Natural compounds as potential inhibitors of SARS−CoV−2 main protease: An *in−silico* study

Amaresh Mishra[1], Yamini Pathak[1], Anuj Kumar[2,3], Surabhi Kirti Mishra[4], Vishwas Tripathi[1][✉]

[1]*School of Biotechnology, Gautam Buddha University, Greater Noida−201310, India*

[2]*Bioinformatics Laboratory, Uttarakhand Council for Biotechnology, Biotech Bhawan, Pantnagar, U.S. Nagar, Uttarakhand−263145, India*

[3]*Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology, Dehradun−248007, India*

[4]*School of Biotechnology, Jawaharlal Nehru University, New Delhi, India*

ABSTRACT

Objective: To explore natural compounds as potential inhibitors against main protease (M^{pro}) of SARS-CoV-2.

Methods: In the current study, systematic molecular docking analysis was conducted using AutoDock 4.2 to determine the binding affinities and interactions between natural compounds and M^{pro}. Selected natural compounds were further validated using a combination of molecular dynamic (MD) simulations and molecular mechanic Poisson-Boltzmann surface area (MM/PBSA) free energy calculations.

Results: Out of twenty natural compounds, four natural metabolites namely, amentoflavone, guggulsterone, puerarin, and piperine were found to have strong interaction with M^{pro} of SARS-CoV-2 based on docking analysis. During MD simulations, all four natural compounds bound to M^{pro} at 50 ns and MM/G/P/BSA free energy calculations showed that all four shortlisted ligands had stable and favorable energies with strong binding to M^{pro} protein.

Conclusions: Guggulsterone is a potential inhibitor of COVID-19 main protease M^{pro}. Further *in vitro* and pre-clinical studies are needed.

KEYWORDS: SARS-CoV-2; COVID-19; Amentoflavone; Guggulsterone; Puerarin; Piperine

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[1]. SARS-CoV-2 is a novel strain of coronavirus and has expanded all over the world in a very short period[2]. The outbreak was declared as a Public Health Emergency of International Concern by WHO on 30 January 2020[3]. To date, there is no specific treatment for this ongoing COVID-19 pandemic. Some preliminary study results investigated the effect of a potential combination of lopinavir and ritonavir on COVID-19 infected patients, which were earlier used in human immunodeficiency virus and SARS CoV or Middle East respiratory syndrome (MERS) coronavirus patients[4,5]. These drugs can specifically target the virus replication cycle. Effective antiviral therapies are still urgently required due to subsequent infection[6]. Because of enormous structural and chemical diversity, availability of more chiral centers, and relative biosafety, natural compounds are considered as an excellent source of drugs for several diseases including viral infections. Around 45% of today's best selling drugs have either originated from natural products or their derivatives[7]. Natural compounds possess antiviral property and could become

a valuable resource. Jin *et al*. have revealed the crystal structure of SARS-CoV-2/COVID-19 main protease (M[pro]/3CL protease, PDB ID-6LU7)[8] which has been structured and repositioned in the Protein Data Bank (PDB) and is publicly accessible. M[pro] of SARS-CoV-2 is reported to play a vital role in virus replication and transcription, which suggests that it could be a promising target for inhibition of the SARS-CoV-2 cycle[9].

In the current study, we have selected several natural compounds based on an extensive literature search of natural compounds with anti-viral property. These natural compounds are presented in Supplementary Table 1. We screened and explored the potential of selected natural compounds to inhibit the M[pro] of COVID-19 using molecular docking, followed by molecular dynamic (MD) simulations, molecular mechanic Poisson-Boltzmann surface area (MM/PBSA) free energy calculations, absorption, distribution, metabolism, and excretion (ADME) screening, drug-likeness, target-specific binding, and toxicity analysis.

2. Materials and methods

A flow chart of pipeline of this study is summarized in Supplementary Figure 1.

2.1. Literature survey and ligands selection

An extensive literature survey was conducted to select natural compounds with antiviral properties from different medicinal plants using PubMed and Google scholar platforms. Based on the literature survey, a total of twenty natural compounds were selected, and their chemical structures were extracted from PubChem[10] repository in SDF format. To prepare the ligands to perform molecular docking, hydrogen atoms were added followed by PDB structure generation by Open Babel program. Further, all the molecules were subjected to the energy minimization and optimization using the universal force field at 200 descent steepest algorithm of Open Babel available in PyRx (https://pyrx.sourceforge.io/) and converted in .pdbqt format.

2.2. Preparation of protease

The 3D coordinates of the main M[pro] of SARS-CoV-2 were obtained from the RCSBPDB repository with PDB ID 6LU7[8]. To prepare the macromolecule for docking, water, and other nonspecific molecules were removed by using UCSF Chimera[11]. Polar hydrogen atoms were added to the 3D structure model of M[pro] to allow intramolecular interactions through hydrogen bonds. The structure optimization and energy minimization were performed by using the SPDB viewer[12]. While clean geometry module embedded in the 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 AutoDock4.2[13]. The box dimensions were kept as 70 Å×70 Å×70 Å with a total of 50 genetic algorithm run. Other docking parameters were set as default. During molecular docking, the amino acid residues including Thr25, Thr26, Gly143, Ser144, His163, His164, and Glu166 were utilized as the binding pocket sites. The protein-ligand interactions were further calculated by the Maestro and Discovery studio programs.

2.4. ADME screening

An *in−silico* tool for analysis of ADME was used to screen the above-mentioned compounds which could be bioactive *via* oral administration. Drug-like properties were calculated based on Lipinski's rule of five using SwissADME prediction (http://www.swissadme.ch/)[14]. Canonical SMILES from PubChem was used to identify ADME properties by SwissADME. Various parameters of compounds were analyzed like lipophilicity, molecular weight, hydrogen-bond donors, hydrogen-bond acceptors, clog *P*-value, Ghosh violations, Lipinski violations, *etc*. Ligands/natural compounds were 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.

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, rationalize possible side-effects, and predict off-targets. In this direction, SwissTargetPredictiontool (https://www.swisstargetprediction.ch) was used[15]. Canonical SMILE for amentoflavone and guggulsterone was entered and was analyzed.

2.6. 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 at 50 ns. The method for the MD simulations of natural compounds-M[pro] complexes was the same as recent studies[16,17]. All simulations of representative natural compounds-M[pro] complexes were conducted with the utilization of GROMOS96 43a1 force field available in

GROMACS 5.1.4 suite[18]. Topology files for ligand molecules were created by using the PRODRG server[19]. The prepared protein complexes were solvated in a cubic box of edge length 10 nm along with simple point charge water molecules. 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 1 000 kJ/mol/nm. Long-range interaction electrostatics[20] was handled by using particle mesh Ewald. 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 the first stage, while in the second stage the restraint weight from the protein and protein-ligand complexes was gradually declined in the NPT (constant number of atoms N, pressure P, and temperature T) ensemble. LINear Constraint Solver constraints were applied to all bonds involving hydrogen atoms[21]. 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[22]. 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.

The conformational stability and flexibility of the complexes were analyzed by 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.

2.7. Free energy calculations by MM/PBSA

The calculations of the binding energy of the M^{pro}-ligand complexes were calculated by using MM/PBSA method. While 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[23]. Only data of the last 10 ns were utilized for the MM-PBSA analysis considering the convergence issue associated with the calculations. In the present study, entropy calculations were not performed 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 the following equations:

$$\Delta G_{MMPBSA} = (G_{complex} - G_{protein} - G_{ligand})_{complex} \qquad (1)$$
$$G_x = E_{MM} - T(S_{MM}) + \Delta G_{solv} \qquad (2)$$
$$E_{MM} = E_{bonded} + E_{coul} + E_{LJ} \qquad (3)$$
$$\Delta G_{solv} = G_{polar} + G_{nonpolar} \qquad (4)$$

Where ΔG_{MMPBSA}: Free energy of molecular mechanics Poisson-Boltzmann surface area; $G_{complex}$: the total free energy of the protein-ligand complex; $G_{protein}$ and G_{ligand} are total free energies of the isolated protein and ligand in solvent; G_x is the free energy of system x that being the ligand, the protein, or the complex; T is the temperature; S_{MM} is the entropy; and angle brackets represent an ensemble average. E_{MM} is the potential energy in vacuum as defined by the molecular mechanics (MM) model, which is composed of the bonded energy terms (E_{bonded}) and nonbonded Coulombic (E_{coul}) and Lennard-Jones (E_{LJ}) terms; ΔG_{solv} = Free energy of solvation is the combination of G_{polar} and $G_{nonpolar}$.

2.8. Toxicity analysis

The prediction of compound toxicities is an important parameter for drug design process. In order to check toxicity of these selected natural compounds, ProTox-Ⅱ web server was used[24]. ProTox-Ⅱ is a kind of virtual lab that 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. Comparison with similar FDA approved drug compound by SwissSimilarity

The compounds which had suitable parameters including ADME & toxicity and the best binding energy were compared with FDA-approved drugs using SwissSimilarity tool (http://www.swisssimilarity.ch)[25].

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

The structure and amino acids were found in the active site pockets of 6LU7.

3.2. ADME

The drug scanning results showed that most of the tested compounds were accepted based on Lipinski's rule of five. These compounds were selected for docking to find their binding affinity with COVID19 main protease M^{pro}. The list of compounds with suitable ADME properties is given in Supplementary Table 2.

3.3. Target prediction

The target prediction analysis was performed for two best compounds, amentoflavone and guggulsterone on the web page. For amentoflavone, it predicted 20% of family AG protein-coupled receptor, 13.3% of 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, and 6.7% of ligand-gated ion channel. While for guggulsterone, it predicted 40% of nuclear receptors, 13.3% of cytochrome P_{450}, 13.3% of secreted protein, 13.3% of oxidoreductase, 6.7% of membrane receptors, 6.7% of fatty acid-binding protein family, and 6.7% of enzymes. The details including target, common name, UniProtID, ChEMBLID, target class, probability, and known actives in 2D/3D are given in Supplementary Table 3 & 4. The possible sites of the target which the compound may bind to were mostly the targets predicted by the software. The probability scores for amentoflavone and guggulsterone were from 1.000 to 0.087 & 1.000 to 0.102, respectively.

3.4. Molecular docking

A total of 18 selected natural compounds were docked with COVID-19 main protease M^{pro} with ritonavir and lopinavir as standard controls. The results indicated a good binding affinity of ritonavir and lopinavir to the COVID-19 main protease M^{pro}. The docking results are given in Table 1 and docking analysis visualization is shown in Figure 1. Molecular docking results between COVID-19 main protease M^{pro} (PDB-6LU7) and selected natural compounds and schematic representation of molecular docking between M^{pro} and top four natural compounds are given in Supplementary Figure 2 & 3.
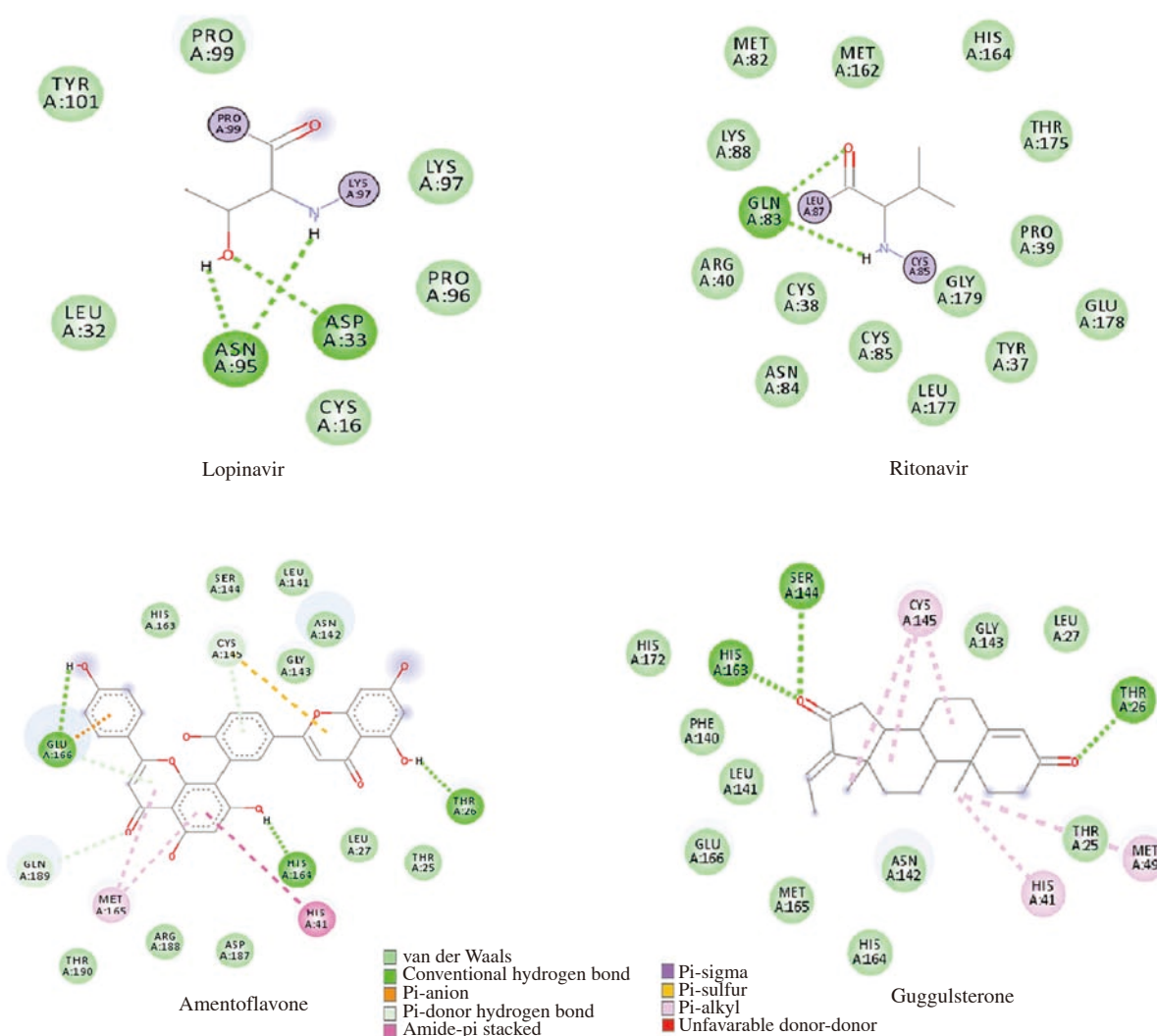

Figure 1. Docking analysis visualization of SARS-CoV-2 main protease M^{pro} (PDB-6LU7) binding with lopinavir, ritonavir, amentoflavone, and guggulsterone. The 3D structures of protein-ligand interactions were visualized by discovery studio programs. The binding residues and their chains were identified from the protein-ligand complex.

3.5. MD simulations

The time-dependent RMSD was calculated from the initial stage of the simulation on 50 ns. The RMSD of the backbone of these 4 complexes was 0.2 to 0.5 nm (Figure 2A), which stabilized at the 35 ns whereas the RMSD of ligands ranged from 0.3 to 1.0 nm (Figure 2B). During MD simulations, RMSF defines the residual flexibility from the average position. The RMSF of the protein ranged from 0.1 to 0.5 nm of all systems (Figure 2C). Some amino acids showed a high-intensity peak, which may represent a loop region. In MD simulations, Rg is a well-known method to determine the compactness of protein. The compactness of a protein is induced by the movement of a ligand. During simulation analysis on 50 ns, the lower flexibility of the Rg is associated with the structural stability of the protein. The Rg values of all phytochemical complexes were 2.20 to 2.05 nm (Figure 2D). The Rg values of all four phytochemical complexes supported their consensus architecture as well as size. The predicted SASA values were found to be associated with the exposures of hydrophobic residues during simulation analysis. SASA plays a principal role in the van der interaction. The SASA values of all systems were 125-150 nm^2 (Figure 3A).

In a complex protein and ligand, hydrogen bonding plays a critical role in determining the strength of interaction. During the simulation time, several hydrogen bonds formed between the donor and the acceptor group (Figure 3C). Two hydrogen bonds were consistently formed during the time of simulation (Figure 3B). Overall observations indicated that all four complexes were stable during simulation.

3.6. Binding free energy calculation

The free binding energy results showed that as compared with other molecules, amentoflavone had maximum binding energy followed by guggulsterone (Table 2). In each molecule, the polar salvation and SASA energy showed moderate effects on binding energy component.

3.7. Toxicity analysis

In–silico toxicities of selected natural compounds were predicted by using ProTox-Ⅱ. As shown in Table 3, none of the selected natural compounds showed potential hepatotoxicity or cytotoxicity except pectolinarin.

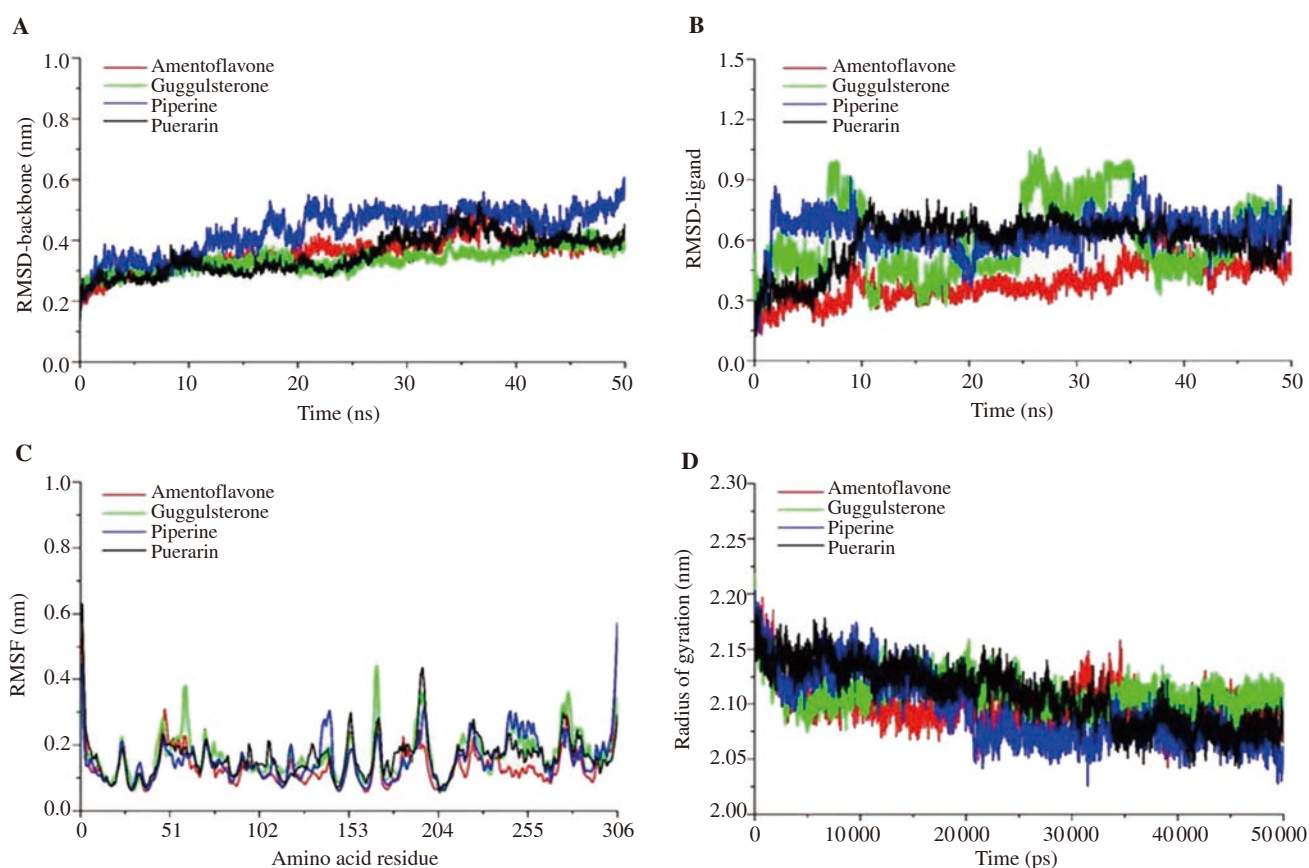We further checked the similarity of two top natural compounds

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 50 ns simulations.

Table 1. Molecular docking analysis results of selected natural compounds against SARS-CoV-2 main protease M^{pro} (PDB-6LU7).

S. No.	Ligand	PubChem ID	Hydrogen bond interactions	Binding energy (ΔG) (kcal/mol)
1	Lopinavir*	CID_92727	ASN95, ASP33	-8.99
2	Ritonavir*	CID_392622	GLN83	-8.39
3	Amentoflavone	CID_5281600	HIS163, GLU166, LEU141, THR190, GLN192	-9.96
4	Guggulsterone	CID_6450278	THR26, SER144, HIS163	-9.67
5	Puerarin	CID_5281807	THR190, ASN142, LEU141, SERA144, CYS145	-8.67
6	Piperine	CID_638024	GLY143	-8.05
7	Maslinic acid	CID_73659	THR24, THR26, GLU166, HIS 164	-7.89
8	Apigenin	CID_5280443	HIS164, ASP 187, THR190	-7.58
9	Epigallocatechin	CID_72277	HIS163, HIS164, LEU141, ARG188, GLU166	-7.24
10	Daidzein	CID_5281708	CYS145, GLU166, THR190	-7.17
11	Xanthohumol	CID_639665	HIS164, CYS145, THR26	-7.01
12	Resveratrol	CID_445154	MET49, LEU141, CYS145	-6.96
13	Luteolin	CID_5280445	HIS163, HIS164, LEU141, THR26, GLY143	-6.90
14	Cyanidin-3-*O*-galactoside	CID_441699	HIS163, HIS164, ASN142, PHE140, CYS145, THR26, TYR54	-6.78
15	Pectolinarin	CID_168849	GLN189, THR190, HIS164	-6.73
16	Herbacetin	CID_5280544	THR190	-6.50
17	Rhoifolin	CID_5282150	THR190, PHE140, HIS163, LEU141, SER144, GLY143, THR26	-6.06
18	Ganomycin B	CID_10246918	GLU166, GLY143	-6.01
19	Phloretin	CID_4788	HIS164, CYS145, PHE140, GLU166, ASP187	-5.86
20	Crocetin	CID_5281232	THR26, THR190, GLN192	-5.71

*standard ligands.
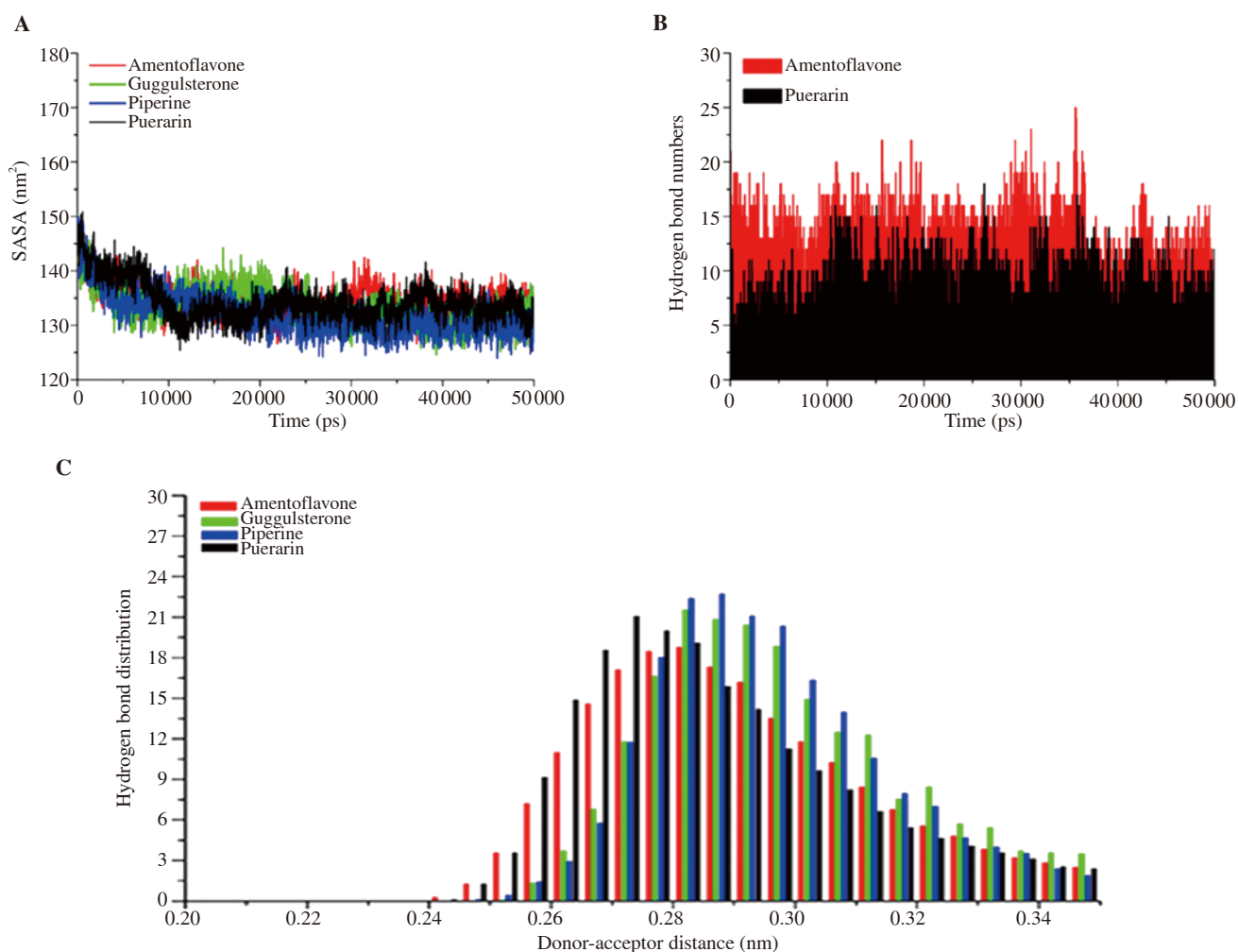
Figure 3. (A) Solvent accessible surface area (SASA); (B) hydrogen bond numbers; (C) hydrogen bond distribution for all four complexes during MD simulations for 50 ns.

Table 2. Binding free energy calculation of four stable complexes during simulation (kJ/mol).

Name of molecules	Van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
Amentoflavone	350.082±17.177	104.696±19.160	239.906±21.251	25.562±1.010	240.434±17.602
Guggulsterone	140.068±15.668	-7.408±9.330	37.621±11.572	11.853±1.070	-121.708±12.423
Piperine	173.545±11.759	-9.373±4.129	50.573±7.610	13.943±1.293	-146.287±11.205
Puerarin	180.787±20.912	-82.405±16.508	148.200±17.298	16.639±1.417	-131.631±20.483

Table 3. Toxicity predictions for selected natural compounds.

S.No.	Compounds	Toxicity class	LD_{50} (mg/kg)	Hepatotoxicity	Cytotoxicity
1	Amentoflavone	5	3919	Inactive	Inactive
2	Guggulsterone	5	2450	Inactive	Inactive
3	Puerarin	4	832	Inactive	Inactive
4	Piperine	4	330	Inactive	Inactive
5	Maslinic acid	4	2000	Inactive	Inactive
6	Apigenin	5	2500	Inactive	Inactive
7	Epigallocatechin	6	10000	Inactive	Inactive
8	Daidzein	5	2430	Inactive	Inactive
9	Xanthohumol	5	3800	Inactive	Inactive
10	Resveratrol	4	1560	Inactive	Inactive
11	Luteolin	5	3919	Inactive	Inactive
12	Cyanidin-3-*O*-galactoside	5	5000	Inactive	Inactive
13	Pectolinarin	5	5000	Inactive	Active
14	Herbacetin	5	3919	Inactive	Inactive
15	Rhoifolin	5	5000	Inactive	Inactive
16	Ganomycin B	4	370	Inactive	Inactive
17	Phloretin	4	500	Inactive	Inactive
18	Crocetin	5	4300	Inactive	Inactive

amentoflavone and guggulsterone with the FDA approved drugs using SwissSimilarity check. For amentoflavone, we did not find any reported similar FDA approved drug. Whereas for guggulsterone, we found 117 FDA proved drugs. The details including drug ID, drug name, similarity score, and molecule structure are given in Supplementary Figure 4 & 5. The structural similarity of other FDA approved drugs with guggulsterone was predicted using SwissSimilarity and the probability score obtained was 0.995 to 0.009.

4. Discussion

COVID-19 pandemic is declared as a public health emergency and is lack of specific treatment[3]. There is a dire need for effective drugs against the novel coronavirus COVID-19. Since the virus is new, information about its molecular mechanism is very limited. The only reference is from the old SARS virus (SARS-CoV-1) that emerged in 2003. Fortunately, in a recent study, Jin *et al*. have revealed the crystal structure of SARS-CoV-2/COVID-19 main protease (M^pro/3CL protease, PDBID-6LU7)[8]. This protease is considered as an important 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 like SARS and MERS[26]. 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. Target prediction suggests that the small compound may have high target attraction towards the specific binding site. To further investigate the molecular docking results, the top four natural compound complexes namely, amentoflavone, guggulsterone, puerarin, and piperine were subjected to 50 ns MD simulations. RMSD of the Cα atoms is related to the stability of the complexes. RMSD value may indicate that the binding of ligands does not lead to the conformation perturbation during the simulation. RMSD of the protein backbone in all systems was small and comparable. The presence of low-intensity peak suggested that binding of the phytomolecules does not affect the stability of the structural region of the enzyme. SASA conformations showed that the binding of ligand molecules does not affect the overall folding of the protein. Lopinavir and ritonavir are well-known protease inhibitors of human immunodeficiency virus[27]. Both drugs were also recommended as repurposed drugs in the treatment of SARS and MERS[5]. Therefore, in this study, we took these drugs as standard reference drugs for comparison. In our *in–silico* prediction experiment, none of the selected compounds showed hepatotoxicity and cytotoxicity except pectolinarin. The compounds with potential to inhibit the viral protease based on the binding energy include 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 compounds, amentoflavone and guggulsterone were the top two with the best binding energy and most satisfactory parameters. Guggulsterone from a Indian traditional ayurvedic medical plant

Commiphora mukul was found to be the most suitable based on comprehensive pharmacokinetic parameters, drug-likeness, and docking analysis. Therefore, we propose guggulsterone as a potential inhibitor of COVID-19 main protease M^{pro}. Further *in−vitro* and *in−vivo* studies are still needed to validate these bioinformatics based findings.

Conflict of interest statement

We declare that there is no conflict of interest.

Authors' contributions

VT conceived the idea, wrote the discussion part,and supervised the entire study. AM and YP contributed to the research tool and wrote the manuscript. AK & SKM did the critical analysis.

References

[1] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395**(10223): 497-506. Doi: 10.1016/S0140-6736(20)30183-5.

[2] Zheng J. SARS-CoV-2: An emerging coronavirus that causes a global threat. *Int J Biol Sci* 2020; **16**(10): 1678. Doi: 10.7150/ijbs.45053.

[3] World Health Organization. WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. [Online] Available from: https://www. who. int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020. [Accessed on 2020 Apr 10].

[4] Lu H. Drug treatment options for the 2019-new coronavirus (2019-nCoV). *Biosci Trends* 2020;**14**(1): 69-71. Doi: 10.5582/bst.2020.01020.

[5] Chu CM, Cheng VC, Hung IF, Wong MM, Chan KH, Chan KS, et al. Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. *Thorax* 2004; **59**(3): 252-256.

[6] Gowtham HG, Monu DO, Ajay Y, Gourav C, Vasantharaja R, Bhani K, et al. Exploring structurally diverse plant secondary metabolites as a potential source of drug targeting different molecular mechanisms of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) pathogenesis: An *in silico* approach. *Sci Rep* 2020; **2**: 1-38.

[7] Nwonu C, Ilesanmi O, Agbedahunsi J, Nwonu P. Natural products as veritable source of novel drugs and medicines: A review. *Int J Herb Med* 2019; **7**(1): 50-54.

[8] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M^{pro} from COVID-19 virus and discovery of its inhibitors. *BioRxiv* 2020. Doi: 10.1101/2020.02.26.964882.

[9] Boopathi S, Poma AB, Kolandaivel P. Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *J Biomol Struct Dyn* 2020; 1-10. Doi: 10.1080/07391102.2020.1758788.

[10] Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. *Nucleic Acids Res* 2016; **44**(D1): D1202-D1213. Doi: 10.1093/nar/gkv951.

[11] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* 2004; **25**(13): 1605-1612. Doi: 10.1002/jcc.20084.

[12] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis* 1997; **18**(15): 2714-2723. Doi: 10.1002/elps.1150181505.

[13] Huey R, Forli S, Hart WE, Halliday S, Belew R, Olson AJ, et al. AutoDock Version 4.2. Doi: 10.1016/j.addr.2012.09.019.

[14] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 1997; **23**(1-3): 3-25.

[15] Daina A, Michielin O, Zoete V. SwissTargetPrediction: Updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res* 2019; **47**(W1): W357-W364. Doi: 10.1093/nar/gkz382.

[16] Gajula MN, Kumar A, Ijaq J. Protocol for molecular dynamics simulations of proteins. *Bio−protocol* 2016; **6**(23): e2051. Doi: 10.21769/bioprotoc.2051.

[17] Khan SA, Zia K, Ashraf S, Uddin R, Ul-Haq Z. Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 *via* integrated computational approach. *J Biomol Struct Dyn* 2020: 1-10. Doi: 10.1080/07391102.2020.1751298.

[18] Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: Fast, flexible, and free. *J Comput Chem* 2005; **26**(16): 1701-1718. Doi: 10.1002/jcc.20291.

[19] Schüttelkopf AW, Van Aalten DM. PRODRG: A tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr D Biol Crystallogr* 2004; **60**(8): 1355-1363. Doi: 10.1107/S0907444904011679.

[20] Abraham MJ, Gready JE. Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *J Comput Chem* 2011; **32**(9): 2031-2040. Doi: 10.1002/jcc.21773.

[21] Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* 1997; **18**(12): 1463-1472.

[22] Berendsen HJ, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput Phys Comm* 1995; **91**(1-3): 43-56. Doi: 10.1016/0010-4655(95)00042-E.

[23] Kumari R, Kumar R, Open Source Drug Discovery Consortium, Lynn A. g_mmpbsa— A GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model* 2014; **54**(7): 1951-1962. Doi: 10.1021/ci500020m.

[24] Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-][: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res* 2018; **46**(W1): W257-W263. Doi: 10.1093/nar/gky318.

[25]Zoete V, Daina A, Bovigny C, Michielin O. SwissSimilarity: A web tool for low to ultra high throughput ligand-based virtual screening. *J Chem Inf Model* 2016; **56**(8):1399-1404. Doi: 10.1021/acs.jcim.6b00174.

[26]Jo S, Kim S, Shin DH, Kim MS. Inhibition of SARS-CoV 3CL protease by flavonoids. *J Enzyme Inhib Med Chem* 2020; **35**(1): 145-151. Doi: 10.1080/14756366.2019.1690480.

[27]Israr M, Mitchell D, Alam S, Dinello D, Kishel JJ, Meyers C. The HIV protease inhibitor lopinavir/ritonavir (Kaletra) alters the growth, differentiation and proliferation of primary gingival epithelium. *HIV Med* 2011; **12**(3): 145-156. Doi: 10.1111/j.1468-1293.2010.00863.x.

[28]Li F, Song X, Su G, Wang Y, Wang Z, Jia J, et al. Amentoflavone inhibits HSV-1 and ACV-resistant strain infection by suppressing viral early infection. *Viruses* 2019; **11**(5): 466. Doi: 10.3390/v11050466.

[29]Gill NS, Kaur N, Arora R. An overview on: *Murraya paniculata* Linn. *Int J Inst Pharm Life Sci* 2014; **4**: 2249-6807.

[30]Wu M, Zhang Q, Yi D, Wu T, Chen H, Guo S, et al. Quantitative proteomic analysis reveals antiviral and anti-inflammatory effects of puerarin in piglets infected with porcine epidemic diarrhea virus. *Front Immunol* 2020; **11**: 169. Doi: 10.3389/fimmu.2020.00169.

[31]Nag A, Chowdhury RR. Piperine, an alkaloid of black pepper seeds can effectively inhibit the antiviral enzymes of Dengue and Ebola viruses, an *in silico* molecular docking study. *Virus Dis* 2020; **31**(3): 308-315. Doi: 10.1007/s13337-020-00619-6.

[32]Mokhtari K, Rufino-Palomares EE, Pérez-Jiménez A, Reyes-Zurita FJ, Figuera C, García-Salguero L, et al. Maslinic acid, a triterpene from olive, affects the antioxidant and mitochondrial status of B16F10 melanoma cells grown under stressful conditions. *Evid Based Complementary Altern Med* 2015; **2015**: 272457. Doi: 10.1155/2015/272457.

[33]Qian S, Fan W, Qian P, Zhang D, Wei Y, Chen H, et al. Apigenin restricts FMDV infection and inhibits viral IRES driven translational activity. *Viruses* 2015; **7**(4): 1613-1626. Doi: 10.3390/v7041613.

[34]Kaihatsu K, Yamabe M, Ebara Y. Antiviral mechanism of action of epigallocatechin-3-*O*-gallate and its fatty acid esters. *Molecules* 2018; **23**(10): 2475. Doi: 10.3390/molecules23102475.

[35]Chung ST, Huang YT, Hsiung HY, Huang WH, Yao CW, Lee AR. Novel daidzein analogs and their *in vitro* anti-influenza activities. *Chem Biodiver* 2015; **12**(4): 685-696. Doi: 10.1002/cbdv.201400337.

[36]Buckwold VE, Wilson RJ, Nalca A, Beer BB, Voss TG, Turpin JA, et al. Antiviral activity of hop constituents against a series of DNA and RNA viruses. *Antivir Res* 2004; **61**(1): 57-62. Doi: 10.1016/S0166-3542(03)00155-4.

[37]Abba Y, Hassim H, Hamzah H, Noordin MM. Antiviral activity of resveratrol against human and animal viruses. *Adv Virol* 2015; **2015**: 184241. Doi: 10.1155/2015/184241.

[38]Zakaryan H, Arabyan E, Oo A, Zandi K. Flavonoids: Promising natural compounds against viral infections. *Arch Virol* 2017; **162**(9): 2539-2551. Doi: 10.1007/s00705-017-3417-y.

[39]Pour PM, Fakhri S, Asgary S, Farzaei MH, Echeverria J. The signaling pathways, and therapeutic targets of antiviral agents: Focusing on the antiviral approaches and clinical perspectives of anthocyanins in the management of viral diseases. *Front Pharmacol* 2019; **10**: 1207. Doi: 10.3389/fphar.2019.01207.

[40]Martínez-Vázquez M, Apan TO, Lastra AL, Bye R. A comparative study of the analgesic and anti-inflammatory activities of pectolinarin isolated from *Cirsium subcoriaceum* and linarin isolated from *Buddleia cordata*. *Plant Med* 1998; **64**(02): 134-137. Doi: 10.1055/s-2006-957390.

[41]Struijs K, Vincken JP, Doeswijk TG, Voragen AG, Gruppen H. The chain length of lignan macromolecule from flaxseed hulls is determined by the incorporation of coumaric acid glucosides and ferulic acid glucosides. *Phytochemistry* 2009; **70**(2): 262-269. Doi: 10.1016/j.phytochem.2008.12.015.

[42]Vázquez LH, Palazon J, Navarro-Ocaña A. The pentacyclic triterpenes α, β-amyrins: A review of sources and biological activities. In: Rao V (ed.) *Phytochemicals–a global perspective of their role in nutrition and health*. IntechOpen; 2012, p. 487-502.

[43]Mothana RA, Jansen R, Jülich WD, Lindequist U. Ganomycins A and B, new antimicrobial farnesyl hydroquinones from the basidiomycete *Ganoderma pfeifferi*. *J Nat Prod* 2000; **63**(3): 416-418. Doi: 10.1021/np990381y.

[44]Behzad S, Sureda A, Barreca D, Nabavi SF, Rastrelli L, Nabavi SM. Health effects of phloretin: From chemistry to medicine. *Phytochem Rev* 2017; **16**(3): 527-533. Doi: 10.1007/s11101-017-9500-x.

[45]Soleymani S, Zabihollahi R, Shahbazi S, Bolhassani A. Antiviral effects of saffron and its major ingredients. *Curr Drug Deliv* 2018; **15**(5): 698-704. Doi: 10.2174/1567201814666171129210654.

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

Figure S1. Flow chart of pipeline utilized in the present study

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

Figure S2. Histogram showing molecular docking results between COVID-19 main protease M^{pro} (PDB-6LU7) and selected natural compounds (the binding energy value ΔG is shown in minus kcal/mol), *Reference compounds.

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

Figure S3. Schematic representation of molecular docking between M^{pro} and top four natural compounds; **(A)** interaction between M^{pro} and Amentoflavone with -9.96 kcal/mol docking energy; **(B)** interaction between M^{pro} and Guggulsterone with docking energy -9.67 kcal/mol; **(C)** interaction between M^{pro} and Puerarin with -8.67 kcal/mol docking energy; (**D**) interaction between M^{pro} and Piperine with -8.05 kcal/mol docking energy. Interactions were visualized using Interactions were visualized using maestro and pymol.

Results: No similar FDA approved drugs was found for Amentoflavone.

Supplementary figure 4: Swiss Similarity check of FDA approved drugs with Amentoflavone.

Library screened FDA approved drugs

Screening method Combined

Date Fri May 29 15:13:33 2020

If you publish these results, please cite the SwissSimilarity website.
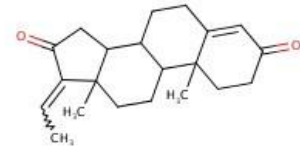
Results

DB00396, Progesterone Score : 0.995	**DB00378, Dydrogesterone** Score : 0.955	**DB00990, Exemestane** Score : 0.718	**DB06710, Methyltestosterone** Score : 0.651
DB00624, Testosterone Score : 0.651	**DB01420, Testosterone Propionate** Score : 0.617	**DB00253, Medrysone** Score : 0.513	**DB00603, Medroxyprogesterone Acetate** Score : 0.399
DB00294, Etonogestrel Score : 0.398	**DB00741, Hydrocortisone** Score : 0.391	**DB00367, Levonorgestrel** Score : 0.355	**DB00717, Norethindrone** Score : 0.355
DB01395, Drospirenone Score : 0.345	**DB09015, Potassium Canrenoate** Score : 0.199	**DB08804, Nandrolone decanoate** Score : 0.182	**DB01185, Fluoxymesterone** Score : 0.151

DB00687, Fludrocortisone **Score : 0.117**	**DB00896, Rimexolone** **Score : 0.115**	**DB06730, Gestodene** **Score : 0.107**	**DB01431, Allylestrenol** **Score : 0.105**
DB06689, Ethanolamine Oleate **Score : 0.105**	**DB00959, Methylprednisolone** **Score : 0.088**	**DB00860, Prednisolone** **Score : 0.088**	**DB00635, Prednisone** **Score : 0.088**
DB00894, Testolactone **Score : 0.065**	**DB00980, Ramelteon** **Score : 0.057**	**DB00421, Spironolactone** **Score : 0.035**	**DB06594, Agomelatine** **Score : 0.030**
DB00202, Succinylcholine **Score : 0.029**	**DB00461, Nabumetone** **Score : 0.028**	**DB01625, Isopropamide** **Score : 0.027**	**DB00546, Adinazolam** **Score : 0.026**
DB01058, Praziquantel **Score : 0.025**	**DB04839, Cyproterone acetate** **Score : 0.024**	**DB00351, Megestrol acetate** **Score : 0.024**	**DB01065, Melatonin** **Score : 0.023**
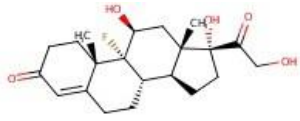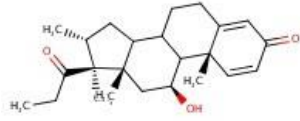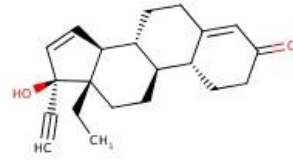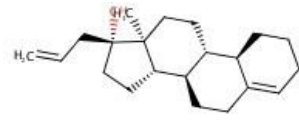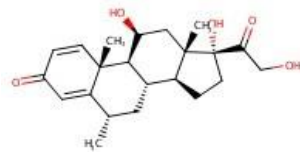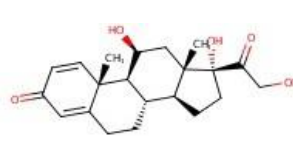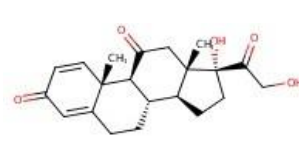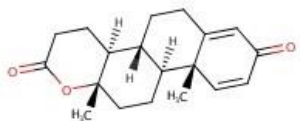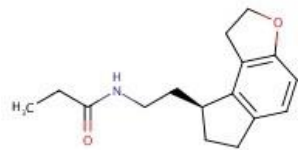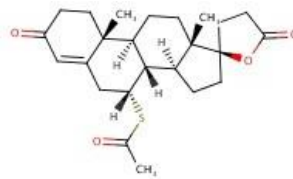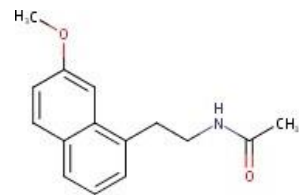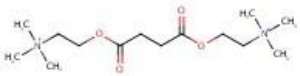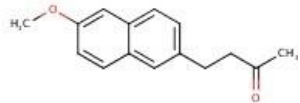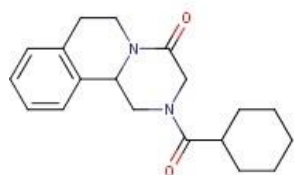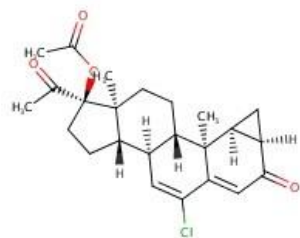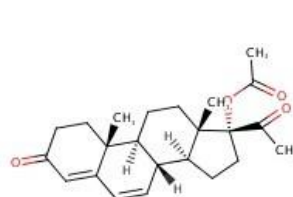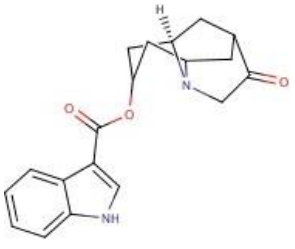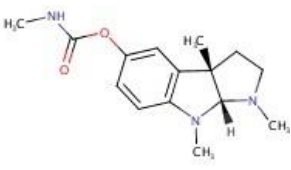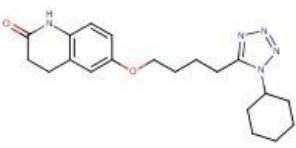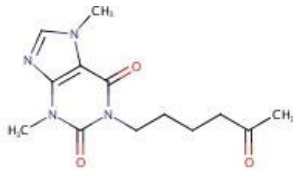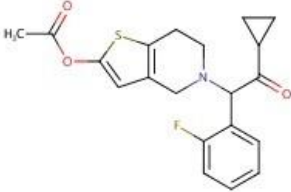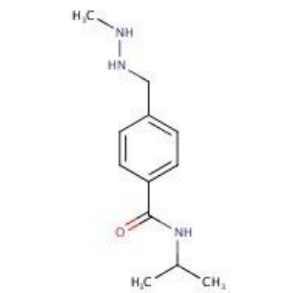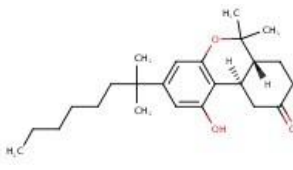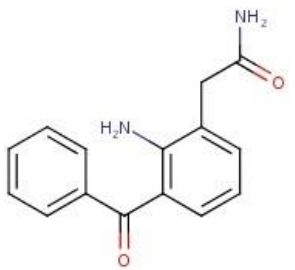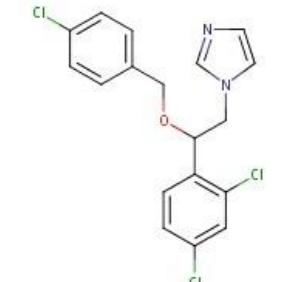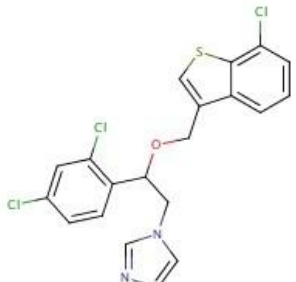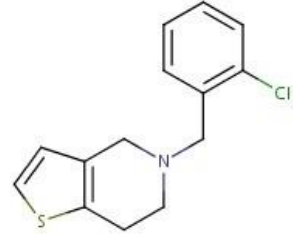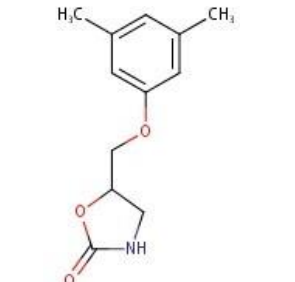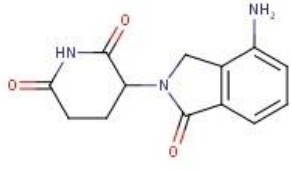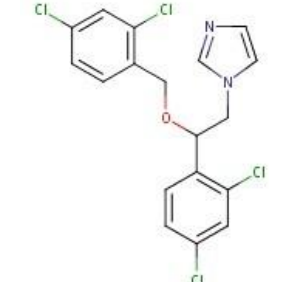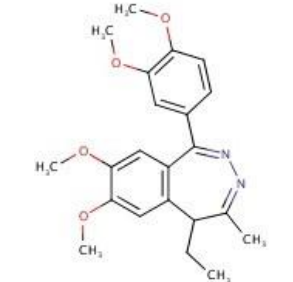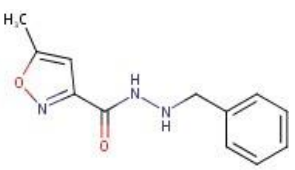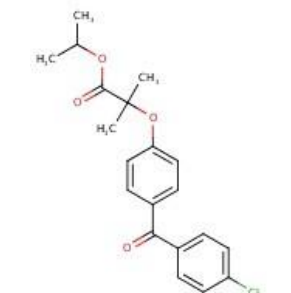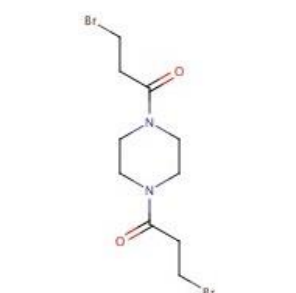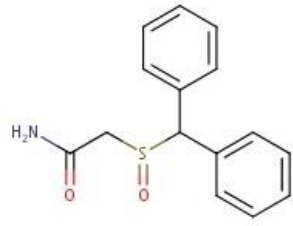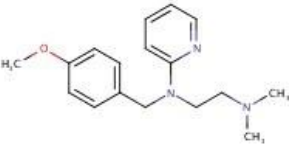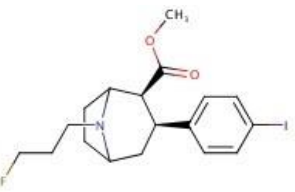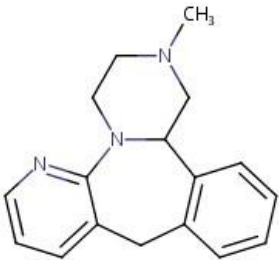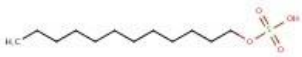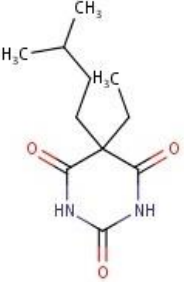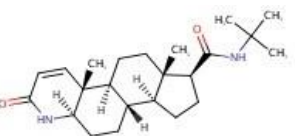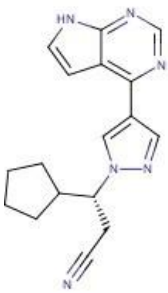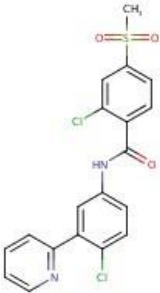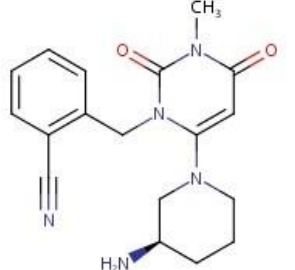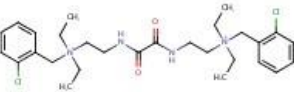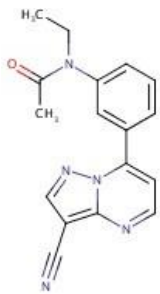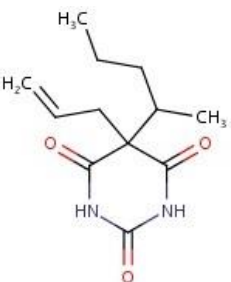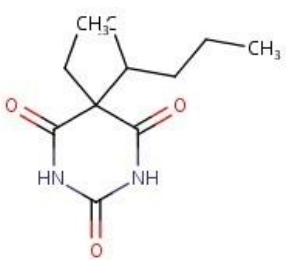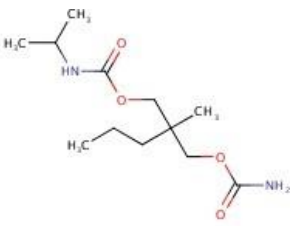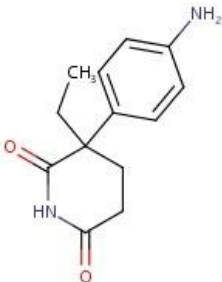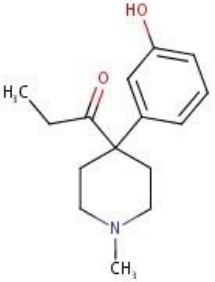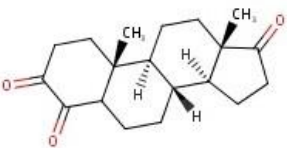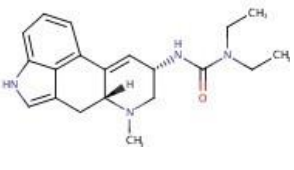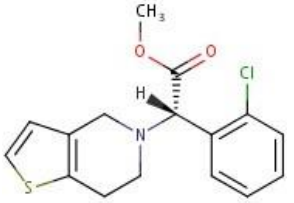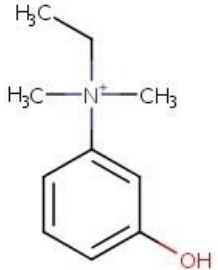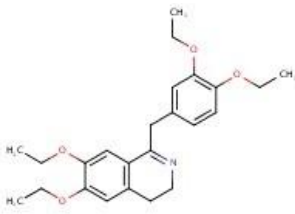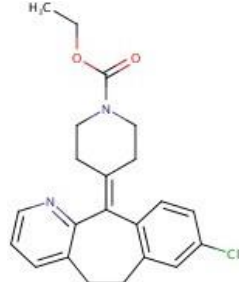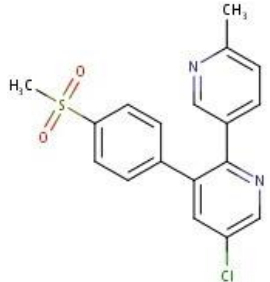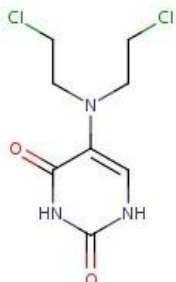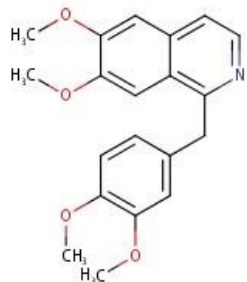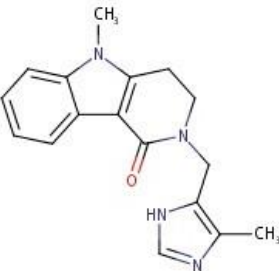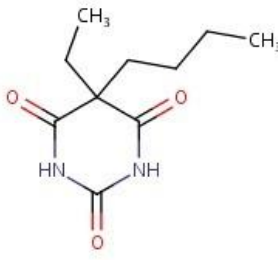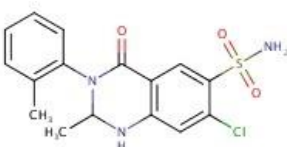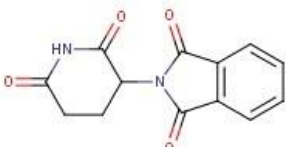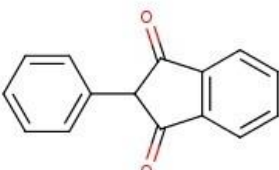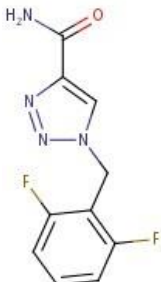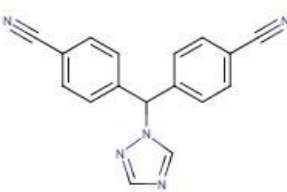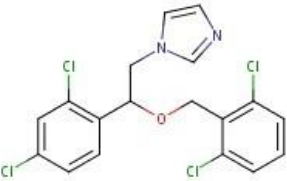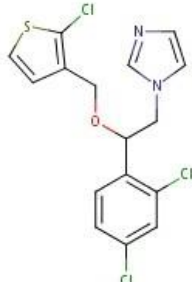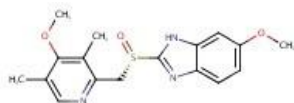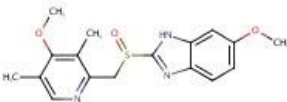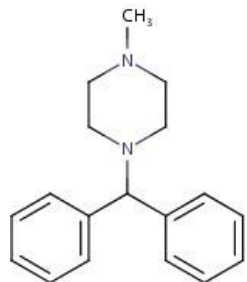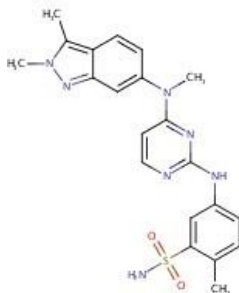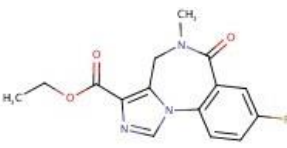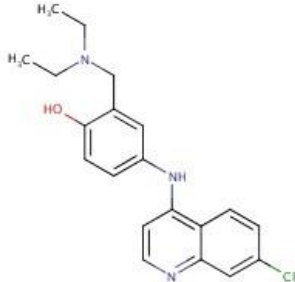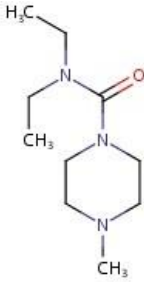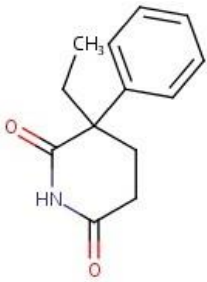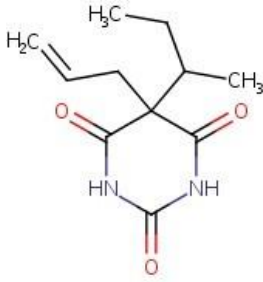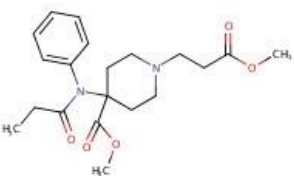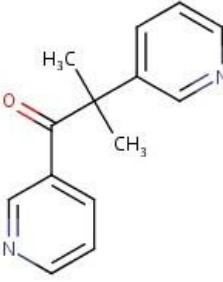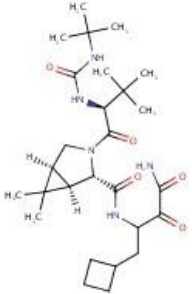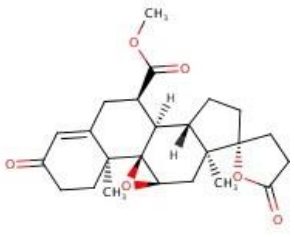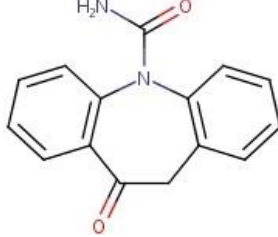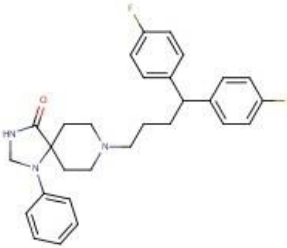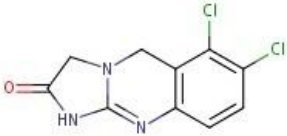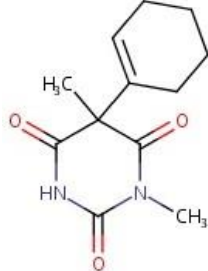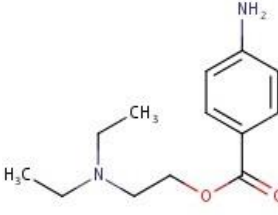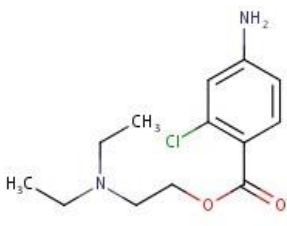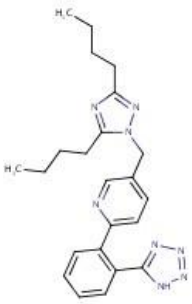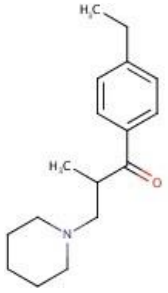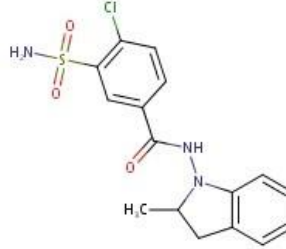
DB00757, Dolasetron	DB00981, Physostigmine	DB01166, Cilostazol	DB00806, Pentoxifylline
Score : 0.022	Score : 0.021	Score : 0.021	Score : 0.021
DB04880, Enoximone	DB06209, Prasugrel	DB01168, Procarbazine	DB00486, Nabilone
Score : 0.020	Score : 0.019	Score : 0.018	Score : 0.017
DB06802, Nepafenac	DB01127, Econazole	DB01153, Sertaconazole	DB00208, Ticlopidine
Score : 0.017	Score : 0.016	Score : 0.016	Score : 0.016
DB00660, Metaxalone	DB00480, Lenalidomide	DB01110, Miconazole	DB08811, Tofisopam
Score : 0.016	Score : 0.015	Score : 0.015	Score : 0.015
DB01247, Isocarboxazid	DB01039, Fenofibrate	DB00236, Pipobroman	DB00745, Modafinil
Score : 0.015	Score : 0.015	Score : 0.015	Score : 0.014

DB06691, Mepyramine **Score : 0.014**	**DB08824, Ioflupane I 123** **Score : 0.014**	**DB00370, Mirtazapine** **Score : 0.014**	**DB00815, Sodium lauryl sulfate** **Score : 0.014**
DB01351, Amobarbital **Score : 0.014**	**DB01216, Finasteride** **Score : 0.013**	**DB08877, Ruxolitinib** **Score : 0.013**	**DB08828, Vismodegib** **Score : 0.013**
DB01394, Colchicine **Score : 0.013**	**DB06203, Alogliptin** **Score : 0.013**	**DB01122, Ambenonium** **Score : 0.013**	**DB00962, Zaleplon** **Score : 0.013**
DB00418, Secobarbital **Score : 0.013**	**DB00312, Pentobarbital** **Score : 0.013**	**DB00395, Carisoprodol** **Score : 0.013**	**DB00357, Aminoglutethimide** **Score : 0.013**
DB06738, Ketobemidone **Score : 0.013**	**DB08905, Formestane** **Score : 0.012**	**DB00589, Lisuride** **Score : 0.012**	**DB00758, Clopidogrel** **Score : 0.012**

DB01010, Edrophonium **Score : 0.012**	**DB06751**, Drotaverine **Score : 0.012**	**DB00455**, Loratadine **Score : 0.012**	**DB01628**, Etoricoxib **Score : 0.012**
DB00791, Uracil mustard **Score : 0.011**	**DB01113**, Papaverine **Score : 0.011**	**DB00969**, Alosetron **Score : 0.011**	**DB01353**, Butethal **Score : 0.011**
DB00524, Metolazone **Score : 0.011**	**DB01041**, Thalidomide **Score : 0.010**	**DB00498**, Phenindione **Score : 0.010**	**DB06201**, Rufinamide **Score : 0.010**
DB01006, Letrozole **Score : 0.010**	**DB08943**, Isoconazole **Score : 0.010**	**DB01007**, Tioconazole **Score : 0.010**	**DB00736**, Esomeprazole **Score : 0.010**
DB00338, Omeprazole **Score : 0.010**	**DB01176**, Cyclizine **Score : 0.010**	**DB06589**, Pazopanib **Score : 0.010**	**DB01085**, Pilocarpine **Score : 0.010**

DB01205, **Flumazenil** **Score : 0.010**	**DB00613**, **Amodiaquine** **Score : 0.010**	**DB00711**, **Diethylcarbamazine** **Score : 0.010**	**DB01437**, **Glutethimide** **Score : 0.010**
DB00849, **Methylphenobarbital** **Score : 0.010**	**DB00306**, **Talbutal** **Score : 0.010**	**DB00899**, **Remifentanil** **Score : 0.010**	**DB00474**, **Methohexital** **Score : 0.010**
DB01011, **Metyrapone** **Score : 0.010**	**DB08873**, **Boceprevir** **Score : 0.009**	**DB00700**, **Eplerenone** **Score : 0.009**	**DB00776**, **Oxcarbazepine** **Score : 0.009**
DB04842, **Fluspirilene** **Score : 0.009**	**DB00261**, **Anagrelide** **Score : 0.009**	**DB01355**, **Hexobarbital** **Score : 0.009**	**DB00721**, **Procaine** **Score : 0.009**
DB01161, **Chloroprocaine** **Score : 0.009**	**DB01342**, **Forasartan** **Score : 0.009**	**DB08992**, **Eperisone** **Score : 0.009**	**DB00808**, **Indapamide** **Score : 0.009**
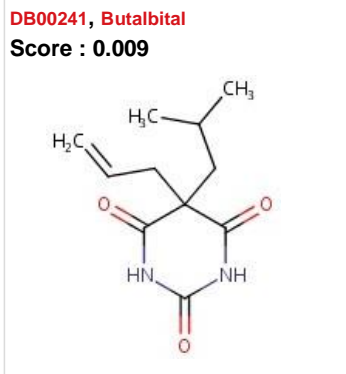
Supplementary figure 5: Swiss Similarity check of FDA approved drugs (drug ID, drug name, similarity score, and molecule structure) with Guggulsterone.

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

S. No.	Natural compound	Medicinal Property	Source	References
1	Amentoflavone	Antioxidant, Anticancer, Antibacterial, Antiviral, Anti-Inflammatory, and UV-Blocking Activities	*Torreyanucifera, Biophytumsensitivum,* and *Selaginellatamariscina*	[28]
2	Guggulsterone	Antibacterial, Antifungal, and Antiviral	*Commiphoramukul*	[29]
3	Puerarin	Anti-Inflammatory, and Anti-Viral Activities	Chinese herb Gegen	[30]
4	Piperine	Anti-Inflammatory, Antiviral, Antipyretic, Immune and Bioavailability	*Piper nigrum* L. (black pepper)	[31]

		Enhancing Qualities		
5	Maslinic acid	Antiviral, Antidiabetogenic, Anti-Inflammatory, and Antimicrobial	*Olea europaea* L.	[32]
6	Apigenin	Anti-Inflammatory, Anticancer, Antioxidative and Antiviral Function	Parsley, artichoke, basil, celery	[33]
7	Epigallocatechin	Antioxidant and antiviral	Green tea (*Cameriasinensis*)	[34]
8	Daidzein	Antiviral and Antidiabetic	Kudzu (*Puerarialobata* (Wild.)	[35]
9	Xanthohumol	Antioxidant, Antiproliferative, and Antiviral Activity	Hop, *Humuluslupulus*	[36]
10	Resveratrol	Antioxidant, Antiviral	Fermented grapes, mulberry, red wine, and	[37]

			peanuts	
11	Luteolin	Antiviral	Leaves of basil (Ocimumbasilicum), parsley (Petroselinum crispum) and spinach (Spinaciaoleracea), seeds of pepper (Capsicum annuum)	[38]
12	Cyanidin-3-o-galactoside	Antiviral	cranberry (*Vacciniummacrocarpon*)	[39]
13	Pectolinarin	Antiviral	*CirsiumSubcoriaceum*	[26, 40]
14	Herbacetin	Antiviral	ramose scouring rush, flaxseed, and Roemeria hybrid	[26, 41]
15	Rhoifolin	Antiviral, antioxidant, anti-inflammatory, antimicrobial, hepatoprotective and anticancer effects	Freshleaves of *Rhus succedanea*	[26, 42]

16	Ganomycin B	Antiviral	*Ganodermapfeifferi*	[43]
17	Phloretin	Anticancer, Antiosteoclastogenic, Anti-Fungal, Antiviral, Anti-Inflammatory, Antibacterial Andestrogenic Activities	apple, kum-quat, pear, strawberry, and vegetables	[44]
18	Crocetin	Antiviral	*Crocus sativus*	[45]

Table S1: Medicinal property of selected natural compounds

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

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 (Å²)	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	Pectolinarin	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 S2. List of compounds with suitable ADME properties

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

SwissTargetPrediction

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
Transitional endoplasmic reticulum ATPase	VCP	P55072	CHEMBL1075145	Primary active transporter	1.0	1 / 1
Placenta growth factor	PGF	P49763	CHEMBL1697671	Unclassified protein	1.0	1 / 1
Vascular endothelial growth factor A	VEGFA	P15692	CHEMBL1783	Secreted protein	1.0	1 / 1
GABA-A receptor; alpha-1/beta-2/gamma-2	GABRA1 GABRB2 GABRG2	P14867 P47870 P18507	CHEMBL2095172	Ligand-gated ion channel	1.0	1 / 1
Serotonin 2c (5-HT2c) receptor	HTR2C	P28335	CHEMBL225	Family A G protein-coupled receptor	1.0	1 / 2
Dopamine D3 receptor	DRD3	P35462	CHEMBL234	Family A G protein-coupled receptor	1.0	1 / 1
Delta opioid receptor	OPRD1	P41143	CHEMBL236	Family A G protein-coupled receptor	1.0	2 / 5
Protein-tyrosine phosphatase 1B	PTPN1	P18031	CHEMBL335	Phosphatase	1.0	6 / 18
Beta-secretase 1	BACE1	P56817	CHEMBL4822	Protease	1.0	10 / 17
Cyclin-dependent kinase 5/CDK5 activator 1	CDK5R1 CDK5	Q15078 Q00535	CHEMBL1907600	Kinase	0.0868858546069	1 / 17
NEDD8-activating enzyme E1 regulatory subunit	NAE1	Q13564	CHEMBL2016431	Unclassified protein	0.0	0 / 1
NADPH oxidase 4	NOX4	Q9NPH5	CHEMBL1250375	Enzyme	0.0	0 / 7
Aldose reductase (by homology)	AKR1B1	P15121	CHEMBL1900	Enzyme	0.0	0 / 61
Monoamine oxidase A	MAOA	P21397	CHEMBL1951	Oxidoreductase	0.0	0 / 6
Tyrosine-protein kinase receptor FLT3	FLT3	P36888	CHEMBL1974	Kinase	0.0	0 / 7
Cytochrome P450 19A1	CYP19A1	P11511	CHEMBL1978	Cytochrome P450	0.0	0 / 16
Estrogen receptor alpha	ESR1	P03372	CHEMBL206	Nuclear receptor	0.0	0 / 24
Cyclin-dependent kinase 1/cyclin B	CCNB3 CDK1 CCNB1 CCNB2	Q8WWL7 P06493 P14635 O95067	CHEMBL2094127	Other cytosolic protein	0.0	0 / 11
Acetylcholinesterase	ACHE	P22303	CHEMBL220	Hydrolase	0.0	0 / 33
Adenosine A1 receptor (by homology)	ADORA1	P30542	CHEMBL226	Family A G protein-coupled receptor	0.0	0 / 23
Cyclooxygenase-2	PTGS2	P35354	CHEMBL230	Oxidoreductase	0.0	0 / 18
Estrogen receptor beta	ESR2	Q92731	CHEMBL242	Nuclear receptor	0.0	0 / 21
Cyclin-dependent	CDK6	Q00534	CHEMBL2508	Kinase	0.0	0 / 4

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
kinase 6						
Adenosine A2a receptor (by homology)	ADORA2A	P29274	CHEMBL251	Family A G protein-coupled receptor	0.0	0 / 11
Tyrosine-protein kinase SYK	SYK	P43405	CHEMBL2599	Kinase	0.0	0 / 3
Glycogen synthase kinase-3 beta	GSK3B	P49841	CHEMBL262	Kinase	0.0	0 / 9
Transthyretin	TTR	P02766	CHEMBL3194	Secreted protein	0.0	0 / 2
Casein kinase II alpha	CSNK2A1	P68400	CHEMBL3629	Kinase	0.0	0 / 2
Cystic fibrosis transmembrane conductance regulator	CFTR	P13569	CHEMBL4051	Other ion channel	0.0	0 / 1
Aldo-keto reductase family 1 member B10	AKR1B10	O60218	CHEMBL5983	Enzyme	0.0	0 / 3
Tankyrase-2	TNKS2	Q9H2K2	CHEMBL6154	Enzyme	0.0	0 / 12
Tankyrase-1	TNKS	O95271	CHEMBL6164	Enzyme	0.0	0 / 27
Tyrosinase	TYR	P14679	CHEMBL1973	Oxidoreductase	0.0	0 / 2
Arachidonate 5-lipoxygenase	ALOX5	P09917	CHEMBL215	Oxidoreductase	0.0	0 / 47
Aryl hydrocarbon receptor	AHR	P35869	CHEMBL3201	Transcription factor	0.0	0 / 1
Estrogen-related receptor alpha	ESRRA	P11474	CHEMBL3429	Nuclear receptor	0.0	0 / 2
P-glycoprotein 1	ABCB1	P08183	CHEMBL4302	Primary active transporter	0.0	0 / 48
Xanthine dehydrogenase	XDH	P47989	CHEMBL1929	Oxidoreductase	0.0	6 / 20
Receptor-type tyrosine-protein phosphatase S	PTPRS	Q13332	CHEMBL2396508	Phosphatase	0.0	0 / 8
AMY1C	AMY1A	P04745	CHEMBL2478	Enzyme	0.0	0 / 1
Carbonic anhydrase II	CA2	P00918	CHEMBL205	Lyase	0.0	5 / 11
G-protein coupled receptor 35	GPR35	Q9HC97	CHEMBL1293267	Family A G protein-coupled receptor	0.0	0 / 2
Death-associated protein kinase 1	DAPK1	P53355	CHEMBL2558	Kinase	0.0	0 / 2
DNA-3-methyladenine glycosylase	MPG	P29372	CHEMBL3396943	Enzyme	0.0	0 / 1
Solute carrier family 22 member 12	SLC22A12	Q96S37	CHEMBL6120	Electrochemical transporter	0.0	0 / 1
Carbonyl reductase [NADPH] 1	CBR1	P16152	CHEMBL5586	Enzyme	0.0	0 / 2
Butyrylcholinesterase	BCHE	P06276	CHEMBL1914	Hydrolase	0.0	0 / 5
Adenosine A3 receptor	ADORA3	P0DMS8	CHEMBL256	Family A G protein-coupled receptor	0.0	0 / 19
Glyoxalase I	GLO1	Q04760	CHEMBL2424	Enzyme	0.0	0 / 4
Poly [ADP-ribose]	PARP1	P09874	CHEMBL3105	Enzyme	0.0	0 / 9

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
polymerase-1						
Matrix metalloproteinase 9	MMP9	P14780	CHEMBL321	Protease	0.0	0 / 2
Matrix metalloproteinase 2	MMP2	P08253	CHEMBL333	Protease	0.0	0 / 2
Matrix metalloproteinase 12	MMP12	P39900	CHEMBL4393	Protease	0.0	0 / 1
Lymphocyte differentiation antigen CD38	CD38	P28907	CHEMBL4660	Enzyme	0.0	0 / 2
DNA topoisomerase I (by homology)	TOP1	P11387	CHEMBL1781	Isomerase	0.0	0 / 1
Arginase-1 (by homology)	ARG1	P05089	CHEMBL1075097	Enzyme	0.0	0 / 2
Carbonic anhydrase I	CA1	P00915	CHEMBL261	Lyase	0.0	3 / 7
Carbonic anhydrase IX	CA9	Q16790	CHEMBL3594	Lyase	0.0	3 / 5
Lysine-specific demethylase 4D-like	KDM4E	B2RXH2	CHEMBL1293226	Eraser	0.0	0 / 2
Arachidonate 15-lipoxygenase	ALOX15	P16050	CHEMBL2903	Enzyme	0.0	0 / 8
Cyclin-dependent kinase 1	CDK1	P06493	CHEMBL308	Kinase	0.0	0 / 10
Arachidonate 12-lipoxygenase	ALOX12	P18054	CHEMBL3687	Enzyme	0.0	0 / 11
Microtubule-associated protein tau	MAPT	P10636	CHEMBL1293224	Unclassified protein	0.0	0 / 1
Vasopressin V2 receptor	AVPR2	P30518	CHEMBL1790	Family A G protein-coupled receptor	0.0	0 / 1
DNA topoisomerase II alpha	TOP2A	P11388	CHEMBL1806	Isomerase	0.0	0 / 1
Insulin receptor	INSR	P06213	CHEMBL1981	Kinase	0.0	0 / 1
Serine/threonine-protein kinase PIM1	PIM1	P11309	CHEMBL2147	Kinase	0.0	0 / 7
Serine/threonine-protein kinase Aurora-B	AURKB	Q96GD4	CHEMBL2185	Kinase	0.0	0 / 4
Dopamine D4 receptor	DRD4	P21917	CHEMBL219	Family A G protein-coupled receptor	0.0	0 / 1
Myosin light chain kinase, smooth muscle	MYLK	Q15746	CHEMBL2428	Kinase	0.0	0 / 1
Myeloperoxidase	MPO	P05164	CHEMBL2439	Enzyme	0.0	0 / 1
PI3-kinase p85-alpha subunit	PIK3R1	P27986	CHEMBL2506	Enzyme	0.0	0 / 1
Liver glycogen phosphorylase	PYGL	P06737	CHEMBL2568	Enzyme	0.0	0 / 1
Tyrosine-protein kinase SRC	SRC	P12931	CHEMBL267	Kinase	0.0	0 / 6
Focal adhesion kinase	PTK2	Q05397	CHEMBL2695	Kinase	0.0	0 / 2

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
1						
Vascular endothelial growth factor receptor 2	KDR	P35968	CHEMBL279	Kinase	0.0	0 / 3
Matrix metalloproteinase 13	MMP13	P45452	CHEMBL280	Protease	0.0	0 / 1
Matrix metalloproteinase 3	MMP3	P08254	CHEMBL283	Protease	0.0	0 / 1
Carbonic anhydrase III	CA3	P07451	CHEMBL2885	Lyase	0.0	0 / 1
Serine/threonine-protein kinase PLK1	PLK1	P53350	CHEMBL3024	Kinase	0.0	0 / 3
PI3-kinase p110-gamma subunit	PIK3CG	P48736	CHEMBL3267	Enzyme	0.0	0 / 1
Protein kinase N1	PKN1	Q16512	CHEMBL3384	Kinase	0.0	0 / 3
Carbonic anhydrase XIV	CA14	Q9ULX7	CHEMBL3510	Lyase	0.0	0 / 1
Hepatocyte growth factor receptor	MET	P08581	CHEMBL3717	Kinase	0.0	0 / 4
Serine/threonine-protein kinase NEK2	NEK2	P51955	CHEMBL3835	Kinase	0.0	0 / 2
Interleukin-8 receptor A	CXCR1	P25024	CHEMBL4029	Family A G protein-coupled receptor	0.0	0 / 1
CaM kinase II beta	CAMK2B	Q13554	CHEMBL4121	Kinase	0.0	0 / 2
ALK tyrosine kinase receptor	ALK	Q9UM73	CHEMBL4247	Kinase	0.0	0 / 4
Serine/threonine-protein kinase AKT	AKT1	P31749	CHEMBL4282	Kinase	0.0	0 / 4
Serine/threonine-protein kinase NEK6	NEK6	Q9HC98	CHEMBL4309	Kinase	0.0	0 / 2
Phospholipase A2 group 1B	PLA2G1B	P04054	CHEMBL4426	Enzyme	0.0	0 / 1
Carbonic anhydrase VA	CA5A	P35218	CHEMBL4789	Lyase	0.0	0 / 1
Tyrosine-protein kinase receptor UFO	AXL	P30530	CHEMBL4895	Kinase	0.0	0 / 4
DNA-(apurinic or apyrimidinic site) lyase	APEX1	P27695	CHEMBL5619	Enzyme	0.0	0 / 1
NUAK family SNF1-like kinase 1	NUAK1	O60285	CHEMBL5784	Kinase	0.0	0 / 2
Aldo-keto reductase family 1 member C2 (by homology)	AKR1C2	P52895	CHEMBL5847	Enzyme	0.0	0 / 1
Aldo-keto reductase family 1 member C1 (by homology)	AKR1C1	Q04828	CHEMBL5905	Enzyme	0.0	0 / 1
Aldo-keto-reductase family 1 member C3 (by homology)	AKR1C3	P42330	CHEMBL4681	Enzyme	0.0	0 / 1
Aldo-keto reductase family 1 member C4	AKR1C4	P17516	CHEMBL4999	Enzyme	0.0	0 / 1

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
(by homology)						
Carbonic anhydrase XIII (by homology)	CA13	Q8N1Q1	CHEMBL3912	Lyase	0.0	0 / 1

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

Amaresh Mishra[a], Yamini Pathak[a], Gourav Choudhir[b], Anuj Kumar[c, d], Surabhi Kirti Mishra[e], Vishwas Tripathi[a]✉

[a]School of Biotechnology, Gautam Buddha University, Greater Noida-201310, India

[b]Department of Botany, Ch. Charan Singh University, Meerut, UP-250004, India

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

[d]Advanced Centre for Computational and Applied Biotechnology, Uttarakhand Council for Biotechnology (UCB), Dehradun-248007, India

[e]School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Email: drvishwastripathi@gmail.com, vishwas@gbu.ac.in

SwissTargetPrediction

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
Androgen Receptor	AR	P10275	CHEMBL1871	Nuclear receptor	1.0	111 / 105
Mineralocorticoid receptor	NR3C2	P08235	CHEMBL1994	Nuclear receptor	1.0	76 / 27
Glucocorticoid receptor	NR3C1	P04150	CHEMBL2034	Nuclear receptor	1.0	124 / 37
Progesterone receptor	PGR	P06401	CHEMBL208	Nuclear receptor	1.0	163 / 50
Pregnane X receptor	NR1I2	O75469	CHEMBL3401	Nuclear receptor	1.0	3 / 3
Cytochrome P450 17A1	CYP17A1	P05093	CHEMBL3522	Cytochrome P450	0.864804765933	67 / 27
Corticosteroid binding globulin	SERPINA6	P08185	CHEMBL2421	Secreted protein	0.634859674355	5 / 20
Sigma opioid receptor	SIGMAR1	Q99720	CHEMBL287	Membrane receptor	0.634859674355	20 / 3
Testis-specific androgen-binding protein	SHBG	P04278	CHEMBL3305	Secreted protein	0.634859674355	3 / 39
Fatty acid-binding protein, liver (by homology)	FABP1	P07148	CHEMBL5421	Fatty acid binding protein family	0.634859674355	1 / 3
Cytochrome P450 19A1	CYP19A1	P11511	CHEMBL1978	Cytochrome P450	0.626506245743	535 / 326
Estradiol 17-beta-dehydrogenase 3	HSD17B3	P37058	CHEMBL4234	Enzyme	0.468147849923	11 / 1
Steroid 5-alpha-reductase 1	SRD5A1	P18405	CHEMBL1787	Oxidoreductase	0.459896352864	69 / 39
Steroid 5-alpha-reductase 2	SRD5A2	P31213	CHEMBL1856	Oxidoreductase	0.459896352864	28 / 74
Nuclear receptor subfamily 1 group I member 3 (by homology)	NR1I3	Q14994	CHEMBL5503	Nuclear receptor	0.301554969452	2 / 2
Niemann-Pick C1-like protein 1	NPC1L1	Q9UHC9	CHEMBL2027	Other membrane protein	0.109945769839	0 / 7
C-C chemokine receptor type 5	CCR5	P51681	CHEMBL274	Family A G protein-coupled receptor	0.109945769839	9 / 1
11-beta-hydroxysteroid dehydrogenase 1	HSD11B1	P28845	CHEMBL4235	Enzyme	0.101613854776	405 / 31
Dual specificity phosphatase Cdc25A	CDC25A	P30304	CHEMBL3775	Phosphatase	0.101613854776	5 / 6
Estrogen receptor alpha	ESR1	P03372	CHEMBL206	Nuclear receptor	0.101613854776	71 / 24
Estrogen receptor beta	ESR2	Q92731	CHEMBL242	Nuclear receptor	0.101613854776	67 / 27
Dopamine transporter	SLC6A3	Q01959	CHEMBL238	Electrochemical transporter	0.101613854776	38 / 1
Protein-tyrosine phosphatase 2C	PTPN11	Q06124	CHEMBL3864	Phosphatase	0.101613854776	0 / 1

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
Aldo-keto reductase family 1 member B10	AKR1B10	O60218	CHEMBL5983	Enzyme	0.101613854776	0 / 6
Butyrylcholinesterase	BCHE	P06276	CHEMBL1914	Hydrolase	0.101613854776	23 / 2
Cathepsin D	CTSD	P07339	CHEMBL2581	Protease	0.101613854776	3 / 1
Prostaglandin E synthase	PTGES	O14684	CHEMBL5658	Enzyme	0.101613854776	7 / 12
Adenosine A3 receptor	ADORA3	P0DMS8	CHEMBL256	Family A G protein-coupled receptor	0.101613854776	16 / 3
Protein-tyrosine phosphatase 1B	PTPN1	P18031	CHEMBL335	Phosphatase	0.101613854776	18 / 42
Cyclooxygenase-1	PTGS1	P23219	CHEMBL221	Oxidoreductase	0.101613854776	40 / 2
11-beta-hydroxysteroid dehydrogenase 2	HSD11B2	P80365	CHEMBL3746	Enzyme	0.101613854776	3 / 24
MAP kinase ERK1	MAPK3	P27361	CHEMBL3385	Kinase	0.101613854776	10 / 1
Arachidonate 5-lipoxygenase	ALOX5	P09917	CHEMBL215	Oxidoreductase	0.101613854776	20 / 19
Nitric oxide synthase, inducible	NOS2	P35228	CHEMBL4481	Enzyme	0.101613854776	61 / 21
Anandamide amidohydrolase	FAAH	O00519	CHEMBL2243	Enzyme	0.101613854776	55 / 14
Vanilloid receptor	TRPV1	Q8NER1	CHEMBL4794	Voltage-gated ion channel	0.101613854776	306 / 2
T-cell protein-tyrosine phosphatase	PTPN2	P17706	CHEMBL3807	Phosphatase	0.101613854776	1 / 15
Phosphodiesterase 4B	PDE4B	Q07343	CHEMBL275	Phosphodiesterase	0.101613854776	97 / 0
Beta-secretase 1	BACE1	P56817	CHEMBL4822	Protease	0.101613854776	25 / 1
Protein farnesyltransferase	FNTA FNTB	P49354 P49356	CHEMBL2094108	Enzyme	0.101613854776	173 / 6
Histamine H1 receptor	HRH1	P35367	CHEMBL231	Family A G protein-coupled receptor	0.101613854776	15 / 0
Platelet activating factor receptor	PTAFR	P25105	CHEMBL250	Family A G protein-coupled receptor	0.101613854776	26 / 0
Protein kinase C eta	PRKCH	P24723	CHEMBL3616	Kinase	0.101613854776	3 / 1
Melatonin receptor 1A	MTNR1A	P48039	CHEMBL1945	Family A G protein-coupled receptor	0.101613854776	531 / 0
Melatonin receptor 1B	MTNR1B	P49286	CHEMBL1946	Family A G protein-coupled receptor	0.101613854776	469 / 0
Carboxylesterase 2	CES2	O00748	CHEMBL3180	Enzyme	0.101613854776	0 / 14
Serotonin transporter	SLC6A4	P31645	CHEMBL228	Electrochemical transporter	0.101613854776	37 / 5
Transient receptor potential cation channel subfamily A member 1	TRPA1	O75762	CHEMBL6007	Voltage-gated ion channel	0.101613854776	6 / 2
Protein-tyrosine phosphatase 1C	PTPN6	P29350	CHEMBL3166	Phosphatase	0.101613854776	0 / 1
Bromodomain-containing protein 4	BRD4	O60885	CHEMBL1163125	Reader	0.101613854776	135 / 0
Bromodomain-	BRD2	P25440	CHEMBL1293289	Reader	0.101613854776	67 / 0

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
containing protein 2						
Bromodomain-containing protein 3	BRD3	Q15059	CHEMBL1795186	Reader	0.101613854776	68 / 0
Cathepsin K	CTSK	P43235	CHEMBL268	Protease	0.101613854776	112 / 0
Cathepsin L	CTSL	P07711	CHEMBL3837	Protease	0.101613854776	72 / 0
Cathepsin (B and K)	CTSB	P07858	CHEMBL4072	Protease	0.101613854776	61 / 0
Poly [ADP-ribose] polymerase-1	PARP1	P09874	CHEMBL3105	Enzyme	0.101613854776	230 / 0
Epoxide hydrolase 1	EPHX1	P07099	CHEMBL1968	Protease	0.101613854776	33 / 0
Epoxide hydratase	EPHX2	P34913	CHEMBL2409	Protease	0.101613854776	259 / 0
Kappa Opioid receptor	OPRK1	P41145	CHEMBL237	Family A G protein-coupled receptor	0.101613854776	31 / 0
Centromere-associated protein E	CENPE	Q02224	CHEMBL5870	Unclassified protein	0.101613854776	1 / 0
Monoamine oxidase A	MAOA	P21397	CHEMBL1951	Oxidoreductase	0.0	55 / 0
Monoamine oxidase B	MAOB	P27338	CHEMBL2039	Oxidoreductase	0.0	137 / 0
Neurokinin 2 receptor	TACR2	P21452	CHEMBL2327	Family A G protein-coupled receptor	0.0	10 / 0
Neurokinin 1 receptor	TACR1	P25103	CHEMBL249	Family A G protein-coupled receptor	0.0	71 / 0
Tachykinin-3	TAC3	Q9UHF0	CHEMBL3707470	Unclassified protein	0.0	5 / 0
G protein-coupled receptor 44	PTGDR2	Q9Y5Y4	CHEMBL5071	Family A G protein-coupled receptor	0.0	6 / 0
Serotonin 2c (5-HT2c) receptor	HTR2C	P28335	CHEMBL225	Family A G protein-coupled receptor	0.0	31 / 0
PI3-kinase p110-alpha subunit	PIK3CA	P42336	CHEMBL4005	Enzyme	0.0	55 / 0
Voltage-gated potassium channel subunit Kv1.5	KCNA5	P22460	CHEMBL4306	Voltage-gated ion channel	0.0	152 / 0
Voltage-gated potassium channel subunit Kv1.3	KCNA3	P22001	CHEMBL4633	Voltage-gated ion channel	0.0	30 / 0
C5a anaphylatoxin chemotactic receptor	C5AR1	P21730	CHEMBL2373	Family A G protein-coupled receptor	0.0	4 / 0
HMG-CoA reductase	HMGCR	P04035	CHEMBL402	Oxidoreductase	0.0	0 / 4
Cytochrome P450 51 (by homology)	CYP51A1	Q16850	CHEMBL3849	Cytochrome P450	0.0	0 / 2
Metabotropic glutamate receptor 1	GRM1	Q13255	CHEMBL3772	Family C G protein-coupled receptor	0.0	46 / 0
Translocator protein (by homology)	TSPO	P30536	CHEMBL5742	Membrane receptor	0.0	117 / 0
FK506-binding protein 1A	FKBP1A	P62942	CHEMBL1902	Isomerase	0.0	46 / 0
Sodium-dependent neutral amino acid transporter B(0)AT2	SLC6A15	Q9H2J7	CHEMBL3351189	Electrochemical transporter	0.0	7 / 0
Cytochrome P450 11B1	CYP11B1	P15538	CHEMBL1908	Cytochrome P450	0.0	165 / 0
HERG	KCNH2	Q12809	CHEMBL240	Voltage-gated ion	0.0	49 / 0

Target	Common name	Uniprot ID	ChEMBL ID	Target Class	Probability*	Known actives (3D/2D)
				channel		
Cytochrome P450 11B2	CYP11B2	P19099	CHEMBL2722	Cytochrome P450	0.0	190 / 0
Alcohol dehydrogenase alpha chain	ADH1A	P07327	CHEMBL1970	Oxidoreductase	0.0	3 / 0
Alcohol dehydrogenase gamma chain	ADH1C	P00326	CHEMBL3285	Oxidoreductase	0.0	2 / 0
DNA polymerase beta (by homology)	POLB	P06746	CHEMBL2392	Enzyme	0.0	0 / 7
Serine/threonine-protein kinase PIM1	PIM1	P11309	CHEMBL2147	Kinase	0.0	78 / 0
Dual-specificity tyrosine-phosphorylation regulated kinase 1A	DYRK1A	Q13627	CHEMBL2292	Kinase	0.0	28 / 0
Casein kinase I gamma 2	CSNK1G2	P78368	CHEMBL2543	Kinase	0.0	2 / 0
Phosphodiesterase 4A	PDE4A	P27815	CHEMBL254	Phosphodiesterase	0.0	56 / 0
Casein kinase I delta	CSNK1D	P48730	CHEMBL2828	Kinase	0.0	4 / 0
Dual specificity protein kinase CLK4	CLK4	Q9HAZ1	CHEMBL4203	Kinase	0.0	12 / 0
Dual specificty protein kinase CLK1	CLK1	P49759	CHEMBL4224	Kinase	0.0	12 / 0
Dual specificity protein kinase CLK2	CLK2	P49760	CHEMBL4225	Kinase	0.0	6 / 0
Dual specificity protein kinase CLK3	CLK3	P49761	CHEMBL4226	Kinase	0.0	2 / 0
Casein kinase I epsilon	CSNK1E	P49674	CHEMBL4937	Kinase	0.0	1 / 0
Casein kinase I isoform gamma-3	CSNK1G3	Q9Y6M4	CHEMBL5084	Kinase	0.0	1 / 0
Serine/threonine-protein kinase PIM3	PIM3	Q86V86	CHEMBL5407	Kinase	0.0	19 / 0
Dual specificity tyrosine-phosphorylation-regulated kinase 1B	DYRK1B	Q9Y463	CHEMBL5543	Kinase	0.0	8 / 0
SPS1/STE20-related protein kinase YSK4	MAP3K19	Q56UN5	CHEMBL6191	Kinase	0.0	1 / 0
Dual specificity protein phosphatase 3	DUSP3	P51452	CHEMBL2635	Phosphatase	0.0	5 / 0
dUTP pyrophosphatase	DUT	P33316	CHEMBL5203	Enzyme	0.0	28 / 0
Quinone reductase 2	NQO2	P16083	CHEMBL3959	Enzyme	0.0	42 / 0