International Pharmacy Acta. 2021;4(1):e9
Original Research Article
OPEN ACCESS
ISSN: 2645-3266

Design and discovery of some novel protease inhibitors against SARS-CoV-2 main protease by molecular docking, drug-likeness and ADME studies: An in-silico approach

Arka Das[a*], Subhasis Banerjee[b], Kalyan Kumar Sen[b]

a. Department of Pharmaceutical Chemistry, Acharya & BM Reddy College of Pharmacy, Soladevanahalli, Bengaluru-560107, Karnataka, India.
b. Department of Pharmaceutical Technology, Gupta College of Technological Sciences, Ashram More, Asansol-713301, West Bengal, India.

Article Info:
Received: June 2021
Accepted: August 2021
Published online: August 2021

*** Corresponding Author:**
Arka Das
Email: arkadas5196@gmail.com

Abstract:

SARS-CoV-2 emerges as a new threat to the life of global population. The high infectivity and spreading rate of the disease make it a pandemic with no such specific drug discovered yet. Considering the high spreading rate of the disease, there is an urgent need for selective anti-SARS-CoV-2 agent. SARS-CoV-2 main protease is an important target involved in transcription of the viral RNA, inhibition of which may lead to virucidal action. Repurposing strategy proved some antivirals to be effective against Mpro, but safety issues are of concern. Identifying lead and computational approaches are the best to consider. The present study incorporates three standard anti-HIV agents Lopinavir, Ritonavir, and Indinavir to undergo pharmacophore modeling. The initial modeling resulted in the selection of few test compounds considering the low RMSD as observed in Zinc database. The Rest of the compounds were designed from the pharmacophoric features of the newly developed model. 20 compounds were subjected to molecular docking. The docking results showed that, compound 20 revealed the highest binding energy (-8.6 kcal/mol), which is even lesser than all the three standards. The other compounds 3, 4, 5, 11 and 19 also responded well to the docking study. These six compounds were further evaluated for their drug-likeness and ADME properties to raise the acceptance level of the lead(s). Further computational study included the Molecular-Dynamic simulation of the compound 20, to ensure the least variations throughout the simulation analysis. The above sequential computational study provides a hypothetical guideline to optimize the lead as effective anti-SARS-CoV-2 agent.

Keywords: Molecular docking; Protease inhibitors; SARS-CoV-2; Main protease; ADME; Drug-likeness

Please Cite this article as: Das A., Banerjee S., Kumar Sen K. Design and discovery of some novel protease inhibitors against SARS-CoV-2 main protease by molecular docking, drug-likeness and ADME studies: An in-silico approach. Int. Pharm. Acta. 2021;4(1):e9
DOI: https://doi.org/10.22037/ipa.v4i1.35246

1. Introduction

2019 novel Coronavirus outbreak first began from the Wuhan province of China, in December 2019[1], causing Coronavirus disease 2019 (COVID-19). Within a very short span of time, it extended towards other parts of China and also to other countries of the world and engulfed a large number of populations, increasing the rate of morbidity day by day. Novel Coronavirus belongs to the family of *Coronaviridae* of order *Nidovirales* and genus *Coronavirus* [2,3]. the rapid global spreading of the disease, its increased rate of infection, hospitalization, and morbidity, compelled World Health Organization (WHO) to declare COVID-19 as a pandemic on March 11, 2020 [4]. Major symptoms of this disease include high fever, cough, muscular fatigue, pulmonary fibrosis, and alveolar cell injury progressing towards pneumonia. The Coronavirus mainly binds to

Angiotensin-Converting Enzyme 2 (ACE-2) of the host body, which is mainly expressed by the alveolar epithelial cells [5]. All these may lead to respiratory distress, chest pain, breathlessness and sometimes may cause silent hypoxia [6], leading to death. As of August 25, 2020, more than 23 million cases were confirmed, and over 810000 death cases were reported worldwide [7]. The International Committee on Taxonomy of Viruses (ICTV) renamed the previous name of novel Coronavirus (2019-nCoV) to Severe acute respiratory syndrome Coronavirus 2 or SARS-CoV-2 [8,9].

In spite of extensive research work, there is no effective medication developed to date to combat COVID-19. Researchers including medicinal chemists are tirelessly searching for potential molecules that can be developed as anti-SARS-CoV-2 agents. Owing to the current emergency situation of the pandemic and the increased rate of confirmed cases every day, there is an urgent need for the development of drug candidates against this virus, within a limited span of time. Unavailability of the appropriate treatment procedure and shortage of time compelled the scientists to focus on the drug repurposing approach, of previously FDA-approved drugs of other pharmacological categories [10,11]. Amongst all the drug targets of SARS-CoV-2, one of the most important targets is the Coronavirus main protease (Mpro), also called 3-Chymotrypsin like protease (3CLPro), mainly involved in the cleavage of two polyproteins PP1a and PP1ab to produce various functional proteins, thus mediating viral replication and transcription [12-14], therefore, inhibiting this enzyme can be a successful strategy to prevent COVID-19. Current situations cannot possibly warrant the development of a safe drug molecule within a short time, which can combat COVID-19. Main protease of SARS-CoV-2 is one of the most important targets, inhibition of which can leads to virucidal actions.

Various antiviral agents have been successfully employed and proved to be effective against Mpro by *in-silico* studies. Shah B *et al*, selected 61 antiviral agents and performed molecular docking on five Mpro proteins, and found that Lopinavir interacted with all five proteins, whereas Ritonavir, and Indinavir showed promising binding affinities with four out of five proteins [15]. Various other studies also proved that Lopinavir, Ritonavir, and Indinavir can be potential drug candidates in targeting Mpro. So, in the current study, we considered these three anti-HIV drugs (protease inhibitors) as standards for developing new derivatives and searching new chemical entities, based on pharmacophore modeling, against Coronavirus Mpro. the alarming situations of the pandemic and immediate requirement for an active anti-SARS-CoV-2 agent, urged the use of computational chemistry and virtual screening as important tools for screening a large number of active compounds from the databases to evolve candidates suitable against COVID-19. This initiative could help accelerating the global efforts to fight against current outbreak.

Pharmacophores are the important functional groups or structural features on a molecule, responsible for optimal interactions with the target proteins and are responsible for the observed therapeutic effect. Pharmacophore modeling includes the 3D spatial arrangement of the functional groups required for its activity. In this study, ligand-based pharmacophore modeling [16] was carried out, by extracting the common pharmacophoric features present in Lopinavir, Ritonavir, and Indinavir, essential for interacting with the target (Mpro). Molecular docking methodology is one of the in-silico methods widely used to screen several compounds for their binding affinity towards a specific macromolecule [17]. Docking predicts the most preferred orientation of a ligand, both geometrically and energetically (best fit) within the active site of a protein. This allows more optimized prediction of binding modes, conformations of the ligands within the binding site of the protein and thus enables accurate prediction of the binding pattern, describing the stability of complexes and estimating the strength of binding [18]. The procedure of docking involves three components namely, identification of the binding site of the protein, search algorithm, and scoring functions. Search algorithm assess all the possible orientations of ligands for appropriate binding within the binding site of the receptor (protein) and scoring function determines the receptor-ligand binding affinities via various mathematical functions, thus predicting binding energies of the complex, lower the binding energy, more stable is the complex with correct binding mode. It is observed that several virtually screened compounds, even after receiving a positive nod through docking often fail to become an eligible clinical candidate due to poor pharmacokinetic profile and noncompliance of the safety and drug-likeness parameters. Therefore, drug-likeness and prediction of Absorption, Distribution, Metabolism, and Excretion (ADME) is an important aspect to consider prior to developing the novel active molecules synthetically. Drug-likeness is an important facet to be accounted for during lead optimization. It describes whether the ligand comply all the criteria to become ‘drug-like’. It basically reflects the bioavailability of a drug, and how a hit molecule can be successfully developed to a lead compound, reducing the adverse drug reactions and can successfully reach the market, in a cost-effective manner. It is often observed that a compound with a greater affinity towards a macromolecule, suffers from a poor pharmacokinetic profile. Therefore, during the early drug discovery, if a virtual pharmacokinetic screening can be done, it can curtail down a huge expense incurred by the company. Absorption, Distribution, Metabolism, Excretion (ADME) prediction study not only looks after the economy but also enhances the development of the safest drug at an explosive pace. Thus early information

regarding drug-likeness and pharmacokinetic profile using various computational tools in lead optimization seems to be the best strategic and economic ways to opt for, at an explosive pace. For drug-likeness, Lipinski's Rule of 5 [19] and Veber's rule [20] were considered. in addition, ADME studies generate certain attributes of a bioactive compound, like; Blood-brain barrier (BBB) permeability, Gastrointestinal (GI) permeability, skin permeability, Cytochrome P450 inhibition, and also some structural alerts (if present). Violation of any of the parameters ousted the respective compound from the pool of active compounds.

This *in-silico* study focused on developing and finding some novel protease inhibitors against SARS-CoV-2 main protease, by ligand-based drug designing (LBDD), considering Lopinavir, Ritonavir, and Indinavir as standard compounds, using molecular docking, drug-likeness, and ADME profile.

2. Materials *&* Methods

Owing to the need for the development of new protease inhibitors, Lopinavir, Ritonavir, and Indinavir were taken for pharmacophore modeling on online webserver. A huge number of hits produced which were screened on the basis of root mean square deviation (RMSD) values. Ten such compounds (1-10) were extracted from the Zinc database [21]. Another set of ten compounds (11-20) were developed considering the common pharmacophoric features of the standard compounds.

2.1. Pharmacophore modeling and database searching

For pharmacophore modeling, we used an online webserver named PharmaGist [22]. At first 3 standard molecules, Lopinavir, Ritonavir,Indinavir,and also the co-crystal ligand Boceprevir were drawn in ChemSketch (ACD 2012: Advanced Chemistry Development, Toronto, Ontario, Canada) and converted to mol2 file using Open Babel software[23], and imported into the webserver, which generates the common 3D features of the molecules. The output was further transferred into another webserver called Zinc Pharmer[24], which utilizes millions of chemical compounds in the Zinc database, in search of compounds with nearly the same pharmacophoric features. Around 99 total hits were generated having nearly the same pharmacophoric features with RMSD ranging from 0.32 – 0.75. To ensemble the structural diversity among various molecules generated in the Zinc database, we sorted out a set of 10 compounds having RMSD< 0.55, starting from 0.32. The lower the RMSD cut-off values compared to input pharmacophores, the higher is the similarity between the pharmacophoric features of the compounds and the standards. A recent article also suggested the selection of molecules with a cut-off RMSD around 0.5, matching with the input pharmacophores [25]. Analysis of the common pharmacophoric features of the 3 standards was done using PyMol molecular viewer [26].

2.2. Target (protein) modeling

Crystal structure of Novel Coronavirus Mpro protein was procured from the Protein Data Bank (PDB, http://www.rcsb.org), having PDB entry: 7BRP, with 1.80 Å resolution and complexed with a protease inhibitor, Boceprevir (co-crystal).This gives the rationale for selecting this protein structure, as we are also aiming to develop novel protease inhibitors against SARS-CoV-2. The protein was modeled by MGL Tools 1.5.6 (Molecular Graphics Laboratory, The Scripps Research Institute, La Jolla, USA), run in an HP system with Windows 7, 64-bit OS, 4GB RAM, and a 1.7 GHz processor. To rule out the biasness of the ligand-receptor interactions during docking, the protein molecule was optimized by energy-minimization using Swiss PDB viewer (SPDBV 4.1.0, Swiss Institute of Bioinformatics) [27]. The energy-minimized protein in pdb format was then subjected to Python Molecular viewer. To preclude the interference of water molecules and the co-crystal, the protein was desolvated and made ligand-free. Bond orders were assigned, polar and missing hydrogens were merged including the addition of partial atomic charges by the Gasteiger method. Initially, the protein was in 'pdb' format which was further converted to 'pdbqt' extension by addition of charges (q) and changing the atom type to Autodock4 (t) compatible mode.

2.3. Ligand (small molecules) modeling

All the compounds (1-20) were drawn using ACD ChemSketch freeware. 2D compounds were converted to pdb (protein data bank) using PRODRG web-server [28]. Then using MGL tools, the ligands were prepared by adding Gasteiger partial atomic charges (q) to every atom and changing the atom type to Autodock4 (t). All the ligands were subsequently made into the 'pdbqt' form. The standard compounds were also processed in the same manner.

2.4. Molecular docking and analysis

Molecular docking study was performed in a Windows 7 based HP system, using Autodock Vina [29]. A grid spacing of 1 Å and x, y, z dimensions of 24 points each, had been considered as the primary set-up for the docking procedure. Grid box resolution was set at x, y, and z centers of 24.729, -13.654, and 17.268 respectively. Docking generates 9 significant conformers for each docked compound, of which the least energetic conformer with the best docking pose, considered to be most active and subjected to subsequent studies.

Docking poses of every ligand-receptor complex were visualized using PyMol molecular viewer, and the 2D interactions between the active conformer of the ligand and the protein were observed using BIOVIA Discovery Studio visualizer (BIOVIA, DassaultSystèmes, Discovery Studio Visualizer, 20.1.0, San Diego, CA, USA, 2020). Interactions of the ligands with the active site residues and their optimized binding poses determine the activity of each compound as a potential Mpro inhibitor.

2.5. Drug-likeness studies and ADME profiling

The current study focuses on evaluating drug-likeness parameters of all the compounds, based on 2 established rules namely, Lipinski's Rule of 5and Veber's rule. All the compounds were checked for their drug-likeness by generating data pertaining to molecular weight, number of hydrogen bond donors/acceptors, polar surface area, number of rotatable bonds, partition coefficient, etc. Violation of any of these two rules was kept as rejection criteria, thus proving to be ineffective against SARS-CoV-2 main protease. Drug-likeness studies were performed using an online web-server named, Swiss ADME (Molecular modeling group, Swiss Institute of Bioinformatics, Lausanne, Switzerland) [30].
Computer-aided drug designing (CADD) is the only way to determine the pharmacokinetic studies of every ligand, much early in the lead optimization stage, and this *in-silico* study focuses on few ADME parameters that all the active molecules should qualify to determine the most active compounds amongst all the actives sorted out during the primary screening. In the current research, this ADME profiling is the secondary screening, whereas the primary screening of actives was based on docking score and drug-likeness. ADME and toxicity studies were performed in the Swiss ADME online webserver.

2.6. Molecular Dynamics (MD) simulation study

MD simulation study of the most active compound was performed using LARMD web-server (http://chemyang.ccnu.edu.cn/ccb/server/LARMD/) [31]. The PDB structures of the protein-ligand docked complexes were extracted from PyMol viewer. Then these were fed into the LARMD server, keeping water as an explicit model and MD time as 1 ns (nanosecond) or 1000 ps (picosecond). The safest and most active molecule selected after primary and secondary screening, was initially subjected for 1000 ps MD run, to determine the stability of the ligand-receptor complex, with an intention to increase the MD runtime if the MD trajectories were not stabilized within the 1000 ps run. All the necessary input files were incorporated and the reaction coordinates were portrayed. A definite task name and password were set and then the job was submitted for the simulation study. The most active compound found out by docking, ADME, and drug-likeness properties was directed towards MD simulation using trajectory analysis and energy calculations (MM/PBSA and MM/GBSA). The calculation of binding free energy (ΔG_{bind}MM/PBSA and MM/GBSA) was done using enthalpy or total energy of the system (ΔE), and solvation entropy ($T\Delta S$) of the system. For the MD study, the ligand-receptor interactional binding mode (Int_mod) module was selected for simulation. The force fields of all nonstandard residue(s) were generated. The RMSD and radius of gyration (Rg) of active compound along with the simulation trajectory were calculated and analyzed. RMSD, Rg, and Root mean square fluctuation (RMSF) were analyzed in the same moduleof LARMD using the AMBER16 program.

3. Results and Discussion

This entire computational study encompasses molecular docking, drug-likeness, ADME studies, and MD simulation of the most active lead. 2D structures of all the 20 compounds considered for this *in-silico* study are presented in Table 1.

3.1. Pharmacophore modeling study

The pharmacophore modeling study was performed using Lopinavir, Ritonavir, Indinavir, and Boceprevir molecules. After performing the pharmacophore modeling of 3 standard compounds and the co-crystal ligand Boceprevir, the common pharmacophoric features evolved out of the study include, 2 H-bond donors, 2 aromatic domains, and 1 hydrophobic region, as shown in Figure 1.

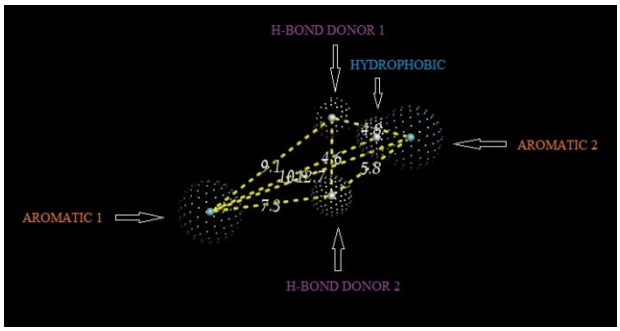


Figure1. Common pharmacophoric features of 3 standard protease inhibitors Lopinavir, Ritonavir, and Indinavir, and the co-crystal ligand Boceprevir, obtained out of pharmacophoric modeling.

3.2. Molecular docking and drug-likeness studies

Docking study generates divergent poses of ligands in which the pose with the best affinity (lowest in terms of binding energy) was chosen as the best pose and

subsequently processed for visualization. In a pool of 20 compounds considered for the study, the first 10 compounds were extracted from the Zinc database and the rest (10 compounds) were designed based on pharmacophoric features of the standards taken for the study. The highly active compounds were identified through multiple screening. Primary screening of all the 20 compounds was done considering the docking score and drug-likeness parameters. Docking study generates a divergent poses of ligand in which the pose with the best affinity (lowest in terms of energy) was chosen as the best pose and subsequently processed for visualization. Docking process was validated prior to initiating the docking study for individual compounds. At first, the co-crystal, Boceprevir was removed from the X-Ray crystallographic structure of the protein (PDB id: 7BRP) and redocked within the active site of the protein. Boceprevir is redocked and used as an external ligand in molecular docking study. 3 standard molecules were also docked and the results are mentioned in Table 2. Docking proves close interaction and good binding affinity of the co-crystal conformers with the active site residues of the protein. RMSD value between the docked active conformer and the X-Ray crystallographic conformation of the co-crystal was observed to be < 1.3Å. According to the literature review, RMSD during docking validation was reported to be around less than 3 Å in one article [32] and less than 1.3 Å in another [33]. The lesser the RMSD, the greater will be the docking accuracy, thus our docking protocol was successfully validated being well within the acceptable range, using Autodock Vina.

Docking reveals that the binding interaction of all the compounds with the active site residues of Mpro protein was well acceptable. The active site analysis divulged the essential composition of amino acids participating with the ligand-binding including Thr-25, His-41, Met-49, Phe-140, Leu-141, Cys-145, His-164, Met-165, Glu-166, Pro-168, Arg-188, Gln-189. The conformer with the least energy of all the docked compounds as generated by Autodock Vina was taken into consideration. Apart from this interaction study, the drug-likeness study was also conducted, to ensure whether the compounds are following the Lipinski's rule and Veber's rule, which a compound requires to comply with to become bioactive. Lipinski's Rule of 5 states that for a compound to be pharmacologically active, it should qualify certain features as set by those scientists, Molecular Weight < 500 dalton, octanol-water partition coefficient (Log P) < 5, Hydrogen (H) bond donors < 5 and H-bond acceptors < 10, while Veber's rule suggests that a compound, before entering into drug discovery pipeline, should possess Rotatable bonds ≤ 10, Topological polar surface area (TPSA) ≤ 140 $Å^2$, and the sum-total of H-bond donors and acceptors ≤ 12. The results fetched from docking analysis in terms of binding energy (Kcal/mol) and the drug-like properties were presented in Table 2.

Considering the structural diversity and the drug-like properties of standard molecules, screening of active compounds was carried out by setting standard protocols. For a compound to be active, there should be 'No violation' in both the Lipinski's rule and Veber's rule, the lower limit set for the summation of H-bond donors and acceptors is not less than 8, and the binding energy (BE) should lie within the range -7.0 to -9.0 Kcal/mol. The results as presented in Table 2 indicate that 10 out of 20 compounds (3, 4, 5, 6, 8, 11, 15, 18, 19, and 20) followed the protocol without any violation.

3.3. In-silico ADME analysis

This virtual assessment was further potentiated by a secondary screening of selected compounds with ADME studies performed in online webserver SwissADME. One of the parameter considered for the ADME profiling is Blood-brain barrier (BBB) permeability, indicates toxic attributes. Though protease inhibitors are not required to penetrate BBB as the central nervous system (CNS) is not their primary site of action, any compounds with the peripheral target, having BBB permeability can cause unwanted CNS-related toxicities [34]. Molecules having high Gastrointestinal (GI) permeability can penetrate the intestinal lining quite easily, making the drug orally bioavailable and it can easily reach the site of action (Coronavirus main protease), crossing the viral membrane. The more negative the Skin permeability (log Kp) value, the lesser the ability of the compounds to penetrate the skin membrane [35]. Metabolic biotransformation is an important phenomenon that safely eliminates the drug from the body, reducing adverse effects due to over-accumulation of the drug inside the body which the Cytochrome P450 (CYP450) isoenzyme family play a major role. Inhibition of CYP450 enzymes may lead to lowering metabolic activity, thus causing drug-drug interactions [36]. Accumulation of the drug or its metabolites may occur due to decreasing in the clearance values leading to several toxicities in the body [37]. To make the study even more extensive, structural alert data was generated, which can readily identify the portion in a molecule that may cause some unwanted effects *in-vivo*. Pan-assay interference compounds (PAINS) [38] alert and a structural alert called Brenk alert are two extraordinary parameters that scrutinize the designed compounds in such a way that any notable area can be readily eliminated at the budding stage of drug discovery. PAINS are mainly the promiscuous compounds that show binding affinity with numerous targets irrespective of the specified target and give rise to false positives and erratic results. Brenk alerts are the structural alerts that may give rise to metabolically unstable and reactive fragments as the compound breaks inside the body, leading to toxicities [39].

All the 10 primarily screened compounds were further tested for the predictive ADME profile to ensure their safety. certain criteria were also kept for a compound to qualify the test. The compounds should possess aqueous solubility (Log S) either soluble or moderately soluble, have high GI absorption, impermeable to BBB, skin permeability (Log Kp) not less than -8 cm/s, should not inhibit at least 3 or more CYP450 isoenzymes, and last but not the least, 'No' PAINS and Brenk alerts. Compounds 3, 4, 5, 11, 19, and 20 were found to be the safest of all, as the data generated from the ADME study lie well within the safe limit. Compound 20, though having moderate aqueous solubility, does not show inhibition of four CYP450 isoenzymes, thus proved to be safer. ADME studies of 10 primary active compounds are shown in Table 3.

3.4. Analysis and interaction study of safe and effective molecules

In this study, 20 compounds were primarily screened by molecular docking and drug-likeness studies. All the actives coming out of this primary screening were again forwarded into a secondary screening of ADME studies. The overall outcome of the study suggests compounds 3, 4, 5, 11, 19, and 20 be the safest and effective lead molecules against SARS-CoV-2 Mpro.

The docking pose within the Mpro active site, the 2D interactions, and the molecular surface view of compound 3 were shown in Figure 2(a), 2(b), and 2(c) respectively. Compound 3 was obtained from the Zinc database with Zinc Id: 08586210. This compound shows strong H-bonding with Glu-166 and pi-pi interactions with Met-165, Glu-166, Cys-145, His-41, and Met-49. Compound 3 shows binding energy of -7.0 Kcal/mol. Figures 2(d), 2(e), and 2(f) represent the docking pose, 2D interactions and molecular surface view respectively of compound 4, which was also fetched from the Zinc database having Zinc Id: 12813413. It shows appreciable interactions with various active site residues of the protein, H-bonding with His-41, and pi-pi stacking with Met-49, Met-165, and Thr-25. Figure 2(g), 2(h), and 2(i) represent the docking pose within the active site, 2D interactions and molecular surface view respectively of compound 5, with Zinc Id: 12813418.

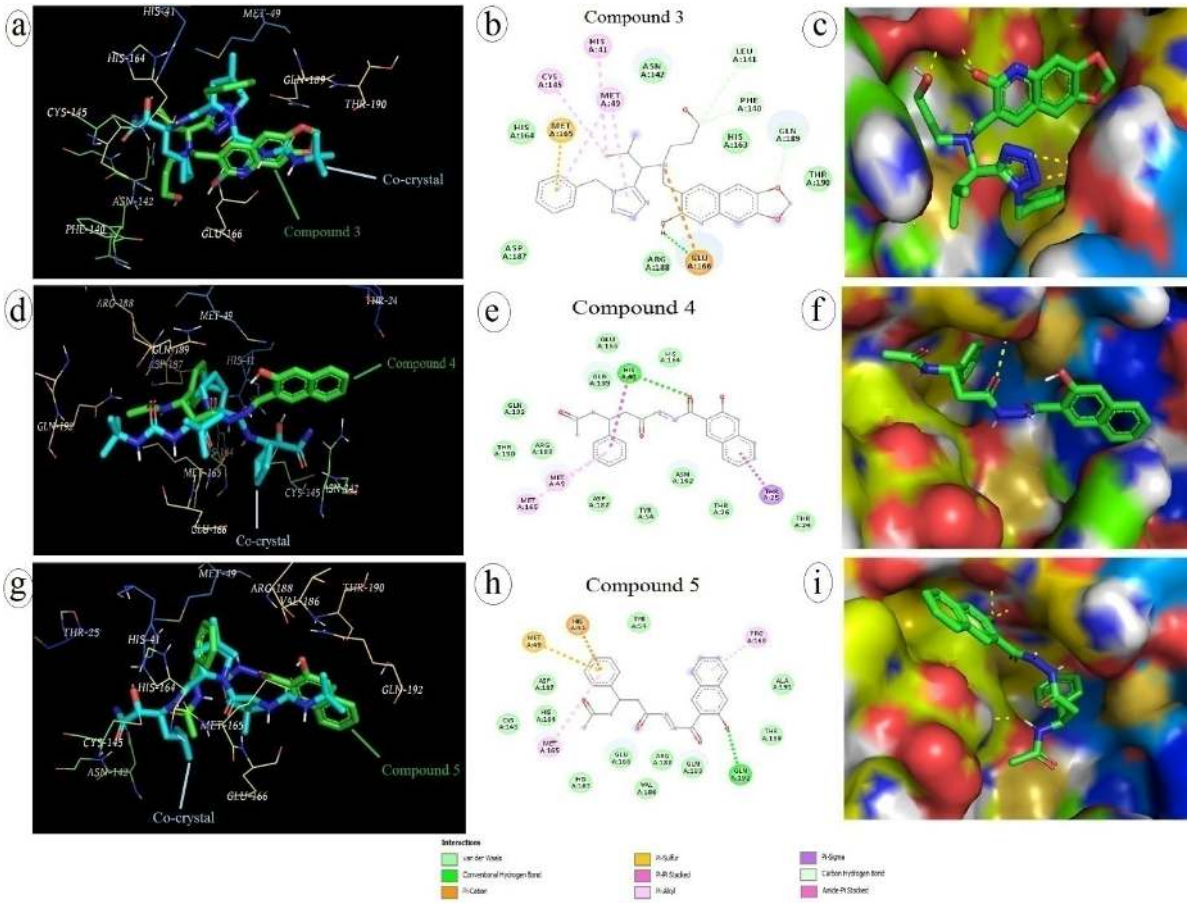


Figure 2. Docking poses of compounds 3, 4, and 5: (**a**) Stereoview of best conformer of compound 3 within the active site of Mpro, superimposed on the native co-crystal; (**b**) 2D interactions of compound 3; (**c**) Molecular surface view of compound 3; (**d**) Stereoview of best conformer of compound 4 within the active site of Mpro, superimposed on the native co-crystal; (**d**) 2D interactions of compound 4; (**f**) Molecular surface view of compound 4; (**g**) Stereoview of best conformer of compound 5 within the active site of Mpro, superimposed on the native co-crystal; (**h**) 2D interactions of compound 5; (**i**) Molecular surface view of compound 5.

This compound also fits well within the binding site domain of the Mpro protein, showing H-bonding with Gln-192 and pi-pi interactions with His-41, Met-49, Met-165, and Pro-168. It is interesting to note that, compounds 4 and 5 are isomeric in nature, 4 being the S isomer and 5 being the R isomer. These 2 compounds differ slightly in their docking binding energies, compound 4 having -7.5 Kcal/mol, whereas -7.3 Kcal/mol was obtained for compound 5S isomer of the compound is slightly more active than the R isomer because of having less energy.

Among the compounds which are not from the Zinc database, created considering the pharmacophoric features of all the standard compounds, compound 11 possesses good binding interactions, including H-bonding with His-41, Met-49, Cys-145, Glu-166, and pi-pi interactions with Thr-25, Met-165. It shows binding energy of -7.1 Kcal/mol. Docking pose within the active site domain, 2D interactions and molecular surface view of compound 11 is presented in Figure 3(a), 3(b) and 3(c) respectively. Another one is compound 19, whose binding energy is -8.2 Kcal/mol. This compound exhibits hydrogen bonding with Glu-166 and pi-pi interactions with Pro-168, Met-49 and Met-165. Docking pose, 2D interactions and molecular surface view of compound 19 were shown in Figure 3(d), 3(e), and 3(f) respectively. Compound 20 exhibits the strongest interaction with the principle amino acid residues of the main protease receptor, having a binding energy of -8.6 Kcal/mol, which is lower than all the 3 standards considered for this study. It shows H-bonding with His-41, Cys-145, Phe-140, Glu-166, and pi-pi stacking interactions with Leu-141, Met-165. Docking pose within the active site, 2D interactions and molecular surface view of compound 20 were shown in Figure 3(g),3(h), and 3(i) respectively. The in-detail scanning of the binding pattern of these pharmacophore-derived compounds with SARS-CoV-2 Mpro accorded that pi-pi stacking and hydrogen bonding interactions are the main arsenals for binding.

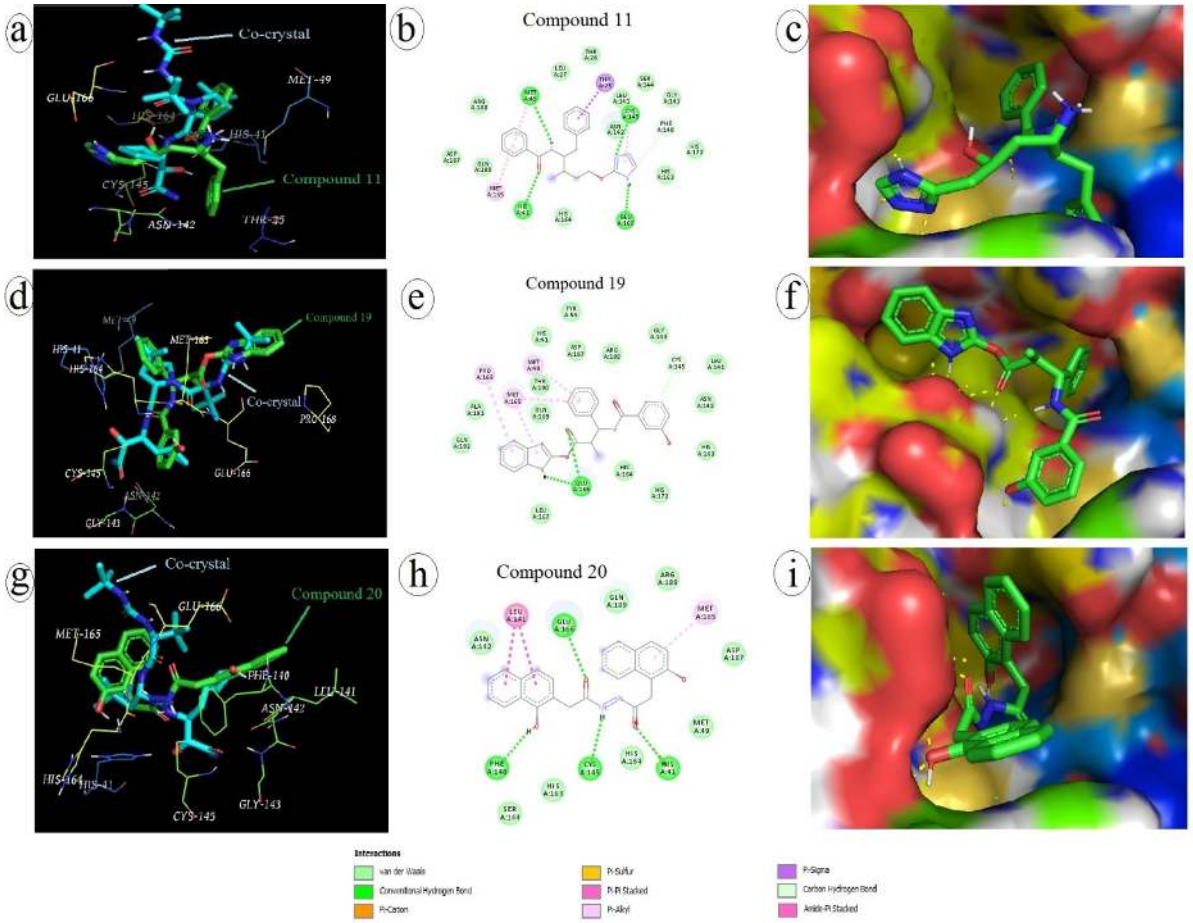


Figure 3. Docking poses of compounds 11, 19, and 20: (**a**) Stereoview of best conformer of compound 11 within the active site of Mpro, superimposed on the native co-crystal; (**b**) 2D interactions of compound 11; (**c**) Molecular surface view of compound 11; (**d**) Stereoview of best conformer of compound 19 within the active site of Mpro, superimposed on the native co-crystal; (**d**) 2D interactions of compound 19; (**f**) Molecular surface view of compound 19; (**g**) Stereoview of best conformer of compound 20 within the active site of Mpro, superimposed on the native co-crystal; (**h**) 2D interactions of compound 20; (**i**) Molecular surface view of compound 20.

3.5. MD simulation analysis of active compound

After a detailed screening of the active compounds by docking, drug-likeness and ADME studies, 6 compounds were found to be highly active and safer lead molecules targeting Coronavirus Mpro. Compound 20 ranked highest in the list possessing best binding affinity, safety and drug-likeness properties. Hence, compound 20 was selected for Molecular Dynamics study and same analysis was carried out on Coronavirus Mpro protein. Using conventional molecular dynamics, the binding free energy (ΔG_{bind}) calculations were carried out by MM/PBSA (Molecular Mechanics/ Poisson Boltzmann Surface Area) and MM/GBSA (Molecular Mechanics/ Generalized Born Surface Area) methods. These calculations give an estimate of overall binding energy between ligand and the receptor and thereby predicting the stability of the complex by a physics-based empirical scoring function. For compound 20, binding free energy (ΔG_{bind}) as calculated by MM/PBSA was -9.94 Kcal/mol and by MM/GBSA was -16.15 Kcal/mol. The 3 standard compounds were also subjected to binding free energy calculation. The ΔG_{bind} as calculated by MM/PBSA for the compounds Lopinavir, Ritonavir and Indinavir was found to be -17.64 Kcal/mol, -21.78 Kcal/mol and -9.72 Kcal/mol respectively. However, for these three compounds, , the values obtained from determining the ΔG_{bind} through MM/GBSA were -14.70 Kcal/mol, -23.30 Kcal/mol and -4.92 Kcal/mol. While comparing the binding energies, Ritonavir is holding better than the other two standards as well as compound 20, but the pose with which the compound 20 is found within the binding domain is well acceptable. Moreover, in comparison with the other standard, Indinavir, compound 20 is having a greater ΔG_{bind} value.

In order to assess the stability and the dynamic behavior of the ligand-receptor complex of compound 20 on Mpro protein, RMSDof ligand-receptor complex, Rg and RMSF were determined. MD simulation study was performed for 1 ns or 1000 ps, in water solvated state with Int-mode, using LARMD webserver. The stability of MD trajectory was confirmed using RMSD and Rg calculations. It was observed that the MD trajectory is almost stable without any noticeable major deviations in RMSD, Rg, and RMSF, in 1000 ps MD runtime, thereby there is no requirement for further increment of MD time. RMSD of protein-Cα atoms and the compound 20 (ligand) lies well within the range of 0-1.5 Å and 0 - 2.0 Å respectively, which is well acceptable for small molecule protease inhibitors. The RMSD curves of receptor and the ligand are nearly superposed, thereby predicting the stability of the complex, and the ligand is not dissociated throughout 1 ns MD simulation analysis. Moderate and stable Rg fluctuations were observed between 24.9 – 25.4 Å. Recognition of the ligand along with the migration process potentially influences the internal atomic fluctuations of the Mpro protein. Thus RMSF of each residue was observed by means of RMSF plot. Regular and routine fluctuations of residues were observed for Mpro-compound 20 complex showing fluctuations between 2 – 12 Å, fluctuations being the greatest between 200 – 400 residue points. Structural flexibility of the Coronavirus Mpro protein and the stability of compound 20 within its active site were proved and validated using MD simulation analysis. This investigation proved that among all the compounds selected for this study, compound 20 is the most active and stable within the binding cavity of the Mpro protein. RMSD, Rg and RMSF plots for compound 20 is depicted in figure 4.

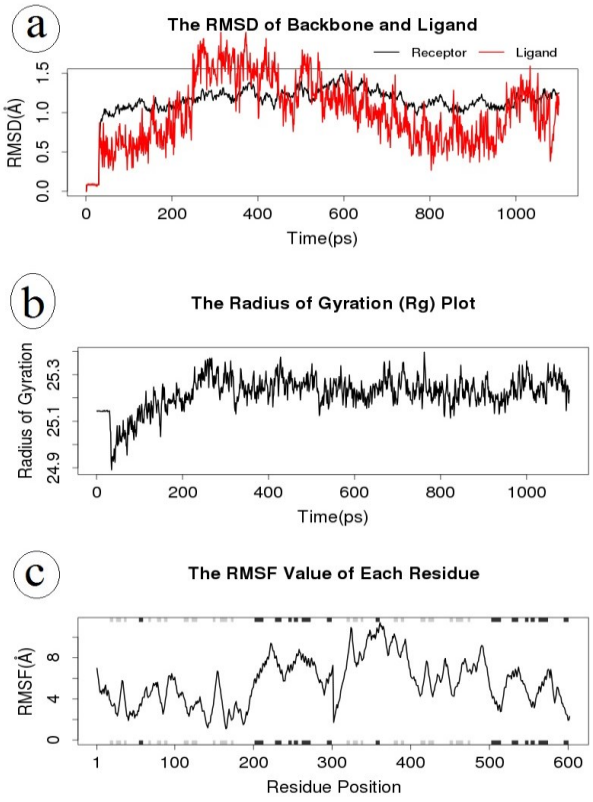


Figure 4. MD simulation study for compound 20 complexed with Mpro protein: (**a**) RMSD plot of the ligand-receptor complex; (**b**) Rg plot of the ligand-receptor complex; (b) RMSF plot of the ligand receptor complex.

4. Conclusion

The drug discovery process of SARS-CoV-2 infection is still in its budding stage. Target specific lead identification is yet to be done. the majority of present drugs for COVID-19 preferentially act on viral main protease, therefore with the strategic application of computational study, like molecular docking, drug likeness, and ADME profile check, we have developed

few protease inhibitors based on the pharmacophoric features of existing antivirals. In order to enrich the basal knowledge during this long process of investigation, computational chemistry has been proved to be the best alternative. The docking study as performed here reveals that among 20 compounds, the active site affinity of compound 20 outweighs all the three standard compounds, which clearly delineates its proper accommodation within binding pocket of Mpro. Apart from this docking studies, compound 20 also possess a good number of drug like properties. Further ADME study made this compound a lead of interest to explore further. Other than compound 20, compounds like 3, 4, 5, 11, and 19 also found to be effective as Mpro inhibitors. Further computational study includes the Molecular-Dynamic (MD) simulation of compound 20 to ensure no or lesser variations throughout the simulation analysis and the stability of the ligand-target complex. Data obtained out of the study will definitely serve as a hypothetical guideline to optimize the lead as effective anti-SARS-CoV-2 agent.

Acknowledgements

The authors declare their deep and sincere gratitude to the Principal of Gupta College of Technological Sciences, West Bengal, India and to the Principal, faculty staffs and all the members of Pharmaceutical Chemistry department of Acharya & BM Reddy College of Pharmacy, Karnataka, India, for their widest support in successful accomplishment of this research work.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest..

Funding/ Support

None.

Authors' ORCIDs

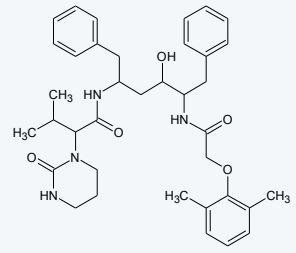
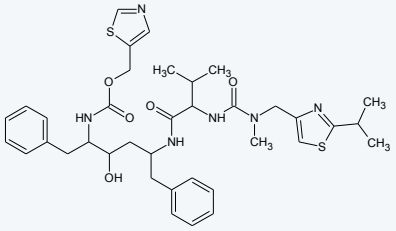
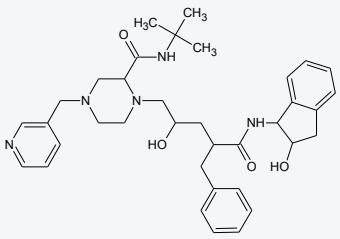
Arka Das:
https://orcid.org/0000-0002-8862-7515

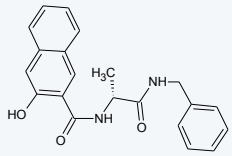
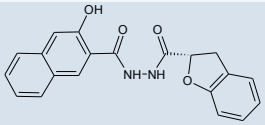
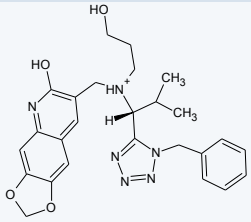
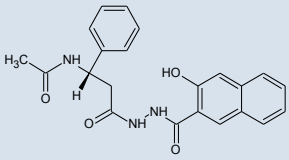
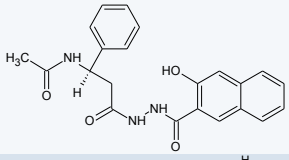
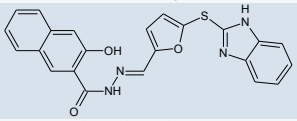
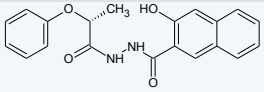
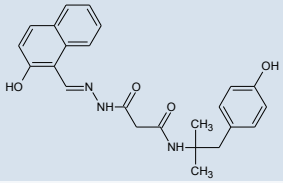
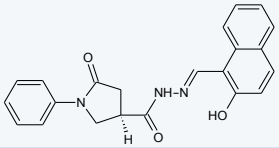
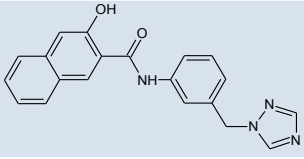
References

1. Lu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J Med Virol* (2020) 92 (4): 401-402. doi: 10.1002/jmv.25678.
2. Xu X, Yu C, Qu J, Zhang L, Jiang S, Huang D, Chen B, Zhang Z, Guan W, Ling Z, Jiang R. Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2. Eur J Nucl Med Mol Imaging (2020) 47: 1275–1280. doi: 10.1007/s00259-020-04735-9
3. Pal M, Berhanu G, Desalegn C, Kandi V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. Cureus (2020) 12 (3): e742. doi: 10.7759%2Fcureus.7423.
4. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed. Ateneo Parmense (2020) 91 (1): 157-160. doi: 10.23750/abm.v91i1.9397.
5. Sun P, Qie S, Liu Z, Ren J, Li K, Xi J (2020) Clinical characteristics of hospitalized patients with SARS- CoV- 2 infection: a single arm meta- analysis. J Med Virol (2020) 92 (6): 612-617. doi: 10.1002/jmv.25735.
6. Wilkerson RG, Adler JD, Shah NG, Brown R. Silent hypoxia: A harbinger of clinical deterioration in patients with COVID-19. Am J Emerg Med (2020). doi: 10.1016/j.ajem.2020.05.044
7. Coronavirus disease (COVID-19) Dashboard. World Health Organization (WHO), Geneva, Switzerland. https://covid19.who.int/ . Accessed 15 July 2020.
8. Gorbalenya AE, Baker SC, Baric R, Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW, Penzar D. Severe acute respiratory syndrome-related coronavirus: The species and its viruses–a statement of the Coronavirus Study Group. biorxiv (2020). doi: 10.1101/2020.02.07.937862.
9. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. Int J Antimicrob Agents (2020) 55 (3): 105924. doi: 10.1016/j.ijantimicag.2020.105924.
10. Wang J. Fast identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study. J Chem Inf Model (2020) 60 (6): 3277-3286. doi: 10.1021/acs.jcim.0c00179.
11. Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. Cell discov (2020). doi: 10.1038/s41421-020-0153-3.
12. Ortega JT, Serrano ML, Pujol FH, Rangel HR. Unrevealing sequence and structural features of novel coronavirus using in silico approaches: The main protease as molecular target. EXCLI J (2020) 19: 400-409. doi: 10.17179%2Fexcli2020-1189.
13. Prajapat M, Sarma P, Shekhar N, Avti P, Sinha S, Kaur H, Kumar S, Bhattacharyya A, Kumar H, Bansal S, Medhi B. Drug targets for corona virus: A systematic review. Indian J Pharmacol (2020) 52 (1): 56-65. doi: 10.4103/ijp.IJP_115_20.
14. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, Duan Y. Structure of Mpro from COVID-19 virus and discovery of its inhibitors. bioRxiv (2020). doi: 10.1101/2020.02.26.964882.
15. Shah B, Modi P, Sagar SR. In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. Life Sci (2020) 252: 117652. doi: 10.1016/j.lfs.2020.117652.
16. Yang SY. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discov Today (2020) 15 (11-12): 444-450. doi: 10.1016/j.drudis.2010.03.013.
17. Yuriev E, Agostino M, Ramsland PA. Challenges and advances in computational docking: 2009 in review. J Mol Recognit (2011) 24 (2): 149-164. doi: 10.1002/jmr.1077.
18. Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010–2011 in review. J Mol Recognit (2013) 26 (5): 215-239. doi: 10.1002/jmr.2266.
19. 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. doi: 10.1016/s0169-409x(00)00129-0.
20. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem (2002) 45 (12): 2615-2623. doi: 10.1021/jm020017n.
21. Irwin JJ, Shoichet BK. ZINC− a free database of commercially available compounds for virtual screening. J Chem Inf Model (2005) 45 (1): 177-182. doi: 10.1021/ci049714+
22. Schneidman-Duhovny D, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. Nucleic Acids Res (2008) 36: W223-228. doi: 10.1093/nar/gkn187.

23. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminformatics (2011) 3 (1): 33. doi: 10.1186/1758-2946-3-33
24. Koes DR, Camacho CJ. ZINCPharmer: pharmacophore search of the ZINC database. Nucleic Acids Res (2012) 40 (W1): W409-414. doi: 10.1093/nar/gks378.
25. Panda SK, Saxena S, Gupta PS, Rana MK. Inhibitors of Plasmepsin X Plasmodium falciparum: Structure-based pharmacophore generation and molecular dynamics simulation. J Mol Liquids (2021):116851. doi: 10.1016/j.molliq.2021.116851.
26. DeLano WL. Pymol: An open-source molecular graphics tool. CCP4 Newsletter on protein crystallography (2002) 40 (1): 82-92.
27. 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.
28. Van Aalten DM, Bywater R, Findlay JB, Hendlich M, Hooft RW, Vriend G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J Computer Aid Mol Des (1996) 10 (3): 255-262. doi: 10.1007/BF00355047.
29. Trott O, Olson A.J. AutoDock Vina: Improving the speed and accuracy of docking. J Comput Chem (2010) 31: 455-461. doi: 10.1002/jcc.21334.
30. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep (2017) 7: 42717. doi: 10.1038/srep42717.
31. Yang, J. F., Wang, F., Chen, Y. Z., Hao, G. F., & Yang, G. F. LARMD: Integration of bioinformatic resources to profile ligand-driven protein dynamics with a case on the activation of estrogen receptor. Brief Bioinformatics (2019) 1-13. doi: 10.1093/bib/bbz141.
32. Moitessier N, Westhof E, Hanessian S. Docking of aminoglycosides to hydrated and flexible RNA. J Med Chem (2006) 49(3):1023-1033. doi:10.1021/jm0508437.
33. Das A, Nandi S, Das K, Banerjee S. Virtual screening of few novel Sulindac derivatives as multi-targeted agents. Drug Discov (2020) 14(33):33-43.
34. Fromm M F. P-glycoprotein: A defense mechanism limiting oral bioavailability and CNS accumulation of drugs. Int. J. Clin Pharmacol Ther (2000) 38 (2): 69-74. doi: 10.5414/cpp38069
35. Potts R O and Guy R H. Predicting Skin Permeability. Pharm. Res. (1992) 9 (5): 663-69.
36. Hollenberg P F. Characteristics and common properties of inhibitors, inducers, and activators of CYP enzymes. Drug Metab Rev (2002) 34 (1-2): 17-35. doi: 10.1081/dmr-120001387.
37. Kirchmair J. Kirchmair J, Göller AH, Lang D, Kunze J, Testa B, Wilson ID, Glen RC, Schneider G. Predicting drug metabolism: experiment and/or computation? Nature Rev. Drug Discov (2015) 14 (6): 387-404. doi: 10.1038/nrd4581.
38. Baell J B and Holloway G A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem (2010) 53 (7): 2719-2740. doi: 10.1021/jm901137j.
39. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, Wyatt PG. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. Chem Med Chem (2008) 3 (3): 435-444. doi: 10.1002/cmdc.200700139.

Table 1. Ligands considered for in-silico analysis

Standard Compounds	Chemical Structures
Standard 1 (Lopinavir)	
Standard 2 (Ritonavir)	
Standard 3 (Indinavir)	

Compounds from Zinc Database	Chemical Structures
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

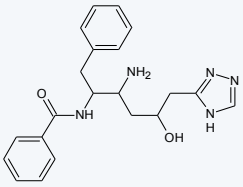
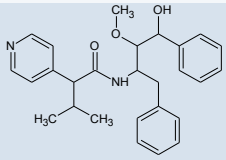
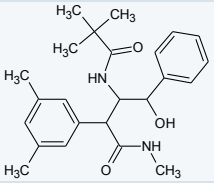
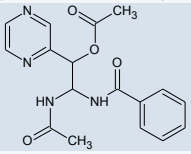
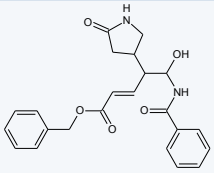
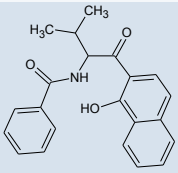
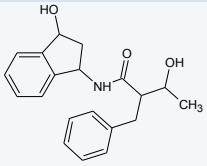
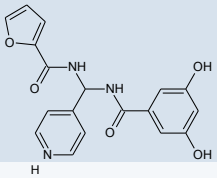
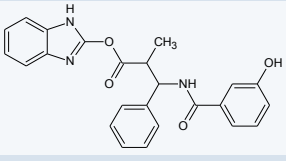
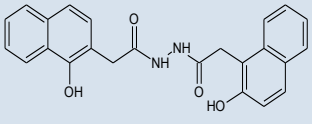
Designed Compounds	Chemical Structures
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

Table 2. Docking output and drug-likeness study of 20 compounds

Docking Score		Drug-Likeness Parameters							
Compound	BE* Kcal/mol	Mol wt (g/mol)	Consensus Log P	H-bond donors	H-bond acceptors	Lipinski's Rule	Rotatable bonds	TPSA[a]	Veber's Rule
Boceprevir (Co-crystal Ligand)	-6.4	519.68	2.09	4	5	1 violation (MW>500)	14	150.70	2 Violation (Rot bonds>10and TPSA> 140 Å^2)
Standard 1 (Lopinavir)	-7.5	628.80	4.53	4	5	1 violation (MW>500)	15	120.00	1 violation (Rot bonds>10)
Standard 2 (Ritonavir)	-7.0	720.94	5.03	4	7	2 violation (MW>500 Log P >5)	18	202.26	2 violation (TPSA>140 Rot bonds>10)
Standard 3 (Indinavir)	-8.4	613.79	2.78	4	7	1 violation (MW>500)	12	118.03	1 violation (Rot bonds>10)
1	-7.7	348.40	3.11	3	3	No violation	7	78.43	No violation
2	-7.5	348.35	2.63	3	4	No violation	5	87.66	No violation
3	-7.0	491.56	2.04	3	8	No violation	10	119.85	No violation
4	-7.5	391.42	2.36	4	4	No violation	9	107.53	No violation
5	-7.3	391.42	2.48	4	4	No violation	9	107.53	No violation
6	-7.9	428.46	3.91	3	5	No violation	6	128.81	No violation
7	-7.8	350.37	2.80	3	4	No violation	7	87.66	No violation
8	-7.1	419.47	3.34	4	5	No violation	9	111.02	No violation
9	-7.3	373.40	2.72	2	4	No violation	5	82.00	No violation
10	-7.6	344.37	2.83	2	4	No violation	5	80.04	No violation
11	-7.1	379.46	1.62	4	5	No violation	10	116.92	No violation
12	-6.8	432.55	3.80	2	4	No violation	11	71.45	1 violation (Rot bonds>10)
13	-6.9	396.52	3.38	3	3	No violation	9	78.43	No violation
14	-6.7	342.35	0.78	2	6	No violation	9	110.28	No violation
15	-7.7	408.45	1.86	3	5	No violation	10	104.73	No violation
16	-6.8	347.41	3.96	2	3	No violation	6	66.40	No violation
17	-7.1	325.40	2.41	3	3	No violation	6	69.56	No violation
18	-7.9	353.33	1.21	4	6	No violation	7	124.69	No violation
19	-8.4	415.44	3.50	3	5	No violation	8	104.31	No violation
20	-8.6	400.43	3.16	4	4	No violation	7	98.66	No violation

Table 3. ADME profiling of 10 primary active compounds

Compounds	Log S* (ESOL)	GI absorption[ƙ]	BBB permeant[ϖ]	Log Kp[α]	CYP450 inhibition					Medicinal Chemistry alerts	
					1A2	2C19	2C9	2D6	3A4	PAINS[µ]	Brenk[€]
3	M. Sol	High	No	-6.73	No	No	Yes	No	Yes	0	0
4	Sol	High	No	-6.61	No	No	Yes	Yes	No	0	0
5	Sol	High	No	-6.61	No	No	Yes	Yes	No	0	0
6	P. Sol	Low	No	-4.67	No	Yes	Yes	Yes	No	1	1
8	M. Sol	High	No	-5.61	No	No	Yes	No	No	1	2
11	Sol	High	No	-7.46	No	No	No	Yes	Yes	0	0
15	Sol	High	No	-7.48	No	No	No	No	No	0	2
18	Sol	High	No	-7.45	No	No	No	Yes	No	0	1
19	M. Sol	High	No	-5.84	No	No	Yes	Yes	No	0	0
20	M. Sol	High	No	-5.88	No	No	No	Yes	No	0	0