolism, excretion, and toxicity that play a significant role to demonstrate the likelihood for the success of a compound to be a drug. ADMET profiling was predicted using the
pkCSM tool.\cite{19} The canonical SMILES molecular structures of the collected compounds were retrieved from PubChem.

2.7. Molecular dynamics simulation of the most potent compound

Molecular dynamics helps to evaluate the physical movement and interactions of all atoms in the protein-ligand complex with themselves and their surroundings (e.g. water). We used MOE version 2015.10 to simulate the ligand-protein complex with the least binding energy pose. The Nosé-Poincaré-Andersen (NPA) equations of motion were employed in the MOE dynamics simulation. The MD default steps and protocols were chosen to optimize the system equilibrium 600 ps and a 500 ps production run was carried out at 310 K. RMSD was calculated as below:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} d_i^2}$$

where $d_i$ is the distance between atom $i$ at two different times and $N$ is the total number of atoms of the complex.

3. RESULTS AND DISCUSSION

3.1. Evaluation of the docking model

Before screening compounds, the docking model's accuracy needs to be evaluated. Co-crystallized ligands X77 and XX5 were separated from the 6W63 (Mpro) and 1R4L (ACE2) complexes, respectively. Then, they were redocked to the active site of the target to determine the root mean square deviation (RMSD) from which to evaluate the suitability of the docking parameters. After redocking the co-crystal ligand, we obtained RMSD values of 0.986 and 0.777 Å, respectively (figure 1). These two values satisfy the condition that RMSD is less than 1.5 Å, proving that the results of molecular docking to the target are reliable.

![Figure 1: Co-crystallized ligand redock results of 6W63 (left) and 1R4L (right)](image)

3.2. Molecular docking of compounds to the target protein

We docked 273 antiviral compounds against four targets: spike protein, ACE2, main protease, and RdRp to screen for their inhibitory potential. Remdesivir is a RNA dependent RNA polymerase (RdRp) inhibitor and the first COVID-19 drug approved by the FDA.\cite{11,12} Artemisinin has also been demonstrated by Moussa Sehailia et al. to act as a spike protein inhibitor.\cite{20} Therefore, in this study, we used remdesivir and artemisinin as positive controls for the RdRp and spike protein, respectively. Regarding the inhibition of main protease and ACE2 receptors by antiviral compounds, the binding energies for two co-crystallization ligands (XX5, X77) after redock will be used for comparison. The interactions between the reference inhibitor and the targets are shown in figure 2.

The docking results obtained 150/273 compounds capable of inhibiting at least two targets compared to positive controls. Among them, there are 82 compounds with negative binding energy than the reference substance on at least three targets. In particular, half of these compounds were antiviral terpenoids with selective inhibition on three targets: protein S, ACE2, RdRp. The results of the 13 most potent natural antiviral compounds are shown in table 1. These compounds inhibit all four targets.
with more negative binding energy than positive controls.

The results of table 1 show that 13 compounds all have good ΔG free energy compared with positive controls: remdesivir (-7.3 kcal/mol), artemisinin (-7.3 kcal/mol), XX5 (-8.7 kcal/mol), artemisinin (-7.3 kcal/mol), XX5 (-8.7 kcal/mol), and X77 (-8.6 kcal/mol). Among them, morellic acid and digitoxin inhibited the spike protein target with the most negative binding energy of -10.3 kcal/mol. Morellic acid also gives the most negative binding energy to the ACE2 target with -12.1 kcal/mol. Of the remaining two targets, the main protease and RdRp, sotetsuflavone gave the most negative binding energy, -9.6 and -9.5 kcal/mol, respectively. The ligand-amino acid interactions between them and the four targets are shown in figure 3.

The results showing the ligand-amino acid interaction of sotetsuflavone and morellic acid with the respective targets showed that these two substances interact with many amino acids in the active regions of the targets, similar to the positive controls. Specifically, morellic acid not only has an alkyl bond with LEU966 similar to artemisinin but also interacts with many other amino acids in the active region such as alkyl bond with VAL976,
hydrogen bonds with THR549, THR547. Similarly, sotetsuflavone also interacts with many amino acids of RdRp targets: PHE793, LYS621, VAL166, PRO620... through π-alkyl bonds, hydrogen bonds. The ligand-amino acid interaction of sotetsuflavone with the main protease is similar to X77, which is π-π, π-alkyl, π-sulfur, hydrogen bonds with important amino acids: HIS41, MET49, CYS145, GLU166, GLY143 etc.

Table 1: The docking results of the 13 most potent natural antivirals and reference compounds

| No. | Name                                         | Pubchem ID | Binding energy with spike protein (kcal/mol) | Binding energy with human ACE2 (kcal/mol) | Binding energy with main protease (kcal/mol) | Binding energy with RdRp (kcal/mol) |
|-----|----------------------------------------------|------------|---------------------------------------------|----------------------------------------|---------------------------------------------|----------------------------------|
| 1   | 3,4-Di-O-cafeoylquinic acid methyl ester     | 10392218   | -8.9                                        | -10.0                                  | -8.9                                        | -8.0                             |
| 2   | (-)-Epigallocatechin gallate                 | 65064      | -8.6                                        | -9.7                                   | -8.6                                        | -8.0                             |
| 3   | Hinokiflavone                                | 5281627    | -9.5                                        | -11.6                                  | -8.9                                        | -9.1                             |
| 4   | Sotetsuflavone                               | 5494868    | -9.3                                        | -11.1                                  | -9.6                                        | -9.5                             |
| 5   | Mulberroside C                               | 190453     | -9.2                                        | -10.5                                  | -9.0                                        | -8.9                             |
| 6   | Daphnoretin                                  | 5281406    | -8.5                                        | -9.7                                   | -8.6                                        | -7.4                             |
| 7   | Isochlorogenic acid A                        | 6474310    | -9.3                                        | -10.5                                  | -9.2                                        | -8.3                             |
| 8   | Isochlorogenic acid C                        | 274951298  | -8.4                                        | -9.9                                   | -8.9                                        | -7.9                             |
| 9   | 1,4-Dicaffeoylquinic acid                    | 12358846   | -9.2                                        | -10.6                                  | -8.7                                        | -7.9                             |
| 10  | Morellic acid                                | 54580250   | -10.3                                       | -12.1                                  | -9.5                                        | -7.8                             |
| 11  | Phoyunnanin C                                | 16220719   | -9.6                                        | -11.3                                  | -8.9                                        | -8.2                             |
| 12  | Digitoxin                                    | 441207     | -10.3                                       | -9.3                                   | -9.5                                        | -9.0                             |
| 13  | Hypericin                                    | 3663       | -9.3                                        | -10.2                                  | -9.2                                        | -8.7                             |
|     | Remdesivir                                   |            |                                              |                                        |                                             | -7.3                             |
|     | Artemisinin                                  |            |                                              |                                        |                                             | -7.3                             |
|     | XX5                                          |            |                                              |                                        |                                             | -8.7                             |
|     | X77                                          |            |                                              |                                        |                                             | -8.6                             |

a. Interaction between morellic acid and spike protein representing 3D and 2D
b. Interaction between morellic acid and ACE2 representing 3D and 2D

c. Interaction between sotetsuflavone and main protease representing 3D and 2D

d. Interaction between sotetsuflavone and RdRp representing 3D and 2D

*Figure 3:* The interaction between the ligands gives the most negative free binding energy and the respective targets
3.3. Lipinski’s rule of five

Lipinski’s rule of five helps identify drug-like and non-drug-like molecules. It predicts with high probability the drug-like effectiveness or failure of molecules complying with 2 or more of the following rules: molecular mass (MW) below 500 Daltons; high lipophilicity (LogP does not exceed 5); no more than 5 donors of hydrogen bonds (HBD); no more than 10 acceptors of hydrogen bonds (HBA1); and molar refractivity (MR) should be between 40-130.

Table 2: The result of Lipinski’s rule of five

| No. | Name                                         | Molecular weight | HBD | HBA1 | logP  | MR        | Drug- likeness |
|-----|----------------------------------------------|------------------|-----|------|-------|-----------|---------------|
| 1   | 3,4-Di-O-caffeoylquinic acid methyl ester    | 530              | 6   | 12   | 1.118 | 129.578   | Yes           |
| 2   | (-)-Epigallocatechin gallate                 | 458              | 8   | 11   | 2.233 | 108.921   | Yes           |
| 3   | Hinokiflavone                                | 538              | 5   | 10   | 5.239 | 140.037   | Yes           |
| 4   | Sotetsuflavone                               | 552              | 5   | 10   | 5.122 | 145.509   | Yes           |
| 5   | Mulberroside C                               | 458              | 5   | 9    | 1.507 | 115.460   | Yes           |
| 6   | Daphnoretin                                  | 352              | 1   | 7    | 2.672 | 90.009    | Yes           |
| 7   | Isochlorogenic acid A                        | 516              | 7   | 12   | 1.030 | 125.200   | Yes           |
| 8   | Isochlorogenic acid C                        | 516              | 7   | 12   | 1.029 | 125.197   | Yes           |
| 9   | 1,4-Dicaffeoylquinic acid                    | 516              | 7   | 12   | 1.029 | 125.198   | Yes           |
| 10  | Morellic acid                                | 560              | 2   | 8    | 5.513 | 151.264   | Yes           |
| 11  | Phoyunnanin C                                | 482              | 4   | 6    | 5.259 | 135.463   | Yes           |
| 12  | Digitoxin                                    | 764              | 5   | 12   | 4.669529 | 200.332   | Yes           |
| 13  | Hypericin                                    | 504              | 6   | 8    | 4.266871 | 134.460   | Yes           |

All these top compounds are satisfied at least two criteria. Then, we focus on analyzing the pharmacokinetic properties including absorption, distribution, metabolism, excretion, and toxicity of these antiviral compounds.

3.4. Prediction of absorption, distribution, metabolism, excretion, and toxicity (ADMET) profile

The seven natural compounds with the best prediction of absorption, distribution, metabolism, excretion, and toxicity are presented in table 3.

The first property is the absorption process, in which human intestinal absorption (HIA) and human colon adenocarcinoma-2 cell line (Caco2) are two crucial parameters that determine the absorption of the drug. A compound has high Caco-2 permeability if it has logPapp > 0.9. The results show that all seven compounds have Caco2 permeability less than 0.9. In which, morellic acid and digitoxin gave the highest Caco2 permeability with values of 0.705 and 0.601, respectively. A substance is considered poorly absorbed if the percentage absorbed in the human intestine is less than 30 %.[19] All compounds in table 3 are very well absorbed from the intestine, especially daphnoretin and hypericin with a maximum absorption percentage of 100 %.

Second, the distribution of substances was evaluated by three parameters: the steady-state volume of distribution (VDss), blood-brain barrier permeability (logBB), and Central Nervous System permeability (logPS). Compounds were said to be well distributed to tissues if logVDss > 0.45 and poorly distributed if logVDss < -0.15. The higher the VDss, the more of a drug is distributed in tissue rather than plasma.[19] Mulberroside C was well distributed in tissues with a logVDss value of 0.55, while hinokiflavone, sotetsuflavone, morellic acid, and hypericin were poorly distributed with values < -0.15. The capability of a drug to cross into the brain is a factor to consider to help reduce toxicity.
and side effects or to improve the effectiveness of drugs whose pharmacological activity is within the brain.\cite{19} The ability to cross the blood-brain barrier of most compounds was poor because their logBB values were less than -1, except daphnoretin and morellic acid. Furthermore, the results showed that all seven compounds failed to enter the CNS because their logPS values were less than -2.

The cytochrome P450 enzymes play a role in the metabolism of a variety of medicines. P450 inhibitors can significantly alter the pharmacokinetics of these drugs. The results showed that six compounds are not inhibitors of CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4. Daphnoretin is an inhibitor of CYP1A2 and CYP3A4.

Regarding elimination, we predicted total clearance and likelihood as a renal OCT2 substrate.

Table 3: Pharmacokinetic and toxicological prediction results

| Properties                | Hinoniflavone | Sotetsuflavone | Mulberroside | Daphnoretin | Morellic acid | Digitoxin | Hypericin |
|---------------------------|---------------|----------------|--------------|-------------|--------------|-----------|-----------|
| Absorption               |               |                |              |             |              |           |           |
| Water solubility (log mol/l) | -2.894       | -2.893         | -2.918       | -3.357      | -4.393       | -4.41     | -2.892    |
| CaCO₂ permeability (log Pa in 10⁻⁶ cm/s) | 0.242         | 0.085          | 0.279        | 0.369       | 0.705        | 0.601     | -0.594    |
| Intestinal absorption (human) (%) | 79.914       | 82.325         | 64.806       | 100         | 82.452       | 74.287    | 100       |
| Skin Permeability         | -2.735        | -2.735         | -2.735       | -2.746      | -2.735       | -2.735    | -2.735    |
| Distribution              |               |                |              |             |              |           |           |
| VDss (human) (log L/kg)   | -1.231        | -1.19          | 0.55         | -0.142      | -0.438       | 0.259     | -0.734    |
| BBB permeability (log BB) | -1.784        | -1.629         | -1.417       | -0.747      | -0.418       | -1.364    | -1.561    |
| CNS permeability (log PS) | -3.031        | -3.531         | -3.444       | -2.265      | -2.717       | -3.553    | -3.443    |
| Metabolism                |               |                |              |             |              |           |           |
| CYP1A2 inhibitor          | No            | No             | No           | Yes         | No           | No        | No        |
| CYP2C19 inhibitor         | No            | No             | No           | No          | No           | No        | No        |
| CYP2C9 inhibitor          | No            | No             | No           | No          | No           | No        | No        |
| CYP2D6 inhibitor          | No            | No             | No           | No          | No           | No        | No        |
| CYP3A4 inhibitor          | No            | No             | Yes          | No          | No           | No        | No        |
| Excretion                 |               |                |              |             |              |           |           |
| Total clearance (log ml/min/kg) | 0.513         | 0.617          | -0.042       | 0.837       | -0.405       | 0.445     | 0.004     |
| Renal OCT2 substrate      | No            | No             | No           | No          | No           | No        | No        |
| Toxicity                  |               |                |              |             |              |           |           |
| AMES toxicity             | No            | No             | No           | No          | No           | No        | No        |
| Oral Rat Acute Toxicity (LD50) | 2.648         | 2.548          | 2.686        | 2.453       | 2.946        | 3.706     | 2.482     |
| Hepatotoxicity            | No            | No             | No           | No          | No           | No        | No        |
| Skin sensitisation        | No            | No             | No           | No          | No           | No        | No        |

© 2022 Vietnam Academy of Science and Technology, Hanoi & Wiley-VCH GmbH www.vjc.wiley-vch.de 218
Organic Cation Transporter 2 (OCT2) is a renal absorption transporter that plays an important role in renal processing and clearance of drugs and endogenous compounds. All substances are not OCT2 substrates. The resulting total clearance is shown in Table 3.

Finally, all seven selected natural antiviral compounds were non-hepatotoxic, non-mutagenic AMES, and non-dermatologically toxic. Currently, compounds sotetsuflavone, mulberroside C, daphnoretin, morellic acid, hypericin have not had many studies on toxicity in humans. Digitoxin is a much-studied cardiac glycoside that is cardiotoxic, similar to digoxin. Hinokiflavone is a mildly cytotoxic compound that inhibits cancer cell growth with IC50 in the range of 15-40 μM.

3.5. Molecular dynamics results

Based on the docking scores and ADMET prediction results, we choose the most potential sotetsuflavone to simulate molecular dynamics. ACE2 and main protease play an important role in the entry and replication of the SARS-CoV-2 virus. In addition, the protein structure of these two targets is also relatively simpler than that of the RdRp and spike protein targets. Therefore, we selected two targets, ACE2 and main protease, to investigate the stability of the binding site of the most potent compound sotetsuflavone on these two targets. We performed molecular dynamics simulations of the enzyme and ligand complexes, using the docking results as the initial configuration.

![Figure 4: RMSD of ACE2-Sotetsuflavone complex (A) and Main protease-Sotetsuflavone complex (B) during 600 ps of molecular dynamics simulation](image)

![Figure 5: Free energy ACE2-Sotetsuflavone complex (A) and Main protease-Sotetsuflavone complex (B) during 600 ps of molecular dynamics simulation](image)

The molecular dynamics simulation results of both complexes show that the stable model with free energy reaches equilibrium. The RMSD values of both complexes are small and stable, showing that there is almost no significant difference in the atomic positions of the complex after 600ps of molecular dynamics simulation. However, the RMSD of the ACE2-sotetsuflavone complex tended to decrease over time. Figure 5 shows that the free energy of ACE2-sotetsuflavone is more stable than...
that of the main protease-sotetsuflavone complex. After 300ps simulation, the free energy of the main protease-sotetsuflavone complex increased and remained stable for the rest of the time. Therefore, the stability of the enzyme-ligand complex should be investigated for a longer time.

All targets including spike protein, ACE2, main protease, RdRp are potential targets in discovering drugs to treat SARS-CoV-2 virus. In this study, we screened for the ability of 175 natural antivirals to inhibit these four targets. The results obtained 13 compounds capable of inhibiting all targets with a much more negative free binding energy than the positive controls. Among them are compounds with good predictive pharmacokinetic results: hinokiflavone, sotetsuflavone, mulberroside C, daphnoretin, morellic acid, digitoxin, and hypericin.

Sotetsuflavone is a biflavonoid with anti-cancer properties.[21] Besides, Paul Coulerie et al. found that the Dengue virus NS5 RNA-dependent RNA polymerase was strongly inhibited, with an IC_{50} values of 0.16µM.[24] Our results also show that sotetsuflavone inhibits the RdRp of SARS-CoV-2 virus with the most negative energy compared to all compounds (-9.5 kcal/mol). We suggest that C3', C6' linkages are important for the ability of SARS-CoV-2 virus to inhibit RdRp, similar to Dengue virus. In addition, this compound also strongly inhibits Spike protein targets (-9.3 kcal/mol), ACE2 (-11.1 kcal/mol), and main protease (-9.6 kcal/mol). The ligand-amino acid interactions of sotetsuflavone in the active regions of the targets were similar to positive controls. Molecular dynamics simulations of the sotetsuflavone complex with two targets (ACE2 and main protease) are also relatively stable, but it is still advisable to investigate for a longer time. Sotetsuflavone is found in many different plant species such as Amentotaxus yunnanensis, Torreya yunnanensis, Dacrydium araucarioides etc.[25] Besides, the predictive pharmacokinetic parameters are relatively good: intestinal absorption is better than 80 %, no hepatotoxicity, AMES, and no skin toxicity. Therefore, this is a potential natural compound to treat COVID-19 in the future.

Morellic acid isolated from Garcinia morella, Garcinia hanburyi, exhibits anti-HIV-1 activity.[36,37] This is a relatively new compound that has not been studied much. Therefore, up to the present time, there have been no published studies on the potential of morellic acid to inhibit the SARS-CoV-2 virus. Our study has shown that morellic acid has the potential to give the most negative binding energy to two spike protein and ACE2 targets out of 273 compounds screened. Morellic acid also gave a very good docking score, inhibiting the main protease target with -9.5 kcal/mol, only inferior to sotetsuflavone (-9.6 kcal/mol). However, the inhibitory activity of this compound on RdRp was not as good as that of other compounds, only approximately equal to that of the positive control artemisinin (-7.8 kcal/mol).

Digitoxin is an FDA-approved drug for the treatment of cardiovascular disease, isolated from Digitalis purpurea.[28] Digitoxin was reported to be active against SARS-CoV-2-infected Vero cells based on their cytotoxic effects, with IC_{50} = 0.23 µM.[29] In addition, several cardiac glycosides have shown antiviral activity through inhibition of NKA and activation of tyrosine kinase (Src). Src regulates nuclear factor kappa B, which is an important transcription factor for SARS-CoV-2.[30] In our study, the results of digitoxin screening with molecular targets of the SARS-CoV-2 virus gave positive results. The docking score showed that the inhibition of all four targets of digitoxin was so good, especially the spike protein target (-10.3 kcal/mol). However, given the known toxicity of digitoxin, more research is needed to evaluate its potential as well as application in the treatment of COVID-19.

Hinokiflavone is a flavonoid compound, isolated from many different species such as Juniperus rigida, Platycladi cacumen, Selaginella hyophytris etc.[31-33] Hinokiflavone has been reported with anticancer, [34] anti-HIV-1 activity with IC_{50} = 65 µM.[35] The molecular docking results of this compound were very potent, strongly inhibiting all four targets, outperforming the positive controls. ADMET predicts relatively well, is also a potential candidate to inhibit SARS-CoV-2.

Similar to hinokiflavone, three natural compounds, mulberroside C, daphnoretin, hypericin also have anticancer effects against some viruses such as herpes simplex type 1, HIV-1, Enterovirus A71, Sindbis etc.[36-39] The SARS-CoV-2 of these compounds has not yet been studied. In this study, all three antiviral compounds gave good docking scores. Although the docking score of daphnoretin is not as good as that of the compounds listed in Table 1, the prediction of human intestinal absorption of this compound is up to 100 %. Given these results, further studies should be conducted to evaluate the potential for the treatment of COVID-19.

4. CONCLUSION

In this study, we found seven natural antiviral compounds: sotetsuflavone, hinokiflavone, mulberroside C, daphnoretin, morellic acid, digitoxin, and hypericin strongly inhibited spike
In silico screening of natural antivirals...

protein, ACE2, main protease and RdRp of the SARS-CoV-2 virus. Among them, sotetsuflavone is the most potent compound due to its strong inhibition of these four targets, drug-like properties, good pharmacokinetic parameters, and low toxicity. In addition, the results of molecular dynamics simulations of the sotetsuflavone-target complex are quite stable. Therefore, it is necessary to conduct further tests on the ability of sotetsuflavone to inhibit SARS-CoV-2 as well as the six compounds mentioned above.

REFERENCES

1. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu. A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med., 2020, 382, 727-733.

2. D. A. Schwartz, A. L. Graham. Potential maternal and infant outcomes from coronavirus 2019-nCoV (SARS-CoV-2) infecting pregnant women: lessons from SARS, MERS, and other human coronavirus infections, Viruses, 2020, 12, 194.

3. A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, Cell Host Microbe., 2020, 27, 325-328.

4. R. Hilgenfeld. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design, FEBS J., 2014, 281, 4085-4096.

5. M. Letko, A. Marzi, V. Munster. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, Nat. Microbiol., 2020, 5, 562-569.

6. Y. Huang, C. Yang, X. -F. Xu, W. Xu, S. -W. Liu. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19, Acta. Pharmacol Sin., 2020, 41, 1141-1149.

7. Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K.-Y. Yuen. Structural and functional basis of SARS-CoV-2 entry by using human ACE2, Cell., 2020, 181, 894-904.

8. L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science, 2020, 368, 409-412.

9. Y. Gao, L. Yan, Y. Huang, F. Liu, Y. Zhao, L. Cao, T. Wang, Q. Sun, Z. Ming, L. Zhang. Structure of the RNA-dependent RNA polymerase from COVID-19 virus, Science, 2020, 368, 779-782.

10. W. Zhu, C. Z. Chen, K. Gorkhov, M. Xu, D. C. Lo, W. Zheng. RNA-dependent RNA polymerase as a target for COVID-19 drug discovery, SLAS Discov., 2020, 25, 1141-1151.

11. D. Rubin, K. Chan-Tack, J. Farley, A. Sherwat. FDA approval of remdesivir-a step in the right direction, N. Engl. J. Med., 2020, 383, 2598-2600.

12. C. J. Gordon, E. P. Tchesnokov, J. Y. Feng, D. P. Porter, M. Götte. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus, J. Biol. Chem., 2020, 295, 4773-4779.

13. C. Ma, M. D. Sacco, B. Hurst, J. A. Townsend, Y. Hu, T. Szeto, X. Zhang, B. Tarbet, M. T. Marty, Y. Chen, J. Wang. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease, Cell Res., 2020, 30, 678-692.

© 2022 Vietnam Academy of Science and Technology, Hanoi & Wiley-VCH GmbH www.vjc.wiley-vch.de
SARS-CoV-2 spike protein is more resistant to hydroxychloroquine: potential repurposing of artenimol for COVID-19. *J. Biomol Struct Dyn.*, 2021, 39, 6184-6194.

21. G. G. Zhang, Y. Jing, H. M. Zhang, E. L. Ma, J. Guan, F. N. Xue, H. X. Liu, X. Y. Sun. Isolation and cytotoxic activity of selaginellin derivatives and biflavonoids from *Selaginella tamariscina*, *Planta Med.*, 2012, 78, 390-392.

22. A. G. Harrison, T. Lin, P. Wang. Mechanisms of SARS-CoV-2 transmission and pathogenesis, *Trends Immunol.*, 2020.

23. S. Wang, Y. Yan, Z. Cheng, Y. Hu, T. Liu. Sotetsuflavone suppresses invasion and metastasis in non-small cell lung cancer A549 cells by reversing EMT via the TNF-α/NF-κB and PI3K/AKT signaling pathway, *Cell Death Discov.*, 2018, 4, 1-11.

24. P. Coulerie, M. Nour, A. Maciuk, C. Eydoux, J.-C. Guillemot, N. Lebouvier, E. Hnawia, K. Leblanc, G. Lewin, B. Canard. Structure-activity relationship study of biflavonoids on the Dengue virus polymerase DENV-NS5 RdRp, *Planta Med.*, 2013, 79, 1313-1318.

25. D. J. Newman, G. M. Cragg, K. M. Snader. Natural products as sources of new drugs over the period 1981-2002, *J. Nat. Prod.*, 2003, 66, 1022-1037.

26. H. N. Murthy, D. Dalawai, Y. H. Dewir, A. Ibrahim. Phytochemicals and Biological Activities of *Garcinia morella* (Gaertn.) Desr.: A Review, *Molecules.*, 2020, 25, 5690.

27. V. Reut rakul, N. Anantachoke, M. Pohmakot, T. Jaipetch, S. Sophsan, C. Yoosook, J. Kasist, C. Nanpaswat, T. Santisuk, P. Tuchinda. Cytotoxic and anti-HIV-1 caged xanthones from the resin and fruits of *Garcinia hanburyi*, *Planta Med.*, 2007, 73, 33-40.

28. E. Navarro, P. Alonso, S. Alonso, J. Trujillo, C. Pérez, M. Toro, M. Ayuso. Cardiovascular activity of a methanolic extract of *Digitalis purpurea* spp. heywoodii, *J. Ethnopharmacol.*, 2000, 71, 437-442.

29. S. Jeon, M. Ko, J. Lee, I. Choi, S. Y. Byun, S. Park, D. Shum, S. Kim. Identification of antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs, *Antimicrob Agents Chemother.*, 2020, 64, e00819-00820.

30. K. F. C. Souza, B. P. T. M. Souza, I. C. N. de Palmer, P. Burth, A. R. Silva, C. F. Gonçalves-de-Albuquerque. Na/K-ATPase as a Target of Cardiac Glycosides for the Treatment of SARS-CoV-2 Infection, *Front Pharmacol.*, 2021, 12.

31. O. Kunert, R. C. Swamy, M. Kaiser, A. Presser, S. Buzzi, A. A. Rao, W. Schühly. Antiplasmodial and leishmanicidal activity of biflavonoids from Indian *Selaginella bryopteris*, *Phytochem Lett.*, 2008, 1, 171-174.

32. B. Zhuang, Z. -M. Bi, Z. -Y. Wang, L. Duan, E. -H. Liu. Chemical profiling and quantitation of bioactive compounds in *Platycladi cacumen* by UPLC-Q-TOF-MS/MS and UPLC-DAD, *J. Pharm. Biomed Anal.*, 2018, 154, 207-215.

33. S. Lee, N. -I. Park, S. -K. Bong, J. Jegal, S. -A. Park, S. -N. Kim, M. H. Yang. Ameliorative effects of Juniperus rigidu fruit on oxazolone- and 2,4-dinitrochlorobenzene-induced atopic dermatitis in mice, *J. Ethnopharmacol.*, 2018, 214, 160-167.

34. J. -F. Goossens, L. Goossens, C. Bailly. Hinokiflavone and Related C–O–C-Type Biflavonoids as Anti-cancer Compounds: Properties and Mechanism of Action, *Nat. Prod. Bioprospect.*, 2021, 1-13.

35. Y. -M. Lin, H. Anderson, M. T. Flavin, Y. -H. S. Pai, E. Mata-Greenwood, T. Pengsuparp, J. M. Pezzuto, R. F. Schinazi, S. H. Hughes, F.-C. Chen. In vitro anti-HIV activity of biflavonoids isolated from *Rhus succedanea* and *Garcinia multiflora*, *J. Nat. Prod.*, 1997, 60, 884-888.

36. P. Agostinis, A. Vantieghem, W. Merlevede, P. A. de Witte. Hypericin in cancer treatment: more light on the way, *Int J Biochem Cell Biol.*, 2002, 34, 221-241.

37. W. S. Ho, J. Y. Xue, S. S. Sun, V. E. Ooi, Y. L. Li. Antiviral activity of daphnoretin isolated from *Wikstroemia indica*, *Phytother Res.*, 2010, 24, 657-661.

38. C. -L. Lu, Y. -M. Li, G. -Q. Fu, L. Yang, J. -G. Jiang, L. Zhu, F. -L. Lin, J. Chen, Q. -S. Lin. Extraction optimisation of daphnoretin from root bark of *Wikstroemia indica* (L.) CA and its anti-tumour activity tests, *Food Chem.*, 2011, 124, 1500-1506.

39. Y. Cao, E. Lei, L. Li, J. Ren, X. He, J. Yang, S. Wang. Antiviral activity of Mulberroside C against enterovirus A71 in vitro and in vivo, *Eur. J. Pharmacol.*, 2021, 174204.

**Corresponding author:** **Bui Thanh Tung**

Department of Pharmacology and Clinical Pharmacy  
VNU University of Medicine and Pharmacy, Vietnam National University Hanoi  
Office 506, Y1 Building, 144 Xuan Thuy, Cau Giay, Hanoi 10000, Viet Nam  
E-mail: tungasia82@yahoo.es.

© 2022 Vietnam Academy of Science and Technology, Hanoi & Wiley-VCH GmbH  www.vjc.wiley-vch.de 222