

Natural compounds as potential inhibitors of novel coronavirus (COVID-19) main protease: An *in silico* study

Amaresh Mishra

Gautam Buddha University, Greater Noida, India

Yamini Pathak

Gautam Buddha University, Greater Noida, India

Gourav Choudhir

²Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

Anuj Kumar

Bioinformatics Laboratory, Uttarakhand Council for Biotechnology (UCB), Biotech Bhawan, Pantnagar, U.S. Nagar, Uttarakhand-263145, India

Surabhi Kirti Mishra

School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Vishwas Tripathi (✉ drvishwastripathi@gmail.com)

Gautam Buddha University, Greater Noida, India

Research Article

Keywords: SARS-CoV-2, COVID19, Anti-viral, Natural compounds, Mpro, novel coronavirus

Posted Date: June 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-22839/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

COVID-19 pandemic has now expanded over 213 nations across the world. To date, there is no specific medication available for SARS CoV-2 infection. The main protease (M^{Pro}) of SARS CoV-2 plays a crucial role in viral replication and transcription and thereby considered as an attractive drug target for the inhibition of COVID-19. Natural compounds are widely recognised as valuable source of antiviral drugs due to their structural diversity and safety. In the current study, we have screened twenty natural compounds having antiviral properties to discover the potential inhibitor molecules against M^{Pro} of COVID-19. Systematic molecular docking analysis was conducted using AutoDock 4.2 to determine the binding affinities and interactions between natural compounds and the M^{Pro} . Out of twenty molecules, four natural metabolites namely, amentoflavone, guggulsterone, puerarin, and piperine were found to have strong interaction with M^{Pro} of COVID-19 based on the docking analysis. These selected natural compounds were further validated using combination of molecular dynamic simulations and molecular mechanic/generalized/Born/Poisson-Boltzmann surface area (MM/G/P/BSA) free energy calculations. During MD simulations, all four natural compounds bound to M^{Pro} on 50ns and MM/G/P/BSA free energy calculations showed that all four shortlisted ligands have stable and favourable energies causing strong binding with binding site of M^{Pro} protein. These four natural compounds have passed the Absorption, Distribution, Metabolism, and Excretion (ADME) property as well as Lipinski's rule of five. Our promising findings based on *in-silico* studies warrant further clinical trials in order to use these natural compounds as potential inhibitors of M^{Pro} protein of COVID.

1. Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which primarily affects the lungs and shows certain types of pneumonia-like symptoms (Huang et al., 2020; Kumar et al., 2020). SARS-CoV-2 is a novel strain of coronavirus, first time emerged in December 2019, during an outbreak in Wuhan, China, and subsequently expanded to all over the world in a very short period of time (World Health Organization, 2020; Hendaus, 2020). The outbreak was declared a Public Health Emergency of International Concern by WHO on 30 January 2020 (WHO, 2020). As on June 02, 2020, this contagious disease has led to over 6, 140, 934 confirmed cases and 373, 548 fatalities (<https://covid19.who.int/>). To date, there is no specific treatment for this ongoing COVID-19 pandemic. Some preliminary study results investigated potential drug combination of Lopinavir and Ritonavir to treat COVID-19 infected patients, which were earlier used in human immunodeficiency virus (HIV) and SARS CoV or Middle East respiratory syndrome (MERS) coronavirus patients (Lu, 2020; Chu et al., 2004). The drugs which can specifically target the virus replication cycle and subsequent infection are urgently required to develop effective antiviral therapies as early as possible. Natural compounds due to the presence of enormous structural and chemical diversity, availability of more chiral centers, and relative biosafety are considered as an excellent source of drugs in several diseases including viral infections. This is further strengthened by the fact that around 45% of today's bestselling drugs have either originated from natural products or their derivatives (Lahlou, 2013). Natural compounds possess antiviral property could become a valuable resource in this regard. Liu et al. (2020) crystallized the COVID-19 main protease (M^{Pro}), which has been structured and repositioned in the Protein Data Bank (PDB) and is publicly accessible. SARS-CoV-2 main protease (M^{Pro}) is reported to play an inevitable role in virus replication and transcription, suggesting it to be a promising target for inhibition of the SARS-CoV-2 cycle (Boopathi et al., 2020; Lahlou, 2013; Xu et al., 2002). Keeping this in mind, in this study we have selected several natural compounds based on extensive literature (Zakaryan et al., 2017; Thayil et al., 2016; Jo et al., 2020).

In the present study, we screened and explored the potential of selected natural compounds to inhibit the M^{Pro} of COVID-19 using molecular docking, followed by MD simulations, MM/G/P/BSA free energy calculations, ADME, drug-likeness, target specific binding, and toxicity analysis validation.

2. Material And Methods

A flow chart of pipeline used in present study is summarized in Figure 1.

- 2.1 Literature survey and ligands selection

An extensive literature survey was conducted to select the natural compounds having antiviral properties from different medicinal plants using PubMed and Google scholar platforms. Based on the literature survey, total twenty natural compounds were selected, and their chemical structures were extracted from PubChem (Kim et al., 2016) repository in SDF format. List of all selected natural compounds along with their corresponding chemical Ids, 2D and 3D structures is presented in Table 1. In order to prepare the ligands to perform molecular docking, hydrogen atoms were added followed by PDB structure generation by OPENBABEL program (O'Boyle et al., 2008). Further, all the molecules were allowed for the energy minimization and optimization using universal force field at 200 descent steepest algorithm of OPENBABEL available in PyRx (<https://pyrx.sourceforge.io/>) and converted in pdbqt format.

- 2.2 Preparation of protease

The 3D coordinates of main M^{Pro} of SARS-CoV-2 was obtained from the RCSB-PDB repository with PDB ID 6LU7 (Jin et al., 2020). In order to prepare the macromolecule for docking, water and other nonspecific molecules were removed by using UCSF CHIMERA (Pettersen et al., 2004). For protein protonation maintaining cellular pH, polar hydrogen atoms were added to the 3D structure model of M^{Pro} . The structure optimization and energy minimization were performed by using SPDB viewer (Guex & Peitsch, 1997). While, clean geometry module embedded in Discovery Studio package was utilized for the side chain angles correction.

- 2.3 Molecular docking

To identify new potential inhibitors against the M^{PRO} of SARS-CoV-2, the site-specific docking-screening of all selected natural compounds were carried out by AutoDock 4.2 (Forli et al., 2012). The box dimensions were kept as 70 Å × 70 Å × 70 Å with total of 50 genetic algorithm run. Other docking parameters were set as default. During performance of molecular docking, the amino acid residues including Thr25, Thr26, Gly143, Ser144, His163, His164, and Glu166 were utilized as the binding pocket sites. Other docking parameters of AutoDock 4.2 were set as default. The protein-ligand interactions were further rendered with the Maestro and Discovery studio programs.

- 2.4 ADME compound screening

An *in-silico* tool for analysis of absorption, distribution, metabolism, and excretion (ADME) was used to screen the above-mentioned compounds which could be bioactive via oral administration. Drug-like properties were calculated using Lipinski's rule of five using SWISSADME prediction (<http://www.swissadme.ch/>) (Lipinski et al., 2012)(Giménez et al., 2010) .

- 2.5 Target prediction

Molecular Target studies are important to find the macromolecular targets of bioactive small molecules. This is useful to understand the molecular mechanisms underlying a given phenotype or bioactivity, to rationalize possible side-effects, to predict off-targets (Enmozhi et al., 2020). In this direction, SwissTarget Prediction tool (<https://www.swisstargetprediction.ch>) was used (Daina et al., 2019). Canonical smile for Amentoflavone and Guggulsterone was entered and was analyzed.

- 2.6 Molecular dynamics (MD) simulations

The four representative docking complexes of ligands with M^{PRO} including amentoflavone, guggulsterone, puerarin, and piperine were used for further refinement using MD simulations analysis. MD simulation studies were carried out to find out the stability and flexibility of the natural compounds-M^{PRO} complexes on 50ns. The method used for the MD simulations of natural compounds-M^{PRO} complexes remains same as earlier described in recent studies (Gajula et al., 2016; Kumar et al., 2018; Jee et al., 2018; Kumar et al., 2020). All simulations of representative natural compounds-M^{PRO} complexes were conducted with the fruitful utilization of GROMOS96 43a1 force field available in GROMACS 5.1.4 suite (Van Der Spoel et al., 2005). Topology files for ligand molecules were created by using PRODRG server (Schüttelkopf & Van Aalten, 2004). The prepared protein complexes were solvated in a cubic box of edge length 10 nm along with SPC water molecules. In order to maintain the system neutrality, adequate numbers of ions were added. To remove the clashes between atoms, system energy minimization calculations were applied with the convergence criterion of 1000 kJ/mol/nm. Long-range interaction electrostatics (Abraham & Gready, 2011) was handled by using PME. For both van der Waals and Coulombic interactions, a cut-off radius of 9 Å was utilized. Equilibration was completed in two different phases. The solvent and ion molecules were kept unrestrained in first stage, while in the second stage the restraint weight from the protein and protein-ligand complexes was gradually declined, in NPT ensemble. LINCS constraints were applied to all bonds involving hydrogen atoms (Hess et al., 1997). The temperature and pressure of the system were kept at 300 K and 1 atm respectively by using Berendsen's temperature and Parrinello-Rahman pressure coupling respectively (Berendsen et al., 1995). The production simulation was initiated from the velocity and coordinates obtained after the last step of the equilibration step. All the systems were simulated for 200 ns and snapshots were taken at every 2 ps interval.

- 2.7 MM/PBSA free energy calculations

The calculations of binding energy of the M^{PRO}-ligand complexes were calculated by using MM/PBSA (Molecular Mechanics Poisson Boltzmann Surface Area) method. While calculations of MM-PBSA the polar part of the solvation energy was calculated by using the linear relation to the solvent accessible surface area. The g_mmpbsa module available in GROMACS was applied for the determination of different components of the binding free energy of complexes (Kumari and Kumar, 2014). Only the last 10 ns of data were utilized for the MM-PBSA analysis to considering the convergence issue associated with the calculations. In present study, entropy calculations were not calculated as they may change the numerical values of the binding free energy reported for the molecules. In the MM-PBSA calculation, the binding free energy between M^{PRO} and a ligand was calculated using following equations:

$$\Delta G_{MM/PBSA} = G_{complex} - G_{protein} - G_{ligand} \quad (1)$$

(1)

$$G_x = E_{MM} - TS_{MM} + \Delta G_{solv} \quad (2)$$

(2)

$$E_{MM} = E_{bonded} + E_{coul} + E_{LJ} \quad (3)$$

(3)

$$\Delta G_{solv} = G_{polar} + G_{nonpolar} \quad (4)$$

(4)

- 2.8 Toxicity analysis

Toxicity analysis of selected natural compounds was done by the ProTox-II web server (Banerjee et al., 2018). ProTox-II is a kind of virtual lab which integrates several parameters like molecular similarity, fragment propensities and most frequent features. It predicts various toxicity endpoints and incorporates a total of 33 models for the prediction of various toxicity aspects of small molecules.

- 2.9 Similar FDA approved drug compound search with SWISS similarity

The compounds which were giving the best binding energy among the selected natural compounds were checked for similarity, if any, with FDA approved drugs using SWISS similarity tool (<http://www.swissimilarity.ch>) (Zoete et al., 2016).

- 2.10 Computational facility details

The MD simulations and corresponding energy calculations were carried out on HP Gen7 server with 48 Core AMD processors and 32GB of RAM.

3. Results

- 3.1 Determination of Active Sites: Table 1 shows the structure and amino acids found in the active site pockets of 6LU7. 6LU7 is the main protease (Mpro) found in COVID-19, which has been structured and repositioned in PDB and can be accessed by the public, as of early February 2020.
- 3.2 ADME (Absorption, distribution, metabolism, and excretion): ADME properties were found by obtaining the canonical smiles from PubChem. These smiles were used to identify ADME properties using SWISS ADME. Then compounds were analyzed on various parameters like lipophilicity, molecular weight, hydrogen-bond donors, hydrogen-bond acceptors, Clog P-value, Ghosh violations, Lipinski violations, etc. Ligands/natural compounds have been selected based on adherence to soft or classical Lipinski's rule of five. The selected ligands that did not incur more than 2 violations of Lipinski's rule were further used in molecular docking experiments with the target protein. The drug scanning results (Table 2) show that most of the tested compounds in this study was accepted by Lipinski's rule of five. These compounds were selected for docking to find their binding affinity with COVID19 main protease Mpro. List of compounds with suitable ADME properties given below:

S.N.	Compound Name	Mol. Weight (g/mol)	Concensus Log Po/w	Num. H-bond acceptors	Num. H-bond donors	Molar Refractivity	Lipinski	Veber	Bioavailability score	Synthetic accessibility (SA)	TPSA (Å ²)	No of rotatable bonds	Solubility (mg/ml)
1	Ritonavir	720.94	5.04	7	4	197.82	No	No	0.17	6.45	202.26	22	6.87e-08
2	Lopinavir	628.8	4.37	5	4	187.92	Yes	No	0.55	5.67	120.00	17	5.57e-08
3	Herbacetin	302.24	1.33	7	5	78.03	Yes	Yes	0.55	3.2	131.36	1	1.73e-01
4	Rhoifolin	578.52	-0.66	14	8	137.33	No	No	0.17	6.33	228.97	6	1.92e+01
5	Guggulsterone	312.45	4.03	2	0	93.54	Yes	Yes	0.55	4.79	34.14	0	8.14e-03
6	Cyanidin-3-o-galactoside	449.38	-1.16	11	8	108.29	No	No	0.17	5.27	193.44	4	5.23e+01
7	Xanthohumol	354.40	3.76	5	3	102.53	Yes	Yes	0.55	3.16	86.99	6	9.26e-03
8	Phloretin	274.27	1.93	5	4	74.02	Yes	Yes	0.55	1.88	97.99	4	1.16e-01
9	Crocetin	328.40	4.21	4	2	98.48	Yes	Yes	0.56	3.99	74.60	8	5.44e+01
10	Pectolarin	622.57	-0.43	15	7	148.29	No	No	0.17	6.63	227.20	8	3.48e+00
11	Apigenin	270.24	2.11	5	3	73.99	Yes	Yes	0.55	2.96	90.90	1	1.07e-02
12	Luteolin	286.24	1.73	6	4	76.01	Yes	Yes	0.55	3.02	111.13	1	4.29e-02
13	Amentoflavone	538.46	3.62	10	6	146.97	No	No	0.17	4.27	181.80	3	1.07e-06
14	Daidzein	254.24	2.24	4	2	71.97	Yes	Yes	0.55	2.79	70.67	1	2.64e-03
15	Puerarin	416.38	0.27	9	6	104.59	Yes	No	4.98	0.55	160.82	3	4.49e-01
16	Epigallocatechin	306.27	0.42	7	6	76.36	Yes	Yes	0.55	3.53	130.61	1	8.42e+00
17	Resveratrol	228.24	2.48	3	3	67.88	Yes	Yes	0.55	2.02	60.69	2	1.18e-01
18	Maslinic acid	472.70	5.24	4	3	137.82	Yes	Yes	0.56	6.22	77.76	1	2.37e-03
19	Piperine	285.34	3.04	3	0	85.47	Yes	Yes	0.55	2.92	38.77	4	2.87e-01
20	Ganomycin B	344.44	4.44	4	3	103.1	Yes	Yes	0.56	3.13	77.76	9	3.13e-02

Table 1. List of compounds with suitable ADME properties

- 3.3 Target prediction

The target prediction analysis was displayed for our two best compounds, Amentoflavone and Guggulsterone on the web page with the following observations the top 15 of the results were given as a pie-chart (Figure 2). The pie chart for Amentoflavone predicts 20% of Family AG protein-coupled receptor, 13.3% Kinase, 13.3% of Enzymes, 13.3% of unclassified protein, 6.7% of Phosphatase, 6.7% of protease, 6.7% of Oxidoreductase, 6.7% of primary active transporter, 6.7% of Secreted protein, 6.7% of Ligand-gated ion channel. The pie chart for Guggulsterone predicts 40% of Nuclear Receptors, 13.3% of Cytochrome P450, 13.3% of Secreted protein, 13.3% of Oxidoreductase, 6.7% of Membrane receptors, 6.7% of Fatty acid-binding protein family, 6.7% of Enzymes. The output table consisting of Target, Common Name, Uniprot ID, ChEMBL-ID, Target Class, Probability, and Known actives in 2D/3D are given in the Supplementary material. The possible sites of the target which the compound may bind to are mostly the targets which are predicted by the software and the probability score for Amentoflavone and Guggulsterone are obtained from 1.0 to 0.0868 & 1.0 to 0.101672 respectively. This makes an inference that the small compound may have high target attraction towards the specific binding site it is directed to.

- 3.4 Molecular docking

Molecular docking is an extensively used *in-silico* way to predict protein-ligand interaction. To perform the docking analysis, the structures and amino acids found in the active site pockets of 6LU7. 6LU7 is the main protease (M^{Pro}) found in COVID-19, which has been structured and repositioned in PDB databank. Thereafter, Ligand-protein docking was performed, and the interactions were determined based on the binding affinity of our compounds. Each individual analysis gave positive results, suggesting that the selected natural compounds may directly inhibit COVID-19 main protease M^{Pro}. The 14 selected natural compounds were docked with COVID-19 main protease M^{Pro} along with the standard ritonavir and lopinavir to compare the results. Further, like previous other findings, our results also indicated a good binding affinity of ritonavir and lopinavir to the COVID-19 main protease M^{Pro}. The results obtained are as follows:

Due to technical limitations, Table 2 cannot be displayed in the text. Please find Table 2 in the supplemental file section.

Table 2. Shows the molecular docking analysis results for selected natural compounds against COVID-19 main protease M^{Pro} (PDB-6LU7).

Figure 3 and Figure 4 can be found in the figures section.

Due to technical limitations, Table 3 cannot be displayed in the text. Please find Table 3 in the supplemental file section.

Table 3. Docking analysis visualization of COVID-19 main protease M^{Pro} (PDB-6LU7) binding with Lopinavir*, Ritonavir*, Amentoflavone, Guggulsterone, Puerarin, Piperine, Maslinic acid, Apigenin, Epigallocatechin, Daidzein, Xanthohumol, Resveratrol, Luteolin, Cyanidin-3-o-galactoside, Pectolarin, Herbacetin, Rhoifolin, Ganomycin B, Phloretin, and Crocetin. The 3D structures of protein-ligand interactions were visualized by discovery studio programs. The binding residues and their chains were identified from the protein-ligand complex as shown in the above images.

- 3.5 Molecular dynamics (MD) simulations

To further investigate the molecular docking results, the top four natural compound complexes namely, amentoflavone, guggulsterone, puerarin, and piperine were subjected to 50ns MD simulations. The conformational stability and flexibility of the complexes have been analyzed by using various parameters such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent accessible surface area (SASA), radius of gyration (Rg), and binding affinity of phytomolecule complexes by using mmpbsa and hydrogen bond formation ability. The RMSD is a commonly used similarity tool to measure the conformational perturbation during the simulation of macromolecule structures. RMSD of the C α Atoms related to the stability of the complexes. The time dependent RMSD from the initial stage of the simulation to 50 ns simulation. The RMSD of the backbone of these 4 complexes lies between 0.231–0.50 nm, which stabilizes at the 35 ns whereas the RMSD of ligands ranged from 0.35–0.96 nm (Figure 5a). RMSD of the protein backbone of all system was small and comparable, which may conclude that the binding of ligands does not lead to the confirmation perturbation during the simulation (Figure 5b). During MD simulations, RMSF define the residual flexibility from the average position. The RMSF of the protein ranged from 0.2–0.4 nm of all systems (Figure 5c). Some amino acid shows the high-intensity pick, which may represent a loop region. The presence of low-intensity pick revealed that binding of the phytomolecules does not affect the stability of the structural region of the enzyme.

In MD simulations, Rg determines the compactness of protein, induced by the movement of a ligand. The lower the flexibility of the Rg during the simulation associated with the structural stability of the protein. The Rg values of all phytochemical's complexes were lies 2.20–2.05 nm (Figure 5d). The Rg values of all four phytochemicals complexes support their consensus architecture as well as size. The SASA associated with the exposures of the hydrophobic residue during the simulation. SASA plays a principal role in the van der interaction. The SASA values of all systems were lies between 125–150 nm². SASA confirmations showed that the binding of ligand molecules does not affect the overall folding of the protein (Figure 6a).

In a complex protein and ligand, hydrogen bonding plays a critical role to determine the strength of interaction. During the simulation time, several hydrogen bonds formed between the donor and the acceptor group (Figure 6c). Two hydrogen bonds consistently formed during the time of simulation (Figure 6b). Over all observations indicated that all four complexes are stable during simulation.

- 3.6 Binding free energy calculation: In recent reports, it has been demonstrated that 100–200 snapshots are enough to calculate the binding free energies (Khan et al., 2020)(Kumar et al., 2020)(Sarma et al., 2020). The free binding energy results showed that as compare to other molecules Amentoflavens has maximum binding energy –240.434±17.602–180.787 kJ/mol whereas Guggulsterone has lower binding energy –121.708±12.423 kJ/mol (Table 5). In each molecule, the polar salvation and SASA energy showed moderate effects on binding energy component.

Name of molecules	Van der Waal energy (K.J./mol)	Electrostatic energy (K.J./mol)	Polar solvation energy (K.J./mol)	SASA energy (K.J./mol)	Binding energy (kJ/mol)
Amentoflavone	-350.08217.177	-104.69619.160	239.90621.251	-25.5621.010	-240.43417.602-180.787
Guggulsterone	-140.06815.668	-7.4089.330	37.62111.572	-11.8531.070	-121.70812.423
Piperine	-173.54511.759	-9.3734.129	50.5737.610	-13.9431.293	-146.287 11.205
Puerarin	-180.78720.912	-82.40516.508	148.20017.298	-16.6391.417	-131.63120.483

Table 4. Binding free energy calculation of four stable complexes during simulation

- 3.7 Toxicity analysis

In-silico toxicities of selected natural compounds were predicted by using ProTox-II. As shown in Table 4, ProTox-II toxicity prediction was done to check the safety of the compounds based on two major toxicity end points, hepatotoxicity & cytotoxicity. According to the toxicity analysis, none of the selected natural compounds showed potential hepatotoxicity or cytotoxicity except Pectolarin which showed potential cytotoxicity.

Due to technical limitations, Table 5 cannot be displayed in the text. Please find Table 5 in the supplemental file section.

Table 5: Toxicity predictions for selected natural compounds

- 3.8 Similarity checks with FDA approved drugs using SWISS similarity

We further checked the similarity of our two top hit natural compounds Amentoflavone and Guggulsterone if any, with the FDA approved drugs using SWISS similarity check. Swiss Similarity web tool is used for rapid ligand-based virtual screening. For Amentoflavone, we did not find any reported similar FDA approved drug in Swiss Similarity database. Whereas for Guggulsterone, we found 117 FDA proved drugs. The output table consisting of Drug ID, Drug name, Similarity score and molecule structure are given in the Supplementary material. The FDA approved drug structure obtained with the similar structure of Guggulsterone which are predicted by the software and the probability score for obtained from 0.995 to 0.009. This makes an inference that that these compounds could be very important and unique with pharmaceutical perspectives and need to be explored at in vitro and subsequent pre-clinical and clinical trials.

Due to technical limitations, Table 6 cannot be displayed in the text. Please find Table 6 in the supplemental file section.

Table 6: Swiss similarity prediction of Amentoflavone and Guggulsterone compounds

4. Conclusion

The ongoing Covid-19 pandemic caused by SARS-CoV-2 has shattered the whole world and created a situation of public health emergency (WHO, 2020). Keeping in the mind the immediate urgency, we are in dire need of some effective drug against the novel coronavirus Covid-19. Since the virus is new to the Human population therefore the information regarding the ins and outs of this virus is very limited. Considering the scenario, we can get the possible lead from the old SARS virus (SARS-CoV-1) emerged in 2003. Fortunately, in a recent study Liu et al. (2020) has revealed the crystal structure of SARS-CoV-2/COVID-19 main protease (M^{Pro}/3CL protease PDB-ID-6LU7 (Jin et al., 2020). This protease is considered an attractive target as it is essential for virus functionality, replication, and entry competence. The main protease M^{Pro} has been investigated as a potential target to inhibit previous coronavirus infections also like SARS and MERS (Jo et al., 2020). This study aimed to screen the natural compounds based on their pharmacokinetic properties, drug likeness and ability to specifically bind to the active sites of SARS-CoV-2 main protease so that these leads can be proposed as potential inhibitor to check the virus replication cycle. Lopinavir and Ritonavir are well known protease inhibitor of HIV (Israr et al., 2011). Both drugs were also recommended as repurposed drug in the treatment of SARS and Middle East respiratory syndrome (MERS) (Chu et al., 2004). Therefore, in this study we have taken these drugs as standard reference drugs to compare the efficacy of the binding of our selected compounds. In our *in-silico* prediction experiment, none of the selected compound showed hepatotoxicity and cytotoxicity except Pectolinarin which showed potential cytotoxicity. The compounds which were found to potentially inhibit the viral protease based on the binding energy were Amentoflavone, Guggulsterone, Puerarin, Piperine, Maslinic acid, Apigenin, Epigallocatechin, Daidzein, Xanthohumol, Resveratrol, Luteolin, Cyanidin-3-o-galactoside, Pectolinarin, Herbacetin, Rhoifolin, Ganomycin B, Phloretin, and Crocetin. Among these Amentoflavone and Guggulsterone were the top two leads showing lowest binding energy and satisfying our studied parameters. Therefore, we propose that these natural compounds may further be validated as potential inhibitors of COVID-19 main protease M^{Pro}. Our promising findings based on preliminary an *in-silico* analysis could become a basis for further studies at *in-vitro* and *in-vivo* levels in order to use these compounds as potential inhibitors of SARS-CoV-2 protease.

Declarations

Conflict of interest

The authors declare no conflict of interest.

Ethical standards

Ethical standards are compulsory for studies relating to human and animal subjects.

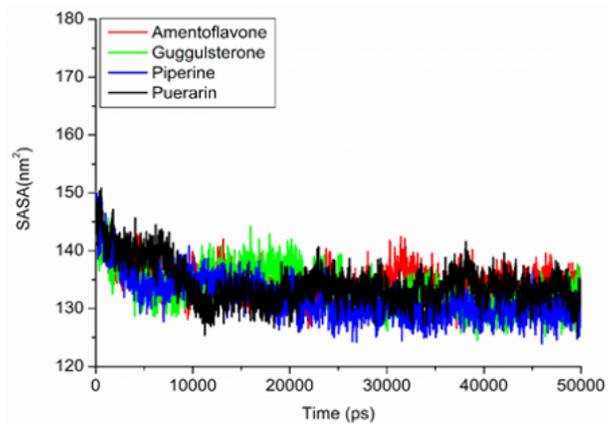
References

- Abraham, M. J., & Gready, J. E. (2011). Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *Journal of Computational Chemistry*. <https://doi.org/10.1002/jcc.21773>
- Banerjee, P., Eckert, A. O., Schrey, A. K., & Preissner, R. (2018). ProTox-II: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gky318>
- Berendsen, H. J. C., van der Spoel, D., & van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*. [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E)
- Boopathi, S., Poma, A. B., & Kolandaivel, P. (2020). Novel 2019 Coronavirus Structure, Mechanism of Action, Antiviral drug promises and rule out against its treatment. *Journal of Biomolecular Structure & Dynamics*. <https://doi.org/10.1080/07391102.2020.1758788>

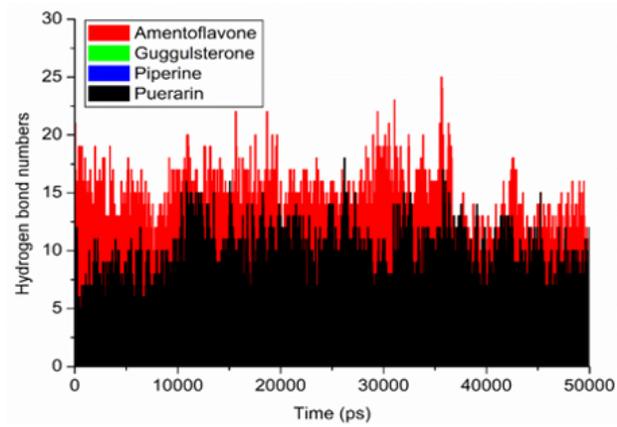
- Chu, C. M., Cheng, V. C. C., Hung, I. F. N., Wong, M. M. L., Chan, K. H., Chan, K. S., Kao, R. Y. T., Poon, L. L. M., Wong, C. L. P., Guan, Y., Peiris, J. S. M., & Yuen, K. Y. (2004). Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. *Thorax*, *59*(3), 252–256. <https://doi.org/10.1136/thorax.2003.012658>
- Daina, A., Michielin, O., & Zoete, V. (2019). SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkz382>
- Enmozhi, S. K., Raja, K., Sebastine, I., & Joseph, J. (2020). Andrographolide As a Potential Inhibitor of SARS-CoV-2 Main Protease: An In Silico Approach. *Journal of Biomolecular Structure & Dynamics*. <https://doi.org/10.1080/07391102.2020.1760136>
- Forli, W., Halliday, S., Belew, R., & Olson, A. (2012). AutoDock Version 4.2. *Citeseer*.
- Gajula, M., Kumar, A., & Ijaq, J. (2016). Protocol for Molecular Dynamics Simulations of Proteins. *BIO-PROTOCOL*. <https://doi.org/10.21769/bioprotoc.2051>
- Giménez, B. G., Santos, M. S., Ferrarini, M., & Dos Santos Fernandes, J. P. (2010). Evaluation of blockbuster drugs under the rule-of-five. *Pharmazie*. <https://doi.org/10.1691/ph.2010.9733>
- Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*. <https://doi.org/10.1002/elps.1150181505>
- Hendaus, M. A. (2020). Remdesivir in the treatment of Coronavirus Disease 2019 (COVID-19): A simplified summary. *Journal of Biomolecular Structure and Dynamics*, *0*(0), 1–10. <https://doi.org/10.1080/07391102.2020.1767691>
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry*. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., ... Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Israr, M., Mitchell, D., Alam, S., Dinello, D., Kishel, J. J., & Meyers, C. (2011). The HIV protease inhibitor lopinavir/ritonavir (Kaletra) alters the growth, differentiation and proliferation of primary gingival epithelium. *HIV Medicine*. <https://doi.org/10.1111/j.1468-1293.2010.00863.x>
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., ... Yang, H. (2020). Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *BioRxiv*. <https://doi.org/10.1101/2020.02.26.964882>
- Jo, S., Kim, S., Shin, D. H., & Kim, M. S. (2020). Inhibition of SARS-CoV 3CL protease by flavonoids. *Journal of Enzyme Inhibition and Medicinal Chemistry*. <https://doi.org/10.1080/14756366.2019.1690480>
- Johns Hopkins Hospital and Medicine. (2020). *Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)*. Johns Hopkins University.
- Khan, S. A., Zia, K., Ashraf, S., Uddin, R., & Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics*. <https://doi.org/10.1080/07391102.2020.1751298>
- Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., & Bryant, S. H. (2016). PubChem substance and compound databases. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkv951>
- Kumar, A., Choudhir, G., Shukla, S. K., Sharma, M., Tyagi, P., Bhushan, A., & Rathore, M. (2020). Identification of phytochemical inhibitors against main protease of COVID-19 using molecular modeling approaches. *Journal of Biomolecular Structure & Dynamics*, *0*(0), 1–21. <https://doi.org/10.1080/07391102.2020.1772112>
- Kumar, A., Kumar, R., Sharma, M., Kumar, U., Prasad Gajula, M. N. V., & Singh, K. P. (2018). Uttarakhand medicinal plants database (UMPDB): A platform for exploring genomic, chemical, and traditional knowledge. *Data*. <https://doi.org/10.3390/data3010007>
- Kumar, A., Kumar, S., Kumar, A., Sharma, N., Sharma, M., Singh, K. P., Rathore, M., & Gajula, M. N. V. P. (2018). Homology Modeling, Molecular Docking and Molecular Dynamics Based Functional Insights into Rice Urease Bound to Urea. *Proceedings of the National Academy of Sciences India Section B - Biological Sciences*. <https://doi.org/10.1007/s40011-017-0898-0>
- Lahlou, M. (2013). The Success of Natural Products in Drug Discovery. *Pharmacology & Pharmacy*. <https://doi.org/10.4236/pp.2013.43a003>
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2012). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. In *Advanced Drug Delivery Reviews*. <https://doi.org/10.1016/j.addr.2012.09.019>
- Lu, H. (2020). Drug treatment options for the 2019-new coronavirus (2019-nCoV). *BioScience Trends*. <https://doi.org/10.5582/bst.2020.01020>

- O'Boyle, N. M., Morley, C., & Hutchison, G. R. (2008). Pybel: A Python wrapper for the OpenBabel cheminformatics toolkit. *Chemistry Central Journal*. <https://doi.org/10.1186/1752-153X-2-5>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. <https://doi.org/10.1002/jcc.20084>
- Sarma, P., Sekhar, N., Prajapat, M., Avti, P., Kaur, H., Kumar, S., Singh, S., Kumar, H., Prakash, A., Dhibar, D. P., & Medhi, B. (2020). In-silico homology assisted identification of inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain). *Journal of Biomolecular Structure & Dynamics*. <https://doi.org/10.1080/07391102.2020.1753580>
- Schüttelkopf, A. W., & Van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*. <https://doi.org/10.1107/S0907444904011679>
- Thayil, Seema, M., & Thyagarajan, S. P. (2016). Pa-9: A flavonoid extracted from plectranthus amboinicus inhibits HIV-1 protease. *International Journal of Pharmacognosy and Phytochemical Research*.
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. In *Journal of Computational Chemistry*. <https://doi.org/10.1002/jcc.20291>
- WHO. (2020). WHO Director-General 's remarks at the media briefing on 2019-nCoV on 11 February. *WHO Director General's Statement*.
- World Health Organization (WHO). (2020). Novel Coronavirus (2019-nCoV). *WHO Bulletin*.
- Xu, Z., Peng, C., Shi, Y., Zhu, Z., Mu, K., Wang, X., & Zhu, W. (2020). Nelfinavir was predicted to be a potential inhibitor of 2019-nCoV main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. In *bioRxiv*. <https://doi.org/10.1101/2020.01.27.921627>
- Zakaryan, H., Arabyan, E., Oo, A., & Zandi, K. (2017). Flavonoids: promising natural compounds against viral infections. In *Archives of Virology*. <https://doi.org/10.1007/s00705-017-3417-y>
- Zoete, V., Daina, A., Bovigny, C., & Michielin, O. (2016). SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *Journal of Chemical Information and Modeling*. <https://doi.org/10.1021/acs.jcim.6b00174>

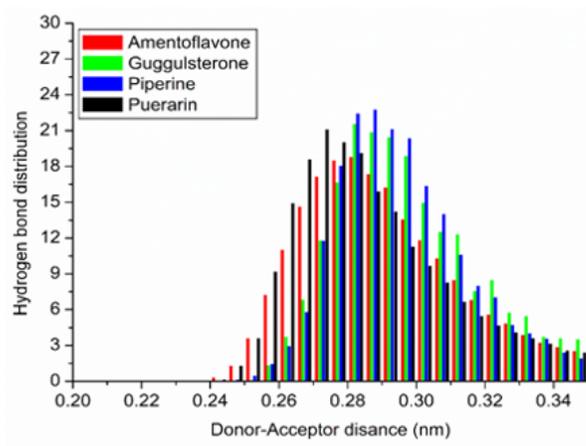
Figures



a

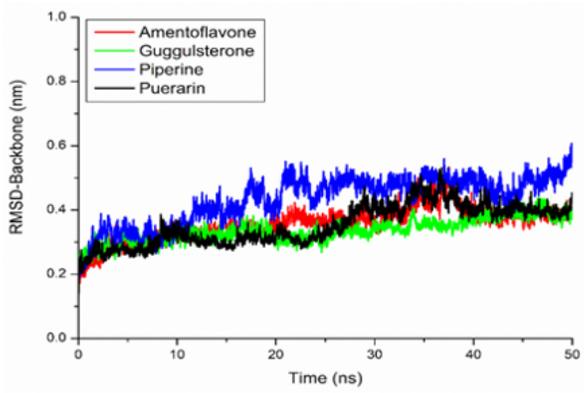


b

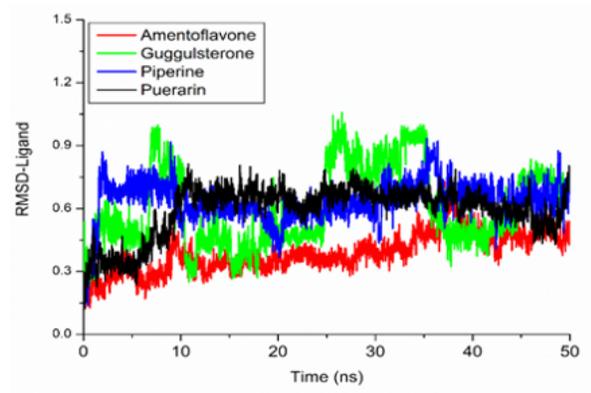


c

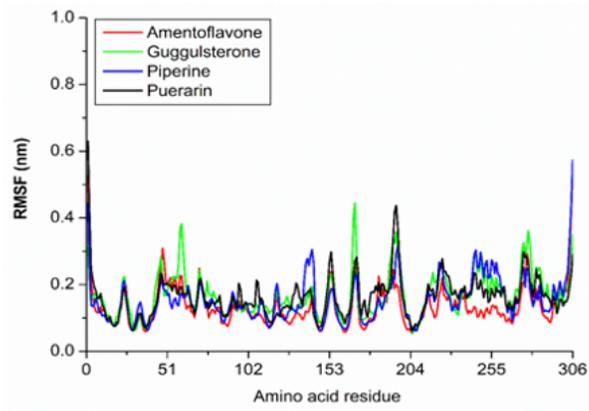
Figure 1
 (a) Solvent accessible surface area (SASA); (b) hydrogen bond numbers; (c) hydrogen bond distribution for all four complexes during MD simulations on 50ns.



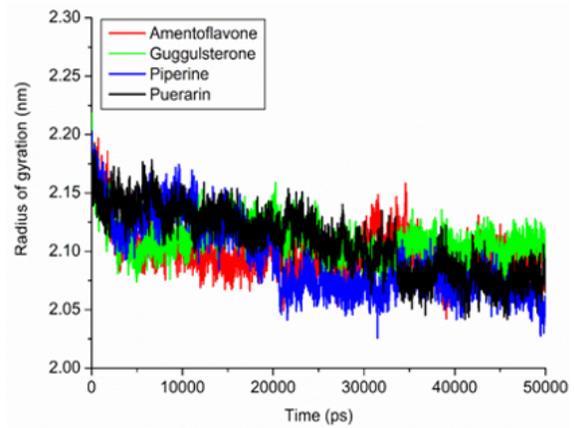
a



b



c



d

Figure 2
 (a) Root mean square deviation (RMSD) backbone; (b) RMSD ligand; (c) root mean square fluctuation (RMSF); (d) radius of gyration for all four complexes over the 50ns simulations. -8.05 kcal/mol docking energy. Interactions were visualized using maestro and pymol.

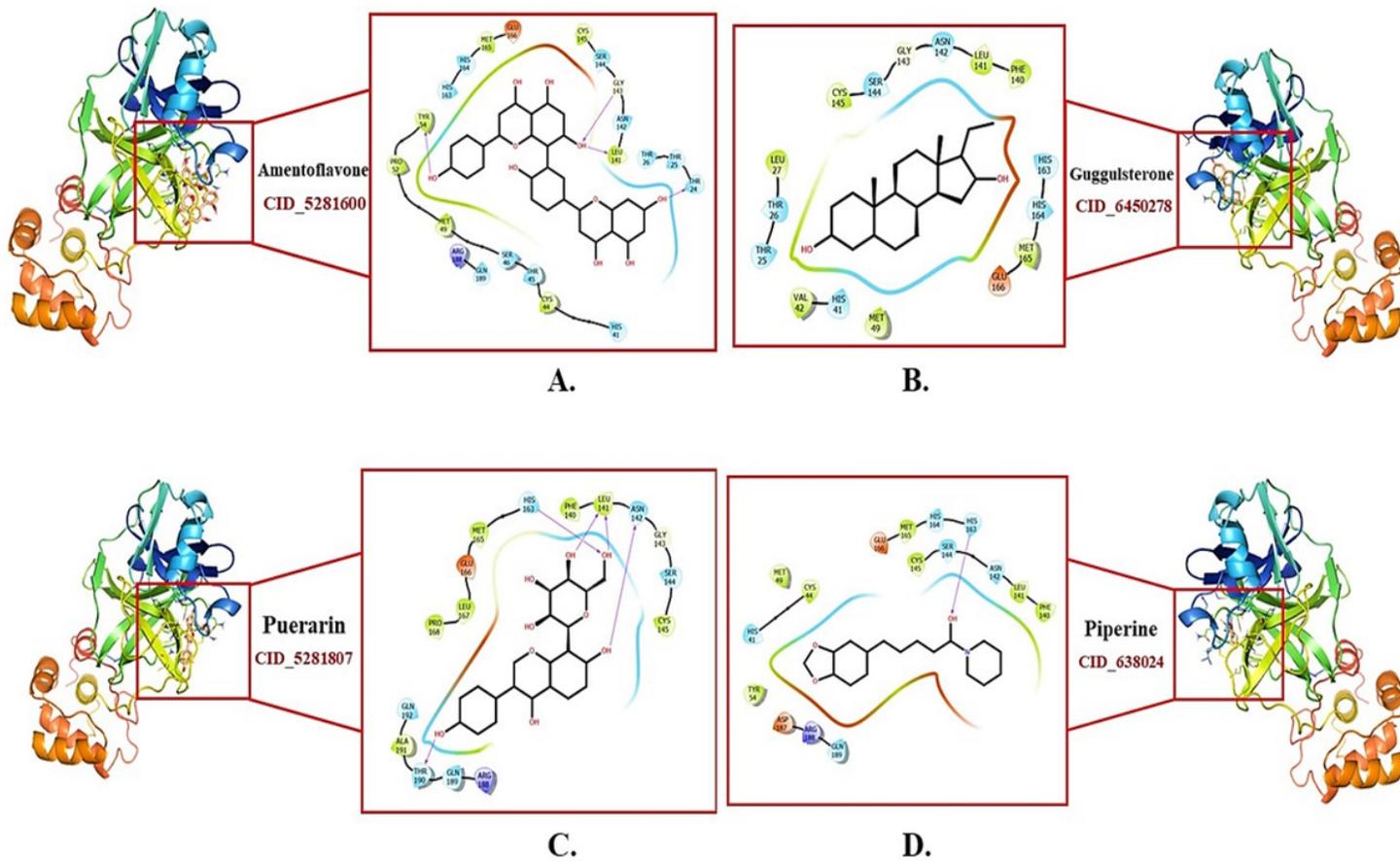


Figure 3
Schematic representation of molecular docking between Mpro and top four natural compounds; (A) interaction between Mpro and Amentoflavone with -9.96 kcal/mol docking energy; (B) interaction between Mpro and Guggulsterone with docking energy -9.67 kcal/mol; (C) interaction between Mpro and Puerarin with -8.67 kcal/mol docking energy; (D) interaction between Mpro and Piperine with -8.05 kcal/mol docking energy. Interactions were visualized using maestro and pymol.

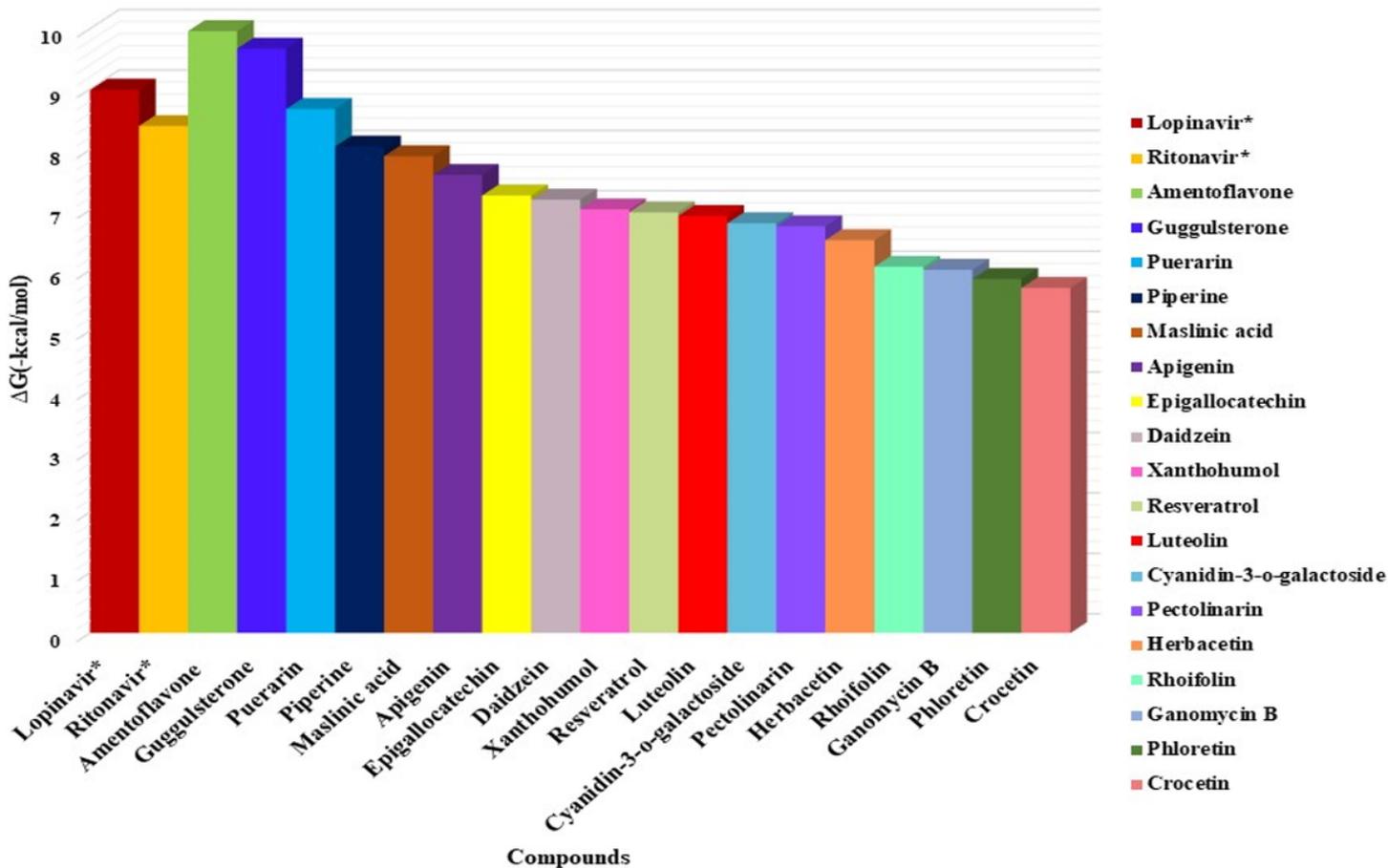


Figure 4
Histogram showing molecular docking results between COVID-19 main protease Mpro (PDB-6LU7) and selected natural compounds (the binding energy value ΔG is shown in minus kcal/mol), *Reference compounds.

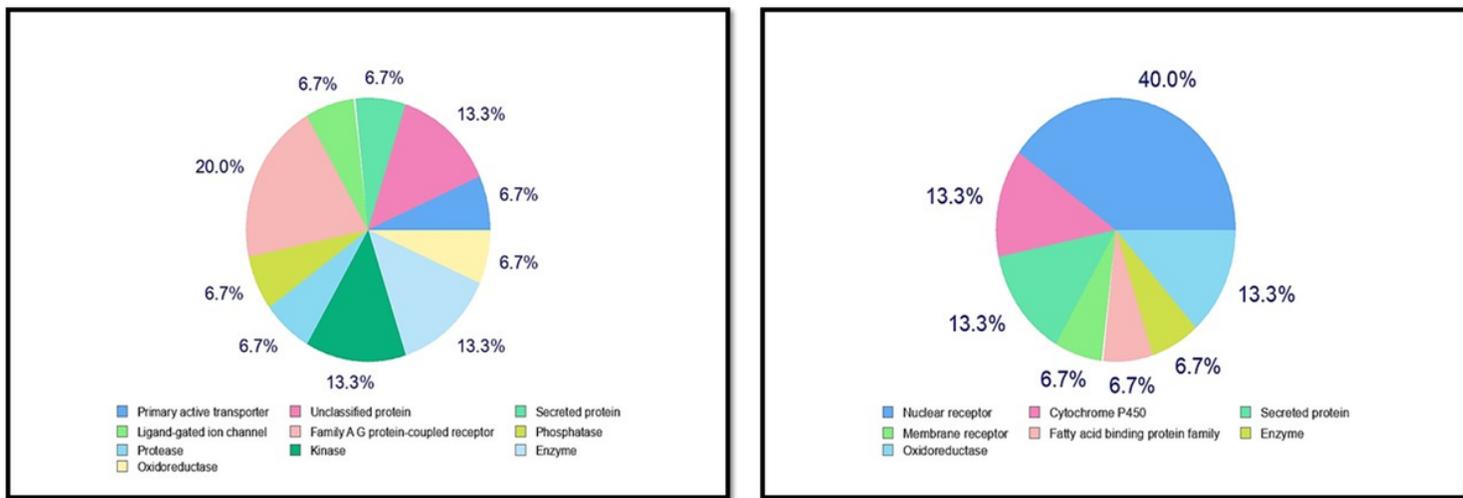


Figure 5
Top-15 of Target Predicted for Amentoflavone and Guggulsterone

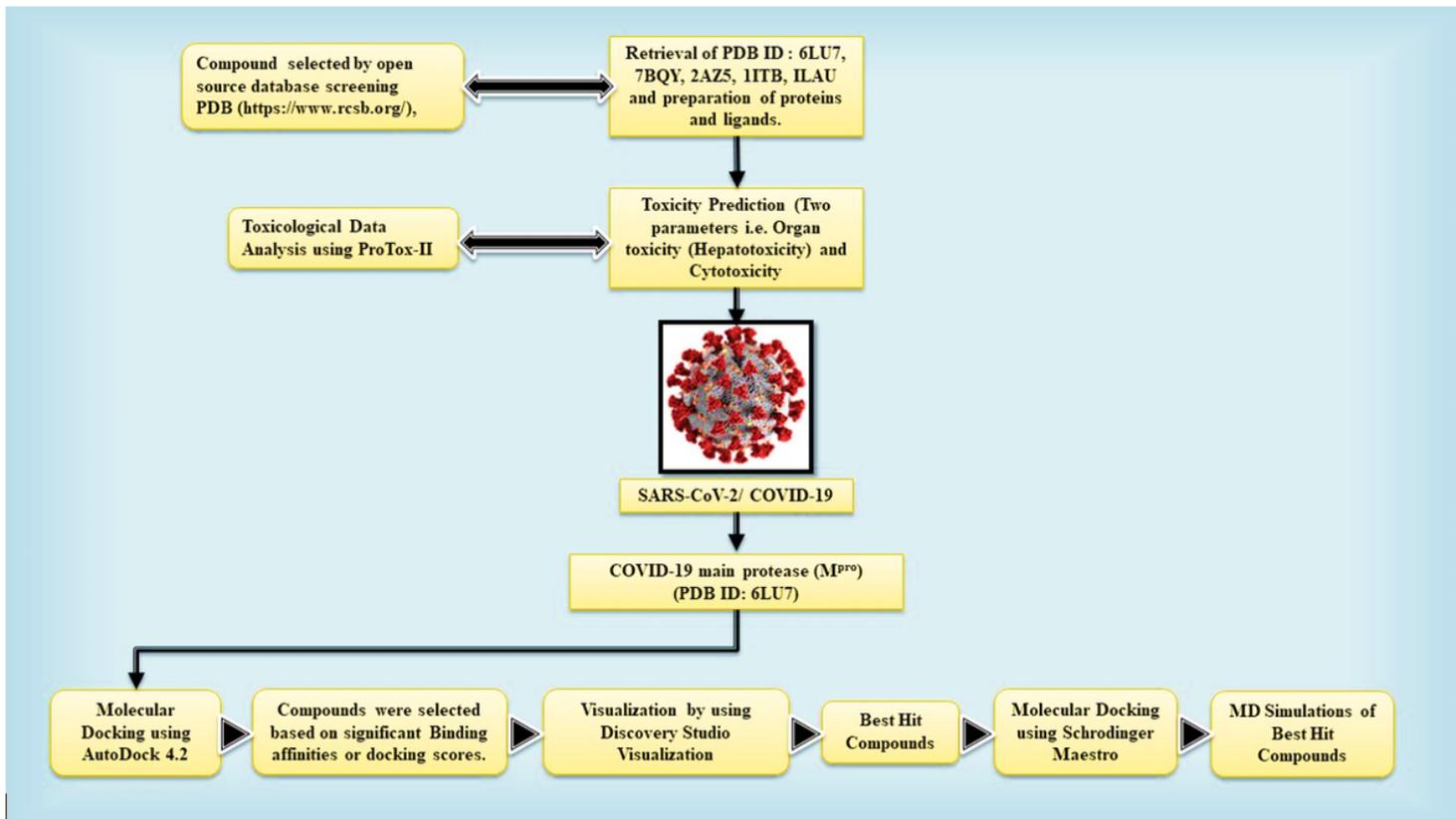


Figure 6

Flow chart of pipeline utilized in present study.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [table6.pdf](#)
- [table5.pdf](#)
- [table3.pdf](#)
- [table2.pdf](#)