Virtual screening and molecular dynamics simulation analysis of Forsythoside A as a plant-derived inhibitor of SARS-CoV-2 3CLpro

Shabana Bibia,b,1,*, Muhammad Saad Khanb,c, Sherif A. El-Kafrawyb,1, Thamir A. Alandijandy, Mai M. Eldalyy, Leena H. Bajraid,f, Esam I. Azhard,e,⇑

a Department of Biosciences, Shifa-Tameer-e-Milat University, Islamabad, Pakistan
b Yunnan Herbal Laboratory, College of Ecology and Environmental Sciences, Yunnan University, Kunming 650091, Yunnan, China
c Department of Biosciences, COMSATS University Islamabad, Sahiwal, Pakistan
d Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia
e Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
f Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is a more severe strain of coronavirus (CoV) that was first emerged in China in 2019. Available antiviral drugs could be repurposed and natural compounds with antiviral activity could be safer and cheaper source of medicine for SARS-CoV-2. 78 natural antiviral compounds database was identified from literature and virtual screening technique was applied to identify potential 3-chymotrypsin-like protease (3CLpro) inhibitors. Molecular docking studies were conducted to analyze the main protease (3CLpro) and inhibitors interactions with key residues of active site of target protein (PDB ID: 6LU7), active site constitute the part of active domain I and II of 3CLpro. 10 compounds with highest dock score were subjected to calculate ADMET parameters to figure out drug-likeness. Molecular dynamic (MD) simulation of the selected lead was performed by Amber simulation package to understand the conformational changes in docked complex. MD simulations analysis (RMSD, RMSF, Rg, BF, HBs, and SASA plots) of lead bounded with 3CLpro, hence revealed the important structural turns and twists during MD simulations from 0 to 100 ns. MM-PBSA/GBSA methods has also been applied for the estimation binding free energy (BFE) of the selected lead-complex. The present study has identified lead compound “Forsythoside A” an active extract of Forsythia suspense as SARS-CoV-2 3CLpro inhibitor that can block the viral replication and translation. Structural analysis of target protein and lead compound performed in this study could contribute to the development of potential drug against SARS-CoV-2 infection.

© 2022 Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Viral infections are considerable threats for the public health. Severe acute respiratory syndrome coronavirus – 2 (SARS-CoV-2) / COVID-19 is currently a critical human ailment spreading around the world. It appeared first in Wuhan city, China, in December 2019, which is famous for several kinds of international import and export (Bogoch et al., 2020; Lu et al., 2020). COVID-19 is badly affecting about two hundred countries around the globe (Hamdy M. Youssef et al., 2020), total reported cases 255,324,963 and the number of deaths are 5,127,696 by the current statistics until 19th November 2021 (Organization World Health, 2021).

SARS-CoV-2 is a single-stranded RNA-virus (29,891 bp), comprising 96% homology to bat coronavirus (CoV) whole genome sequence, and presented 79.6% sequence similarity with SARS-CoV

https://doi.org/10.1016/j.jsps.2022.05.003
1319-0164/© 2022 Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
SARS-CoV-2 translates spike (S protein) - associated receptor domain which could interact to the human angiotensin-converting enzymes 2 (ACE-2) and encourages the fusion of membrane and approval of virus into cells for example in the lung by endocytosis (Corum and Zimmer, 2020; Ou et al., 2020; Yan et al., 2020; Zhou et al., 2020). When a virus enters the cells, just like other viruses SARS-CoV-2 also controlled the nucleic acid machinery and weaken the synthesis of new viral proteins, and also involves in the assembly of the proteins and therefore take control of viral replication (Chen et al., 2020).

The key viral proteins of SARS-CoV-2 are papain-like protease (PLpro), 3-chymotrypsin-like protease (3CLpro), spike S proteins and RNA-dependent RNA polymerase (Wu et al., 2020). The S protein able to bind with ACE-2, permitting the virus to move into host. Efforts are continuously being applied to repropose several promising antiviral drugs, therefore, hydroxychloroquine, chloroquine phosphate (Cortegiani et al., 2020), arbidol (Khamitov et al., 2008), lopinavir (Yao et al., 2020), and remdesivir (Holshue et al., 2020; Wang et al., 2020) have been tested and presented significant biological activities against SARS-CoV-2 infections. Current information of clinical trials of SARS-CoV-2 drugs is available in database “ClinicalTrials.gov” (Pang et al., 2020). Alternative medicine and traditional Chinese medicine have shown promising results for SAR-CoV-2 infection in China (Yang et al., 2020). Indeed, phytochemicals isolated from traditional plants are always of great importance for the management of several diseases (Patel et al., 2020; Suwannarach et al., 2020). With all previous effort, specific antiviral therapy for COVID-19 is still missing.

Bioinformatics and mathematical modeling have been conducted to hypothesized the methods/protocols to discover the dynamics of transmission of the COVID-19 in the human population, and suggest the strategies to control the multiplication of new cases in Saudi Arabia, and such kind of mathematical models could give the idea to work for the large dataset in future (Yousef et al., 2021, 2022; Hamdy M. Yousef et al., 2020; Hamdy M. Yousef et al., 2020). Structural modeling of reproposed antiviral drugs has been the hot research topic after the emergence of COVID-19 (Bibi et al., 2022), because it’s very important technique of drug design and discovery, these studies usually predict the importance of drug and target interaction, target with disease interactions, and drug with disease interactions, and ultimately significant to find out the potential treatments in less time (March-Vila et al., 2017). Integration of ligand-based (chemicals/drug-based) and structure-based (enzyme/protein based) methods has been studied in several drug design studies previously and hence proved its significance in drug discovery for many diseases (Bian et al., 2017; Honarpvlar et al., 2014; Wilson and Lill, 2011). So that in this paper, an integrated drug design or computer-aided drug design approach (Bibi and Sakata, 2017a; Khan et al., 2021) was applied to selected chemical database for understanding the mechanism of protein–ligand interactions with the key residues, which could be helpful in finding potential natural main protease 3CLpro inhibitors with the significant activity to treat COVID-19.

2. Materials and methods

Several experimental attempts were made to investigate inhibitors against SARS-CoV-2 by utilizing various techniques explained in the previous literature (Bibi et al., 2021, 2020; Biswas et al., 2021). From different proteins of SARS-CoV-23CLpro is a unique target for identification of potential inhibitors as this protease is crucial for viral replication and therefore a promising target for drug (Anand et al., 2003; Mody et al., 2021). For inhibitors screening which could be further studied and developed as antivirals compounds through enzymatic assay optimum conditions identification are necessary and that’s conditions could be determined by biophysical and biochemical characterization of 3CLpro. Biochemical conditions data suggest that in the enzymatic assay of 3CLpro in the absence of salt basic pH or neutral conditions to be used (Ferreira and Rabeh, 2020). Results showed that against 3CLpro numerous natural products have shown encouraging inhibitory activities. Natural products like flavonoids (Park et al., 2016), terpenoids (J. Y. Park et al., 2012), diarylheptanoids (J.-Y. Park et al., 2012), and coumarins (Park et al., 2016) are potential inhibitors of the SARS-CoV proteases. By using the antiviral drugs database, computer-aided drug design (CADD) protocols were used to screen out potential SARS-CoV-2 3CLpro inhibitors (Fig. 1).

2.1. Construction of chemical database for in silico screening

Recent literature is available on the studies conducted to identify promising drugs for CoV infections (Chou et al., 2003; Farag et al., 2020; Hilgenfeld, 2014). Virtual screening techniques are highly recommended to find out the potential drugs for fatal and infectious diseases (Bibi and Sakata, 2016; Elmezayen et al., 2020; Jin et al., 2020a; Wang, 2020). An integrated CADD scheme has been employed using 78 natural compounds database to highlight the potential drug to combat COVID-19 infection. By utilizing the previous literature, botanical information, two-dimensional (2D) structures of selected 78 compounds along with the bioactivities for CoV infections have been enlisted in Table 1. The structure of each compound is drawn by using the ChemDraw software (Mills, 2006) and information of each structure was crosschecked from PubChem database (Kim et al., 2016) to reduce the chance of ambiguity and saved in SDF format for further analysis.

2.2. Selection of target protein

The selection of appropriate target protein for the drug design pipeline explains the important parameters necessary to clarify the action of bound ligands which could selectively inhibit the activity of 3CLpro to supress the replication of SARS-CoV-2 activity in the host (Jin et al., 2020b). With PDB ID 6LU7, a target protein is selected to perform a molecular docking experiment in this study to evaluate the protein–ligand binding interactions (Jin et al., 2020b).

2.3. Molecular docking analysis

Molecular docking is an important CADD application for the identification of protein–ligand interactions to understand the molecular mechanism of small drug-like entities in cellular pathways (Bibi and Sakata, 2016; Khan et al., 2021). Information of selected 78 natural antiviral compounds used for molecular docking is presented in Table 1, and a three-dimensional (3D) structure of selected target protein is presented in Fig. 2. Target protein (PDB ID: 6LU7) in pdb format was imported to the MOE software (Vilar et al., 2008). Heteroatoms, 3D protonation, and water molecules along with the default ligand (N3) attached in the 3CLpro target protein were removed to prepare protein for the docking procedure shown in Fig. 2. An active site is identified in the selected 3CLpro selected protein (6LU7) which constitute the region among active domain I and II as shown in Fig. 2 and structural optimization is performed by following parameters; such as the addition of hydrogen atoms, energy minimization with respect to 0.1 gradients, MMFF94X force field, chiral constraints and geometrical parameter. By using the surfaces and maps panel module, transparency of the front and the back surface is adjusted and resulted in the information of significant residues in the selected substrate binding site of 6LU7 protein in the native conformation (Jin et al., 2020b).
2020b; Naik et al., 2020). MOE software creates a database of natural 78 antiviral compounds collected from different reported studies to perform molecular docking simulations and saved with mdb extension for further analysis. Top-ranked poses are subjected for refinement and calculation of binding free energies (ΔG) which is evaluated by scoring function (GBVI/WSA dg) (Yousafi et al., 2020). A reliable scoring scheme that results in the docking score of the correct binding poses is established by a number of molecular interactions (hydrogen and hydrophobic interactions) (Aldeghi et al., 2016). MOE database of the docked complex was visualized carefully for understanding the mode of binding interactions of natural 3CLpro inhibitor bound in the selected pocket of the target protein.

2.4. Pharmacokinetic /ADMET profile estimation

The selected 10 top-scored compounds were used for the calculation of ADMET (absorption, distribution, metabolism, excretion, and toxicity) profile, and that is considered as an essential criterion for the drug-like screening of chemical compounds (Bibi and Sakata, 2017b; Fatima et al., 2018; Ntie-Kang et al., 2013). For the purpose of ADMET profile estimation of selected chemical compounds, SwissADME (Daina et al., 2014) and DataWarrior tools (Sander et al., 2015) were used.

2.5. Lead identification

An integrated drug design pipeline is designed in this study to identify the potential lead compound which could target 3CLpro for SARS-CoV-2 infections. 3CLpro is an acceptable target of CoV and have presented significance in blocking the viral replication and transcription procedure and becomes a striking therapeutic option for drug design of CoV infections (Hilgenfeld, 2014). The selected lead compound from the natural source presented significant results and fulfill maximum parameters of drug design, which includes biological activity, drug-likeness, pharmacokinetic profile, and best bounding to the active site residues of the protein, hence limiting the virus replication and inactivate the main protease 3CLpro of SARS-CoV-2.

2.6. Molecular dynamics simulations

The lead compound complex was used for further MD simulation analysis at 100 ns to understand the behaviour of the docked complex and conformational stability in the system for practical applications. MD simulation protocols were divided into three major categories such as preparation of parameter files as input, followed by the pre-processing phase, and end with the simulation time-phase (Ahmad et al., 2019).

In the initial phase of file preparation, by the use of AMBER20’s antechamber module (Lee et al., 2020), complex libraries with the parameters allied with 3CLPRO and Forsythoside A. In TIP3P solvation box, the complex system was solvated at 12 Å which was accomplished by the Leap module of AMBER. Molecular interactions were resolute by ff14SB force field (Case et al., 2014). For charge neutralization, Na+ ions were added.

While in the pre-processing step of the system, the selected complex energies were improved by the numerous turns, completely established number of hydrogen atoms were abated for five hundred steps, minimization of system energy was recorded for thousand steps with limitation of two hundred kcal/mol – Å2 on the residual arrangement, moreover, the minimization step was followed by the complete atoms minimization atleast once more for thousand steps with pragmatic restriction of five kcal/mol –Å2 used for the system carbon alpha atoms, and three-hundred steps of minimization were applied for the system non-heavy atoms with the restriction of hundred kcal/mol –Å2. The complex was heated progressively to three-hundred K by NVT ensemble parameter, preserved by the Langevin dynamics (Izaguirre et al., 2001) and SHAKE algorithm (Kráutler et al., 2001) to confine the hydrogen bonding. The equilibration stage of the complex was retrieved at hundred-ps. The system pressure during the simulation period was maintained using NPT ensemble and limit the system to Cx atoms of five kcal/mol – Å2. The MD simulation plots of hundred-ns were generated using
Table 1
Selected dataset of 78 natural compounds that possess coronavirus activity.

| Compound          | Chemical structure | CID      | Source                | Inhibitory constant (IC$_{50}$ or EC$_{50}$) | Ref                  |
|-------------------|--------------------|----------|-----------------------|----------------------------------------------|----------------------|
| 4′-O-methylbavachalcone | ![Structure](image) | 42607530 | *Psoralea corylifolia* | 10.1 μM                                       | (Kim et al., 2014)   |
| Aloe-emodin       | ![Structure](image) | 10207    | *Isatis indigotica*   | 8.3 μM                                        | (Lin et al., 2005)   |
| Amentoflavone     | ![Structure](image) | 5281600  | *Torreya nucifera*    | 8.3 μM                                        | (Ryu et al., 2010)   |
| Apigenin          | ![Structure](image) | 5280443  | *Houttuynia cordata*  | 0.4 μM                                        | (Yin et al., 2011)   |
| Bavachinin        | ![Structure](image) | 10337211 | *Psoralea corylifolia* | 38.4 μM                                       | (Kim et al., 2014)   |
| Beta-sitosterol   | ![Structure](image) | 222284   | *Isatis indigotica*   | 1.210 μM                                      | (Lin et al., 2005)   |
| Bilobetin         | ![Structure](image) | 5315459  | *Torreya nucifera*    | 72.3 μM                                       | (Ryu et al., 2010)   |
| Blancoxanthone    | ![Structure](image) | 11703574 | *Calophyllum blancoi* | 3 μM                                          | (Shen et al., 2005)  |
| Compound       | Chemical structure | CID       | Source               | Inhibitory constant (IC\textsubscript{50} or EC\textsubscript{50}) | Ref                |
|----------------|--------------------|-----------|----------------------|-------------------------------------------------|--------------------|
| Celastrol      | ![Celastrol structure](image1.png) | 122724    | Triterygium regeli   | 10.3 μM                                          | (Ryu et al., 2010) |
| Cepharanthine  | ![Cepharanthine structure](image2.png) | 10206     | Stephania tetrandra  | 0.83 μM                                          | (Kim et al., 2019) |
| Corylifol A    | ![Corylifol A structure](image3.png) | 25056407  | Psoralea corylifolia | 32.3 μM                                          | (Kim et al., 2014) |
| Dihydrotanshinone I | ![Dihydrotanshinone structure](image4.png) | 11425923  | Salvia miltiorrhiza  | 14.4 μM                                          | (J. Y. Park et al., 2012) |
| Emodin         | ![Emodin structure](image5.png) | 3220      | Rheum officinale     | 200 μM                                           | (Ho et al., 2007)  |
| Eucalyptol     | ![Eucalyptol structure](image6.png) | 2758      | Eucalyptus           | 0.61 nM                                          | (Sharma and Kaur, 2020) |
| Fangchinoline  | ![Fangchinoline structure](image7.png) | 73481     | Stephania tetrandra  | 1.01 μM                                          | (Kim et al., 2019) |

(continued on next page)
| Compound      | Chemical structure | CID    | Source          | Inhibitory constant (IC₅₀ or EC₅₀) | Ref          |
|---------------|--------------------|--------|-----------------|------------------------------------|--------------|
| Forsythoside A| ![Forsythoside A](image1.png) | 5281773| *Forsythia suspense* | 0.64 mM (Li et al., 2011) | (Li et al., 2011) |
| Ginkgetin     | ![Ginkgetin](image2.png) | 5271805| *Torreya nucifera* | 32.0 μM (Ryu et al., 2010) | (Ryu et al., 2010) |
| Hesperetin    | ![Hesperetin](image3.png) | 72281  | *Isatis indigotica* | 365 μM (Lin et al., 2005) | (Lin et al., 2005) |
| Iguesterin    | ![Iguesterin](image4.png) | 46881919| *Triterygium regelii* | 2.6 μM (Ryu et al., 2010) | (Ryu et al., 2010) |
| Indigo        | ![Indigo](image5.png) | 10215  | *Isatis indigotica* | 752 μM (Lin et al., 2005) | (Lin et al., 2005) |
| Tingenone     | ![Tingenone](image6.png) | 101520 | *Triterygium regelii* | 9.9 μM (Ryu et al., 2010) | (Ryu et al., 2010) |
| Isobavachalcone| ![Isobavachalcone](image7.png) | 5281255| *Psoralea corylifolia* | 7.3 μM (Kim et al., 2014) | (Kim et al., 2014) |
| Compound      | Chemical structure | CID      | Source            | Inhibitory constant (IC$_{50}$ or EC$_{50}$) | Ref                      |
|---------------|--------------------|----------|-------------------|--------------------------------------------|--------------------------|
| Jensenone     | ![Chemical Structure](image) | 11594161 | Eucalyptus        | Not determined                            | (Yang et al., 2010)      |
| Jubanine G    | ![Chemical Structure](image) | 122216365 | Ziziphus jujuba   | 13.41 μM                                   | (Kang et al., 2015)      |
| Jubanine H    | ![Chemical Structure](image) | 122216366 | Ziziphus jujuba   | 4.49 μM                                    | (Kang et al., 2015)      |
| Leptodactylone| ![Chemical Structure](image) | 442134   | Boenninghausenia sessilicarpa | 450 μM                                   | (Yang et al., 2007)      |
| Luteolin      | ![Chemical Structure](image) | 5280445  | Galla chinensis   | 10.6 μM                                    | (Yi et al., 2004)        |
| Lycorine      | ![Chemical Structure](image) | 72378    | Lycoris radiata   | 15.7 nM                                    | (Li et al., 2005)        |
| Methyl tanshinonate | ![Chemical Structure](image) | 14610613 | Salvia miltiorrhiza | 21.1 μM                                   | (J. Y. Park et al., 2012) |

(continued on next page)
| Compound          | Chemical structure | CID      | Source                | Inhibitory constant (IC$_{50}$ or EC$_{50}$) | Ref                  |
|-------------------|--------------------|----------|-----------------------|---------------------------------------------|----------------------|
| Myricetin         | ![Myricetin](image1.png) | 5281672  | Scutellaria baicalensis | 2.71 μM                                      | (Yu et al., 2012)    |
| Neobavaisoflavone | ![Neobavaisoflavone](image2.png) | 5320053  | Psoralea corylifolia   | 18.3 μM                                      | (Kim et al., 2014)   |
| Nummularine B     | ![Nummularine B](image3.png) | 51017057 | Ziziphus jujuba        | 6.17 μM                                      | (Kang et al., 2015)  |
| Pristimerin       | ![Pristimerin](image4.png) | 159516   | Tritygium regelii      | 5.5 μM                                       | (Ryu et al., 2010)   |
| Procyanidin A2    | ![Procyanidin A2](image5.png) | 6325839  | Cinnamomi cortex       | 29.9 μM                                      | (Zhuang et al., 2009) |
| Procyanidin B1    | ![Procyanidin B1](image6.png) | 11250133 | Cinnamomi cortex       | 41.3 μM                                      | (Zhuang et al., 2009) |
| Psoralidin        | ![Psoralidin](image7.png) | 5281806  | Psoralea corylifolia   | 4.2 μM                                       | (Kim et al., 2014)   |
| Compound                  | Chemical structure | CID       | Source                  | Inhibitory constant (IC<sub>50</sub> or EC<sub>50</sub>) | Ref                        |
|---------------------------|--------------------|-----------|-------------------------|----------------------------------------------------------|----------------------------|
| Pyranojacareubin          | ![Chemical structure](image1.png) | 15307925  | Calophyllum blancoi     | 15 μM                                                    | (Shen et al., 2005)        |
| Quercetin                 | ![Chemical structure](image2.png) | 5280343   | Houttuynia cordata      | 5.6 μM                                                   | (Yin et al., 2011)         |
| Quercetin 7-rhamnoside    | ![Chemical structure](image3.png) | 5748601   | Houttuynia cordata      | 0.03 μM                                                  | (Yin et al., 2011)         |
| Rosmariquinone            | ![Chemical structure](image4.png) | 160142    | Salvia miltiorrhiza     | 21.1 μM                                                  | (J. Y. Park et al., 2012)  |
| Sciadopitysin             | ![Chemical structure](image5.png) | 5281696   | Torreya nucifera        | 38.4 μM                                                  | (Ryu et al., 2010)         |
| Scutellarein              | ![Chemical structure](image6.png) | 5281697   | Scutellaria baicalensis | 0.86 μM                                                  | (Yu et al., 2012)          |
| Silvestrol                | ![Chemical structure](image7.png) | 11787114  | Aglaia foveolata        | 1.3 μM                                                   | (Müller et al., 2018)      |
| Compound               | Chemical structure | CID   | Source            | Inhibitory constant (IC$_{50}$ or EC$_{50}$) | Ref                                 |
|------------------------|--------------------|-------|-------------------|---------------------------------------------|-------------------------------------|
| Sinigrin               | ![Sinigrin structure](image) | 6911854 | *Isatis indigotica* | 217 μM                                      | (Lin et al., 2005)                  |
| Curcumin               | ![Curcumin structure](image) | 969516 | *Curcuma longa*   | 0.2 μM                                      | (J. Sharifi-Rad et al., 2020)       |
| Tanshinone I          | ![Tanshinone I structure](image) | 114917 | *Salvia miltiorrhiza* | 38.7 μM                                    | (J. Y. Park et al., 2012)           |
| Tanshinone IIA        | ![Tanshinone IIA structure](image) | 164676 | *Salvia miltiorrhiza* | 89.1 μM                                    | (J. Y. Park et al., 2012)           |
| Tanshinone IIB        | ![Tanshinone IIB structure](image) | 318797 | *Salvia miltiorrhiza* | 24.8 μM                                    | (J. Y. Park et al., 2012)           |
| Tetrandrine           | ![Tetrandrine structure](image) | 73078  | *Stephania tetrandra* | 0.33 μM                                    | (Kim et al., 2019)                  |
| Tetra-O-galloyl-D-glucose | ![Tetra-O-galloyl-D-glucose structure](image) | 5153644 | *Galla chinensis* | 4.5 μM                                      | (Yi et al., 2004)                   |
| Compound                  | Chemical structure | CID       | Source                      | Inhibitory constant (IC₅₀ or EC₅₀) | Ref                          |
|---------------------------|--------------------|-----------|-----------------------------|-----------------------------------|------------------------------|
| Theaflavin monogallates   | ![Theaflavin monogallates](image) | 169167    | *Camellia sinensis* (black tea) | Not determined                  | (Lung et al., 2020)         |
| Theaflavin 3,3'-digallate | ![Theaflavin 3,3'-digallate](image) | 135403795 | *Camellia sinensis* (black tea) | 9.5 µM                           | (Chen et al., 2005)         |
| Tomentin A                | ![Tomentin A](image) | 71659627  | *Paulownia tomentosa*       | 6.2 µM                           | (Cho et al., 2013)          |
| Tomentin B                | ![Tomentin B](image) | 71659628  | *Paulownia tomentosa*       | 6.1 µM                           | (Cho et al., 2013)          |
| Tomentin C                | ![Tomentin C](image) | 71659765  | *Paulownia tomentosa*       | 11.6 µM                          | (Cho et al., 2013)          |
| Tomentin D                | ![Tomentin D](image) | 71659766  | *Paulownia tomentosa*       | 12.5 µM                          | (Cho et al., 2013)          |
| Tomentin E                | ![Tomentin E](image) | 71659767  | *Paulownia tomentosa*       | 5.0 µM                           | (Cho et al., 2013)          |
| Tryptanthrin              | ![Tryptanthrin](image) | 73549     | *Strobilanthes cusia*       | 0.06 µg/mL                       | (Tsai et al., 2020)         |
| Thymoquinone              | ![Thymoquinone](image) | 10281     | *Nigella sativa*            | 0.3 µM                           | (Srinivasan, 2018)          |

(continued on next page)
| Compound            | Chemical structure | CID   | Source     | Inhibitory constant (IC₅₀ or EC₅₀) | Ref               |
|---------------------|--------------------|-------|------------|-----------------------------------|-------------------|
| Thymohydroquinone   | ![Structure](image) | 95779 | Nigella sativa | 3.1 μM                           | (Srinivasan, 2018) |
| Dithymoquinone      | ![Structure](image) | 398941| Nigella sativa | Not determined                    | (Srinivasan, 2018) |
| P-CYMENE            | ![Structure](image) | 7463  | Nigella sativa | 0.4354 μM                        | (Srinivasan, 2018) |
| Carvacrol           | ![Structure](image) | 10364 | Nigella sativa | 1.06 μM                          | (Srinivasan, 2018) |
| Terpinen-4-ol       | ![Structure](image) | 11230 | Nigella sativa | 30 μM                            | (Srinivasan, 2018) |
| Anethole            | ![Structure](image) | 637563| Nigella sativa | 4650 μM                          | (Srinivasan, 2018) |
| Longifolene         | ![Structure](image) | 289151| Nigella sativa | Not determined                    | (Srinivasan, 2018) |
| Cuminaldehyde       | ![Structure](image) | 326   | Cuminum cyminum | 50 μM                            | (Srinivasan, 2018) |
Table 1 (continued)

| Compound  | Chemical structure | CID     | Source          | Inhibitory constant (IC\textsubscript{50} or EC\textsubscript{50}) | Ref             |
|-----------|--------------------|---------|-----------------|---------------------------------------------------------------|-----------------|
| Rutin     | ![Rutin structure](image.png) | 5280805 | Morus nigra     | 0.012 µM                                                      | (Tokusoglu, 2016) |
| Myricetin | ![Myricetin structure](image.png) | 5281672 | Morus nigra     | 0.012 µM                                                      | (Tokusoglu, 2016) |
| Quercetin | ![Quercetin structure](image.png) | 5280343 | Morus nigra     | 0.00011 µM                                                   | (Tokusoglu, 2016) |
| Kaempferol| ![Kaempferol structure](image.png) | 5280863 | Morus nigra     | 0.028 µM                                                      | (Tokusoglu, 2016) |
| Chlorogenic acid | ![Chlorogenic acid structure](image.png) | 1794427 | Morus nigra     | 0.2 µM                                                      | (Tokusoglu, 2016) |
| Ferulic acid | ![Ferulic acid structure](image.png) | 445858  | Morus nigra     | 0.2 µM                                                      | (Tokusoglu, 2016) |
| Gallic acid | ![Gallic acid structure](image.png) | 370     | Morus nigra     | 0.06 µM                                                      | (Tokusoglu, 2016) |

(continued on next page)
Table 1 (continued)

| Compound            | Chemical structure | CID  | Source            | Inhibitory constant (IC50 or EC50) | Ref       |
|---------------------|--------------------|------|-------------------|------------------------------------|-----------|
| (+)-catechin        | ![Chemical structure](image1) | 9064 | Morus nigra       | 1.61 μM                            | (Tokusoglu, 2016) |
| (−)-Epicatechin     | ![Chemical structure](image2) | 72276| Morus nigra       | 0.407 μM                           | (Tokusoglu, 2016) |
| Epigallocatechin    | ![Chemical structure](image3) | 72277| Morus nigra       | 0.48 μM                            | (Tokusoglu, 2016) |
| Cyclomorusin        | ![Chemical structure](image4) | 5481969 | Morus nigra       | 1.7 μM                             | (Tokusoglu, 2016) |

Fig. 2. Structural representation of selected target protein 3-chymotrypsin-like protease (3CLpro) PDB ID: 6LU7 presented with the identified domains and target protein active site (structure covered by black net) by molecular operating environment (MOE) software.
a time scale of two-fs. Bound and unbound interactions analyses were performed by keeping the limiting distance of eight-Å. CPPTRAJ module (Roe and Cheatham, 2013) was cast-off at the end to do the statistical calculation for multiple structural considerations to ensure stability of the lead complex selected complex. Visual molecular dynamic (VMD) tool is used for the MD simulation trajectories (Humphrey et al., 1996).

2.7. Binding free energy (BFE) estimation

The interaction energy and solvation free energy for the selected 3CLPRO receptor, Forsythoside A, 3CLPRO″ Forsythoside A complex were estimated by using the significant module of MMPBSA.py (Miller et al., 2012) of software AMBER20. Moreover, it is reported that an average value of the system was recorded as gross BFE of the system. The BFE was calculated with MMPBSA module and its corresponding MM-GBSA calculations of AMBER modules with aims to develop the difference between the bound and unbound presentation of solvated conformations of the same molecule (Humphrey et al., 1996). Mathematically, the BFE can be calculated through Eq. (1).

\[ \Delta G_{\text{bind solv}} = \Delta G_{\text{bind vacuum}} + \Delta G_{\text{solv solv}} - \Delta G_{\text{3CLPRO} - \text{Forsythoside A}} - \Delta G_{\text{solv solv}} - \Delta G_{\text{solv frothyoside A}} + \Delta G_{\text{solv 3CLPRO}} \]  

(1)

For all three states of the system, the solvation energy was estimated by resolving either Poisson Boltzmann (PB) or Generalized Born (GB) equation and thus it will give the electrostatic role of the solvation state. It also permits the calculation of empirical terms for hydrophobic assists as shown in Eq. (2).

\[ \Delta G_{\text{solv solv}} = G_{\text{electrostatic}} \varepsilon = 80 - G_{\text{electrostatic}} \varepsilon = 1 + \Delta G_{\text{hydrophobic}} \]  

(2)

Calculation of the average interaction energy between 3CLPRO and Forsythoside A gives to delta- \[ \Delta G_{\text{vacuum}} \] (Eq. (3)).

\[ \Delta G_{\text{vacuum}} = \Delta E_{\text{molecular mechanics}} - T \Delta S \]  

(3)

3. Results

3.1. Database of selected natural compounds

Reported natural antiviral compounds isolated from different plant species were used to generate a database by MOE software (Vilar et al., 2008; Yousafi et al., 2020). 78 compounds with their chemical structures, PubChem identifier (CID), botanical information, and inhibitory constant values with respect to antiviral studies including SARS-CoV and MERS-CoV infections also are enlisted in Table 1. Each compound has been studied for anti-viral diseases and blocked virus entry into the host, virus replication, and translation activities. Therefore, in this study, computational techniques have been used to screen anti-SAR-CoV-2 3CLpro drugs, which could bind significantly within the substrate-binding pocket.

3.2. Protein-ligand interaction analysis

The most important step in CADD is molecular docking performed by MOE software (Vilar et al., 2008), ligand and protein structure are the preliminary requirement to start the molecular docking procedure so the previous knowledge of target protein is utilized, 3CLpro is a molecular target for the development of anti-COVID-19 drug (Hilgenfeld, 2014; Wu et al., 2020). The 3D structural information of the 3CLpro enzyme (PDB ID: 6LU7) has been used. The prepared enzyme/ protein structure without any bounded ligand and considerable substrate-binding active site for the molecular docking and interaction analysis was identified from literature (Jin et al., 2020b) as shown in Fig. 2.

Database of 78 natural compounds generated by MOE with.mdb extension was docked in the same active site by Dock module of MOE software (Vilar et al., 2008), lead compound ‘Forsythoside A’ with the best binding poses, and molecular interactions with the significant amino acids initiate the mechanism of inhibition of 3CLpro for SARS-CoV-2 infections are presented graphically in Fig. 3. Thr26 (A), Cys145 (A), Met165 (A), and Glu166 (A) are the key amino acids that play role in the hydrogen bonding interactions between the lead compound and the selected active binding pocket of 3CLpro.

By the evaluation of selected active binding site following residues (Thr24 (A), Thr25 (A), Thr26 (A), His41 (A), Met49 (A), Tyr54 (A), Phe140 (A), Leu141 (A), Asn142 (A), Gly143(A), Ser144 (A), Cys145 (A), His163 (A), His164 (A), Met165 (A), Glu166 (A), Leu167(A), Pro168 (A), His172 (A), Asp187 (A), Arg188 (A), Ala189 (A), Thr190(A), Ala191(A), Gln192(A)) were highlighted for significant activity (Jin et al., 2020b; Naik et al., 2020). Table 2 describes the summary of molecular docking analyses of selected top 10 compounds. Selected ten compounds have demonstrated the best binding poses with the highest dock score within the range of −8.6628 to −7.8328. All compounds have demonstrated a significant docked score along with the number of hydrogen bonds (HBs) within the range of 4.5 Å.

Compound 1 "Jubanine H" and Compound 2 "Forsythoside A" have presented very good docking results above threshold hold values of acceptable dock score which is −8.0 Kcal/mol and also presented significant binding interactions with the key amino acid residues of the active site of the target protein. Moreover, top scored six compounds enlisted in Table 2 have presented good energy value above −8.0 Kcal/mol and binding interactions with significant residues of substrate binding active site of the selected 3CLpro target enzyme.

3.3. Pharmacokinetic /ADMET profile estimation

Several studies have proved the importance of pharmacokinetic / ADMET profile estimation for the screening of databases to identify potential drug-like and a lead-like compounds that could be better tolerable in the design and development of new drugs (Bibi and Sakata, 2017b; Ntie-Kang et al., 2013; Sepay et al., 2020). ADMET properties are calculated by SwissADME (Daina et al., 2017) and Datawarrior tools (Sander et al., 2015). Physicochemical properties of the selected 10 compounds, such as partition coefficient/lipophilic parameters (logP values), hydrogen bond acceptor (HBA), hydrogen bond donor (HDB), total polar surface area (TPSA), molar refractivity (MR), and rotatable bond (RB) were calculated for each compound. As initial phase drug discovery protocols promote the calculation of drug likeness, water-solubility, pharmacokinetics, toxicity estimations along with the medicinal chemistry perspective as mentioned for top-scored ten compounds in Table 3.

Due to the large chemical structure of natural compounds in the selected dataset, physicochemical properties described by Lipinski (Lipinski et al., 1997) and Veber's (Veber et al., 2002) in the theory of drug likeness are violating one or more parameters shown in Table 3. The lipophilic profile and water solubility of the selected ten compounds are moderate or quite low values. Toxicity estimation of 10 selected compounds presented significant results in the form of mutagenic, tumorigenic, reproductive and irritant effects while four compounds showed minor violations. Five compounds have presented no alerts in the medicinal chemistry parameter evaluation while five compounds have presented one or two alerts. The pharmacokinetic properties of selected 10 compounds vary with respect to other.
drug-like parameters because most of the plant-derived compounds don’t fulfil the Lipinski rule of drug-likeness due to their large chemical structures (Katiyar et al., 2012; Yousafi et al., 2020).

3.4. Lead identification

Out of ten selected compounds based on top dock scores, it is hard to select a lead compound with all acceptable ADMET...
Table 3
Summary of pharmacokinetic/ADMET properties of 10 selected compounds.

| ADMET properties | Jubanine H | Forsythoside A | Theaflavin 3,3'-digallate | Jubanine G | Silvestrol | Nummularine B | Tetrandrine | Beta-sitosterol | Theaflavin monogallates | Ginkgetin |
|------------------|------------|----------------|--------------------------|------------|------------|---------------|-------------|----------------|------------------------|---------|
| Physicochemical Properties |            |                |                          |            |            |               |             |                |                        |         |
| MW               | 585.73     | 624.59         | 868.70                   | 571.71     | 654.66     | 591.70        | 622.75      | 414.71         | 716.60                 | 566.51  |
| RB               | 11         | 11             | 8                        | 10         | 11         | 10            | 4           | 6              | 5                      | 5       |
| HBA              | 7          | 15             | 20                       | 7          | 13         | 7             | 8           | 1              | 16                     | 10      |
| HBD              | 4          | 9              | 13                       | 4          | 4          | 4             | 0           | 1              | 11                     | 4       |
| MR               | 172.44     | 148.42         | 215.39                   | 167.64     | 161.84     | 173.34        | 186.07      | 133.23         | 179.69                 | 155.91  |
| TPSA             | 138.10     | 245.42         | 351.12                   | 138.10     | 171.83     | 138.10        | 61.86       | 20.23           | 284.36                 | 159.80  |
| B.S              | 0.17       | 0.17           | 0.17                     | 0.17       | 0.17       | 0.17          | 0.17        | 0.55           | 0.55                   | 0.17    |
| Lipophilicity    |            |                |                          |            |            |               |             |                |                        |         |
| LogP o/w (iLOGP) | 4.39       | 2.41           | 0.73                     | 2.23       | 4.35       | 3.16          | 4.87        | 4.79           | 1.75                   | 3.94    |
| LogP o/w (XLOGP3)| 2.89       | −0.50          | 4.71                     | 2.53       | 1.64       | 3.12          | 6.66        | 9.34           | 1.79                   | 5.69    |
| LogP o/w (WLOGP) | 0.55       | −1.12          | 3.53                     | 0.16       | 1.36       | 0.37          | 5.75        | 8.02           | 2.55                   | 5.74    |
| Consensus LogP o/w| 2.10       | −0.55          | 1.86                     | 1.40       | 1.94       | 1.86          | 5.41        | 7.19           | 1.33                   | 4.34    |
| Water solubility |            |                |                          |            |            |               |             |                |                        |         |
| Class            | Moderately soluble | Soluble                 | Moderately soluble       | Moderately soluble | Moderately soluble | Moderately soluble | Insoluble | Poorly soluble | Moderately soluble | Poorly soluble |
| Pharmacokinetics |            |                |                          |            |            |               |             |                |                        |         |
| BBB permeability | No         | No             | No                       | No         | No         | No            | No          | No             | No                     | No      |
| GI absorption    | High       | Low            | Low                      | Low        | High       | Low           | High        | Low            | Low                    | Low     |
| CYP1A2 inhibitor | No         | No             | No                       | No         | No         | No            | No          | No             | No                     | No      |
| P-gp substrate   | Yes        | Yes            | Yes                      | Yes        | Yes        | Yes           | No          | No             | No                     | No      |
| CYP2C19 inhibitor| No         | No             | No                       | No         | No         | No            | No          | No             | No                     | No      |
| CYP2C9 inhibitor | No         | No             | Yes                      | No         | No         | No            | No          | No             | Yes                    | Yes     |
| CYP2D6 inhibitor | No         | No             | No                       | No         | No         | No            | No          | No             | No                     | No      |
| CYP3A4 inhibitor | Yes        | No             | Yes                      | Yes        | Yes        | Yes           | No          | No             | Yes                    | No      |
| Log Kp (Skin permeation) | −7.82 | −10.46 | −8.25 | −7.99 | −9.13 | −7.69 | −5.37 | −2.20 | −9.40 | −5.72 |
| Toxicity estimation |            |                |                          |            |            |               |             |                |                        |         |
| Mutagenic        | None       | None           | None                     | None       | None       | None          | None        | None           | Low                    | None    |
| Tumorigenic      | None       | None           | None                     | None       | None       | None          | None        | None           | None                   | None    |
| Reproductive effects | None | None           | None                     | None       | None       | None          | None        | None           | None                   | None    |
| Irritant effects | Low        | None           | None                     | Low        | None       | Low           | None        | None           | None                   | None    |
| Medicinal Chemistry |            |                |                          |            |            |               |             |                |                        |         |
| PAINS            | No alert   | 1 alert        | 1 alert                  | No alert   | No alert   | No alert      | No alert    | No alert       | No alert               | No alert |
| Brenk            | No alert   | 2 alert        | 2 alert                  | No alert   | No alert   | No alert      | No alert    | No alert       | 1 alert                | No alert |
| Synthetic accessibility | 7.38 | 6.37 | 6.36 | 7.17 | 6.60 | 6.93 | 7.01 | 6.30 | 5.80 | 4.50 |
parameters until it is tested to formulate the final activity and dose of the selected drug. Because these are natural compounds in the selected dataset having large chemical structures, So ADMET parameters usually violate but as the molecular docking results are promising, the selected lead compound have generated 6 HBs and very notable binding energy (-8.6225 kcal/mol) and presented interacts with the significant residues of the active site (Fig. 3). Lead compound's medicinal status is confirmed from the literature also and founds effective with respect to potential anti-viral drugs for avian infections of bronchitis virus (Li et al., 2011; Wang et al., 2021), influenza virus (Law et al., 2017), and CoV (Bailly, 2021). Some recent studies highlight the importance of our selected lead compound "Forsythoside A" an active chemical isolated from Forsythia suspense and famous for different pharmacological activities and most commonly used traditional antiviral medicine (Bailly, 2021). While in this integrated drug design study, Forsythoside A has presented significant biochemical and structural activity as 3CLpro inhibitor for SARS-CoV-2 infections.

3.5. Molecular dynamics simulations

To validate the structural behaviour of lead compounds within the substrate-binding active cavity of 3CLpro was used to run MD simulations at 100 ns. 3CLpro protein stability is the description of all the net forces to determine whether the protein will remain in folded state or assume non-native congregating structures. Superimposition of 3CLpro at 0 ns and 100 ns is demonstrated in Fig. 4. Therefore, comparing the protein–ligand complexes conformation at different level of stimulation until 100 ns provides highly useful structural insights that help to understand the possible changes in the ligand pose generated as a result. Hence, a RMSD of 1.303 Å was noted during complex comparison at 0 ns (tan colour) and 100 ns (blue colour) simulation, which indicates minor divergence in the structural conformation due to movement in ligand pose depicted in Fig. 4.

During the simulation, 2D plots were generated to explain the fluctuating behaviour of the docked complex at different time frames during MD simulation productions. These plots are important for the MD simulations statistical analysis, hence it could be significant to decode the backbone stability and flexibility of the residues during the different time frames of MD simulations (Ismail et al., 2020). Results show that the RMSD calculates the small atoms convergence from a reference state (Fig. 5A), while the RMSF of a protein explained the protein’s dynamic nature that contributed to the system's overall versatility and residual mobility from its mean position (Fig. 5B). It is noted that the average RMSF value was 2.21 Å and the average RMSF value was 1.14 Å for the complex of Forsythoside A drug with 3CLpro target protein and almost similar residues were involved in the binding interactions at 0 ns and 100 ns, which seems maximum correlation with the docking results, and minor fluctuations of the 3CLpro protein residues were observed from the initial state during MD simulation runs.

The radius of gyration (Rg) was evaluated to confirm the compactness and equilibrium conformation of the system, it is expected that the high and low values of radius of gyration describe the magnitude of system compactness and system less tight packing. It also explains the system of concern if the weather is in order or not as required. A high compact system encourages MD simulations system stability. The estimated average Rg of the system was 22.70 Å with the maximum value of 38.45 Å at 19.19 ns reflecting the higher-ordered and compact nature of the system (Fig. 6A). The thermal residual deviation was assessed afterwards by beta-factor (BF) (Fig. 6B), therefore the results seem effectively correlated with RMSF and hereafter approves the stability of the system. BF and RMSF are complemented each other in terms of overall system stability and residual flexibility. The average BF of the system analyzed is 39.56 Å with a maximum of 346.31 Å.

The frequency of the HBs in each time frame of the MD simulation plots can be visualized in Fig. 7A. Hydrogen bonding plays an important role in overall complex stability, an increase in HBs increases the binding affinity of the drug molecule for the protein binding site and thus strengthen the complex stability. These HBs were extracted by means of VMD hydrogen bond plugin. The estimated maximum number of HBs between SARS-CoV-2 3CLpro residues and Forsythoside A atom were 2 and the minimum number of HBs was 1 respectively during MD simulation analysis until 0–100 ns. Solvent accessible surface area (SASA) that explains the thermodynamic stability of the protein was also calculated through the MD simulation run. SASA is a very useful analysis in which changes in the accessibility of the protein to solvent can be determined. A high peak of 157 nm2 was observed at 30 ns and later stabilized with an average value of accessibility 151.42 nm2 in Fig. 7B.

Fig. 4. Superimposition of 3CLpro complex with “Forsythoside A” at 0 ns and 100 ns. The calculated RMSD value is 1.303, complex at 0 ns is presented in a tan colour and 100 ns is presented in blue colour. Critical structural changes are shown in zoomed view.
3.6. Binding free energies (BFE) calculations

The understanding of net BFE or total energy (delta/ΔT energy) of the molecules in both simulation models is important to describe the favourable receptor - ligand bounded complex in biomolecular modeling. In the explicit solvent system, the net BFE for the 3CLpro - Forsythoside A bounded complex is −40.0981 kcal/mol (GB model) and −36.5514 kcal/mol (PB model), respectively. Special emphasis is for GB model and PB model based energy estimations during 100 ns MD simulation run, the estimated gas energy for the system is noticed as −116.2404 kcal/mol for GB model, hereafter for PB model, gas energy is noticed as −116.2404 kcal/mol that also the same value without any noticeable difference. Solvation free energy in the case of GB is 76.1423 kcal/mol while in the case of PB it is 79.6891 kcal/mol.

The electrostatic impact to the system assessed by molecular mechanics (MM) force field is very much acceptable to the net energy as −76.8056 kcal/mol. Similarly, the Van der Waals impact from MM also associated with the system stability estimated as −39.4349 kcal/mol. The electrostatic energy impact (presented as EGB and EPB) to the ΔG solvation energy is considerable factor focused to a non-acceptable impact in GB solvation energy. The surface area energy (ESURF) estimated is −5.3735 kcal/mol for GB model. In PB, ENPOLAR and EDISPER are the repellant and
attractive force based energies and is estimated as $-4.4689 \text{ kcal/mmol}$ and Zero kcal/mol. For each model, MM-based energy components for the selected bounded conformation of 3CLpro receptor and Forsythoside A are tabulated in Table 4.

### 4. Discussion

As it is a global pandemic, SARS-CoV-2 infections emerged as pneumonia and fever in China and now reached in other countries.
and a large number of people are affected and died in the last few months, WHO has described the large figures of SARS-CoV-2 / COVID-19 patients, deaths and recoveries around the globe along with the updates of vaccinations around the world (Organization World Health, 2021). As the virus was not limited to only one country or a state so its different stains make it difficult to develop any drug or vaccine which could completely treat the SARS-CoV-2 infection (Liu et al., 2020). Possible available treatments are antiviral drugs to treat different patients and another successful treatment is plasma technique as the final option for COVID-19 patients (Liu et al., 2020). Precautionary vaccination, wearing face- masks, use of sanitisers are highly encouraged to minimize the spread of COVID-19.

The previously available drugs repurposing is the most acceptable methodologies in the current scenario when there is no specific treatment for COVID-19, several in silico investigations have been performed to point out the potential drug-like therapeutics (Wu et al., 2020; Zhang et al., 2020). China has prescribed the Chinese traditional medicine and found effective outcomes in the control of COVID-19 along with the strict safety measures (Yang et al., 2020). The Chinese herbs were already famous for the treatment of respiratory diseases (Wang and Liu, 2014). Table 1 shows phytochemicals extracted from most of the Chinese medicinal plants, each compound is studied previously for CoV infections and other viral diseases, so this study is conducted to understand the mechanism of drug interaction with 3CLpro of SARS-CoV-2. It’s highly required to find out a potential drug which should be from the natural source because it would be safer and better tolerable with historical background to treat different illnesses and acceptable in different cultures (Mani et al., 2020).

The selected target protein (PDB ID: 6LU7) for this study is 3CLpro a novel target for SARS-CoV-2 infections, significant research has been conducted on achieving the main protease inhibitors which could be potential drugs for SARS-CoV-2 and several pharmaceutical companies has developed antiviral drugs from different sources for SARS-CoV-2 and most of them are yet in clinical trials (Liu et al., 2020; Tu et al., 2020). 3CLpro inhibitors for SARS-CoV-2 could block the entry of the virus into the host and stops their replication and translation activity (Báez-Santos et al., 2015; Ziebuhr et al., 2000). Molecular docking is performed to screen a potential 3CLpro inhibitor for SARS-CoV-2 by MOE software (Vilar et al., 2008) and resulted dock scores represent the best binding capacity of ligands in the substrate-binding active site of the target protein. Summary of docking results of 10 selected compounds is shown in Table 2 and 2D plots of protein–ligand interactions for the best-docked ligand is shown in Fig. 3.

ADMET profile of the selected top-scored 10 compounds is shown in Table 3. Each compound is evaluated with respect to drug-likeness, lipophilicity, water-solubility, pharmacokinetics, toxicity estimations along with the medicinal chemistry prospective to decrease the chance of rejection of a selected lead compound (Forsythoside A) in the drug development phase. Identified lead compound followed the maximum parameters of ADMET profile with a little violation. In this study, the selected lead compound highlighted the molecular mechanism of 3CLpro inhibition to provide the direction for the novel significant therapeutics in the prospective of SARS-CoV-2 infection. Several phytochemical studies have presented the idea that most of the natural compounds are highly acceptable but do not fulfill the criteria of drug-likeness, so it’s important to check the toxicity of compounds which is acceptable in the case of selected lead compounds (Katiyar et al., 2012; M. Sharifi-Rad et al., 2020). According to the molecular docking outcomes, it is concluded that the selected lead compound “Forsythoside A” from a natural source could be a significant 3CLpro inhibitor of SARS-CoV-2 infection (Fig. 3).

Selected lead compound by the In silico screening (Figs. 4–7) was subjected to MD simulations at 100 ns. Fig. 4 shows the superimposed structural analysis at 0 ns and 100 ns MD simulations of the selected lead compound bound to 3CLpro of SARS-CoV-2. It is clearly illustrating the surface of 3CLpro and conformational changes of the lead compound during MD simulations until 100 ns. To understand conformational changes of the selected complex during MD simulations at the atomic level, RMSD, RMSF, Rg, BF, HBs and SASA plots of 3CLpro bound with the selected lead compound are explained significantly and also described the superimposition of lead compound binding poses at 10 ns interval within 3CLpro cavity approving the selected pocket correlates the docked results.

RMSD measures of the 3CLpro-lead compound showed an increase for a short time span at 1.8 – 3.6 Å from the start of simulations till 0.8 ns then shows very little deviations. The system gets stabilize at average values of RMSD 2.21 Å is shown in Fig. 5 (A); RMSF plot showed fluctuation between 50 and 300 amino acid residues, most of the fluctuation were observed in the loop region in and around the binding cavity that might aid in the suitable conformational pose for binding with lead compound.

Comparative analysis of the plots generated at different ns of 3CLpro bound (3CLpro-lead compound) with unbound-state of 3CLpro revealed important structural turns and twists. By the application of CADD strategy, the following retrieved information will help the researcher to move towards the development of the selected lead compound “Forsythoside A” as a potential drug for SARS-CoV-2 infection.

5. Conclusions

Forsythoside A, a natural compound extracted from Forsythia suspensa, is selected as a lead compound by an integrated drug design pipeline. It is a medicinal compound affiliated with multiple pharmacological activities. Molecular docking investigation estimated ADMET profile and MD simulation analysis at 0–100 ns has depicted the maximum potential of lead compound as a substantial anti-SARS-CoV-2 drug with respect to 3CLpro inhibition mechanism. Moreover, its significance with respect to the medical perspective of viral infections has been studied previously while this in silico study confirms the activity of Forsythoside A as a 3CLpro inhibitor and could be repurposed for SARS-CoV-2 infections. It is highly recommended to test Forsythoside A in the laboratory for its inhibitory activity against 3CLpro of SARS-CoV-2.

Funding

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under grant no. (GCV19-46-1441). The authors, therefore, acknowledge with thanks to DSR for technical and financial support.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to express their gratitude to Dr. Mazhar Sadiq and Prof. Mohammad Amjad Kamal for their assistance and support to complete this work.
Li, H., Wu, J., Zhang, Z., Ma, Y., Liao, F., Zhang, Y.u., Wu, G., 2011. Forsythoside A inhibits the avian infectious bronchitis virus in cell culture. Phytomer Res. 25 (3), 338–342. https://doi.org/10.1007/s12265-011-9152-0.

Li, S.Y., Chen, L., Zhu, H., Huang, H.Y., Wang, H., Wang, L., Zhang, X., Hua, S.N., Yu, J., Xiao, P.G., Li, R.S., Tan, X., 2005. Identification of natural compounds with antiviral activities against SARS-associated coronavirus. Antiviral Res. 67, 18–23. https://doi.org/10.1016/j.antiviral.2005.02.001.

Lin, C.-W., Tsai, F.-J., Tsai, C.-H., Lai, C.-C., Wan, L., Ho, T.-Y., Hsieh, C.-C., Chao, P.-D., 2005. Anti-SARS coronavirus 3C-like protease effects of Isaits indigotica root and plant-derived phenolic compounds. Antiviral Res. 68 (1), 36–42. https://doi.org/10.1016/j.antiviral.2005.02.003.

Lipskina, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 23 (1–3), 3–25.

Lu, H., Zhu, Y., Li, H., Wang, Q., Li, Y., Watts, S.P., Carter, J.L., Smoot, J., Greig, A.C., Daniels, A.D., Jersey, S., Albaia, D., 2020. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. ACS Cent. Sci. 6 (3), 315–331. https://doi.org/10.1021/acscentsci.0c00272.

Lu, H., Stratton, C.W., Tang, Y.-W., 2020. Outbreak of pneumonia of unknown etiology in Wuhan. J Med Virol 92 (4), 401–402. https://doi.org/10.1002/jmv.25768.

Lung, J., Lin, Y.-S., Yang, Y.-H., Yeh, C.-Y., Shu, L.-H., Cheng, Y.-C., Li, H.T., Wu, C.-Y., 2020. The potential chemical structure of anti-SARS-CoV-2 RNA-dependent RNA polymerase. J. Med. Virol 92 (6), 693–697. https://doi.org/10.1002/jmv.25769.

Izaguirre, A., Catarello, D.P., Wozniak, J.M., Skeel, R.D., 2001. Languine stabilization of molecular dynamics. J. Chem. Phys. 114 (5), 2090–2098. https://doi.org/10.1063/1.1332996.

Mani, J.S., Johnson, J.B., Steel, J.C., Broszczak, D.A., Neilsen, P.M., Walsh, B.K., Naiker, M., 2013. Natural product-derived pyranoxanthones as potential agents against coronaviruses: A review. Virus Res. 284, 197989. https://doi.org/10.1016/j.virusres.2020.197989.

March-Vila, E., Pinzi, L., Sturm, N., Tinveilla, A., Engkvist, O., Chen, H., Rastelli, C., 2017. On the integration of in silico drug design methods for drug repurposing. Front. Pharmacol. 298.

Miller, B.R., McGee, T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E., 2012. The SARS-CoV protease catalytic residues: Theoretical and experimental results. J. Chem. Theory Comput. 8 (9), 3314–3321. https://doi.org/10.1021/ct300418h.

Suwannarach, N., Kumla, J., Sujarit, K., Pattananandecha, T., Sanjumun, C., Lumyong, S., 2020. Natural bioactive compounds from fungi as potential candidates for anti-viral and immunomodulatory activities. Adv. Biochem. Biophys. 4, 12523136.v1.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., Zhou, Q., 2020. Structural basis for the interaction of SARS-CoV-2 by full-length human ACE2. Science (80-. 367 (6485), 1444–1448. https://doi.org/10.1126/science.abc6342.

Wang, J., 2020. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. J. Chem. Inf. Model. 60 (6), 3277–3286. https://doi.org/10.1021/acs.jcim.0c00131.

Wang, Y., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., Xiao, J., 2020. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 30 (3), 269–271. https://doi.org/10.1038/s41422-020-0282-0.

Vega-Rendon, D., Chazot, P.L., Kurt, B., Boyunegmez Tumer, T., Catarina Moreira, A., Sharopov, R., Rasoulzadeh, M., Martin, N., Cho, W.C., Calina, D., Sharifi-Rad, J., 2020b. Natural bioactive compounds from fungi as potential candidates for anti-viral and immunomodulatory activities. Adv. Biochem. Biophys. 4, 12523136.v1.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., Zhou, Q., 2020. Structural basis for the interaction of SARS-CoV-2 by full-length human ACE2. Science (80-. 367 (6485), 1444–1448. https://doi.org/10.1126/science.abc6342.

Wang, J., 2020. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. J. Chem. Inf. Model. 60 (6), 3277–3286. https://doi.org/10.1021/acs.jcim.0c00131.

Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., Xiao, J., 2020. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 30 (3), 269–271. https://doi.org/10.1038/s41422-020-0282-0.

Wang, X., Li, X., Wang, X., Chen, L., Ning, E., Fan, Y., Wang, H., Chen, T., Wang, W., 2021. Experimental study on the preparation and treatment of avian infectious bronchitis. Vet. Sci. Rev. 35, 523–531. https://doi.org/10.1177/0264410X211009.0002.

Whitley, R.J., Lill, M.A., 2011. Integrating structure-based and ligand-based approaches for computational drug design. Future Med. Chem. 3 (6), 735–750.

Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, X., Xu, Y., Li, M., Li, X., Zheng, M., Chen, L., Li, H., 2020. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharmacol. Sin. 10 (5), 766–788.

Yan, R., Zhang, Y., Li, X., Xu, L., Guo, Y., Zhou, Q., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science (380). 367 (6485), 1444–1448.
Yang, Q.-Y., Tian, X.-Y., Fang, W.-S., 2007. Bioactive coumarins from Boenninghausenia sessilicarpa. J. Asian Nat. Prod. Res. 9 (1), 59–65. https://doi.org/10.1080/10286020500558297.

Yang, Y., Islam, M.S., Wang, J., Li, Y., Chen, X., 2020. Traditional Chinese medicine in the treatment of patients infected with 2019-new coronavirus (SARS-CoV-2): A review and perspective. Int. J. Biol. Sci. 16 (10), 1708–1717. https://doi.org/10.7150/ijbs.45538.

Yang, Z., Wu, N., Fu, Y., Yang, G., Wang, W., Xu, Y., Effert, T., 2010. Anti-infectious bronchitis virus (ibv) activity of 1, 8-cineole: Effect on nucleocapsid (n) protein. J. Biomedol. Struct. Dyn. 28 (3), 323–330. https://doi.org/10.1080/07391102.2010.10507362.

Yao, T.-T., Qian, J.-D., Zhu, W.-Y., Wang, Y., Wang, G.-Q., 2011. In vitro and in vivo effects of Houttuynia cordata on infectious bronchitis virus. Avian Pathol. 40 (5), 491–498. https://doi.org/10.1080/03079457.2011.605107.

Yousaf, Q., Batool, J., Khan, M.S., Perveen, T., Sajid, M.W., Hussain, A., Mehmood, A., Saleem, S., 2020. In Silico Evaluation of Food Derived Bioactive Peptides as Inhibitors of Angiotensin Converting Enzyme (ACE). Int. J. Pept. Res. Ther. 27 (1), 341–349. https://doi.org/10.1007/s10989-020-10090-y.

Youssef, H.M., Alghamdi, N., Ezzat, M.A., El-Bary, A.A., Shawky, A.M., 2021. A proposed modified SEIR epidemic model to analyze the COVID-19 spreading in Saudi Arabia. Alexandria Eng. J. 61 (3), 2456–2470. https://doi.org/10.1016/j.aje.2021.06.095.

Youssef, H.M., Alghamdi, N.A., Ezzat, M.A., El-Bary, A.A., Shawky, A.M., 2020a. A modified SEIR model applied to the data of COVID-19 spread in Saudi Arabia. AIP Adv. 10, 7018–7044. https://doi.org/10.1063/5.0029698.

Youssef, H.M., Alghamdi, N.A., Ezzat, M.A., El-Bary, A.A., Shawky, A.M., 2020b. A new dynamical modeling SEIR with global analysis applied to the real data of spreading COVID-19 in Saudi Arabia. Math. Biosci. Eng. 17, 7018–7044.

Yu, M.-S., Lee, J., Lee, J.M., Kim, Y., Chin, Y.-W., Lee, J.-C., Keum, Y.-S., Jeong, Y.-J., 2012. Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, nsP13. Bioorganic Med. Chem. Lett. 22 (12), 4049–4054. https://doi.org/10.1016/j.bmcl.2012.04.081.

Zhang, D.-H., Wu, K.-L., Zhang, X., Deng, S.-Q., Peng, B., 2020. In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus. J. Integr. Med. 18 (2), 152–158. https://doi.org/10.1016/j.jim.2020.02.005.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X.L., Zheng, X.-S., Zhao, K., Chen, Q.-J., Deng, F., Liu, L.-L., Yan, B., Zhan, F.-X., Wang, Y.-Y., Xiao, G.-F., Shi, Z.-L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579 (7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7.

Zhuang, M., Jiang, H., Suzuki, Y., Li, X., Xiao, P., Tanaka, T., Ling, H., Yang, B., Saitoh, H., Zhang, L., Qin, C., Sugamura, K., Hattori, T., 2009. Procyanidins and butanol extract of Cinnamomi Cortex inhibit SARS-CoV infection. Antiviral Res. 82 (1), 73–81. https://doi.org/10.1016/j.antiviral.2009.02.001.

Ziebuhr, J., Snijder, E.J., Gorbalenya, A.E., 2000. Virus-encoded proteases and proteolytic processing in the Nidovirales. J. Gen. Virol. https://doi.org/10.1099/02317-8453.