Finding potent inhibitors against SARS-CoV-2 main protease through virtual screening, ADMET, and molecular dynamic simulation studies

Rajarshi Roy, Md Fulbabu Sk, Nisha Amarnath Jonniya, Sayan Poddar, Parimal Kar*

Discipline of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Khandwa Road, MP 453552, India

*Corresponding author: Parimal Kar| Email: parimal@iiti.ac.in| Phone: +91 731 2438700 (Ext. 550)

Abstract

Coronavirus disease 2019 (COVID-19) is caused by SARS-CoV-2, and more than 15 million people across the globe have been infected with it. Currently, no antiviral drug or vaccine is available to treat this disease. This underscores an urgent need for developing a drug against SARS-CoV-2. The main protease or chymotrypsin-like cysteine protease ($3CL^{pro}$) of SARS-CoV-2 is considered an essential protein for maintaining the viral life cycle and, therefore, a potential target for drug development. Recently, Ton and co-workers[1] have identified the top 1,000 potential ligands for $3CL^{pro}$ by screening 1.3 billion compounds from the ZINC15 library. In the current study, we have further screened these 1000 compounds using structure-based virtual screening and identified 9 compounds having glide score ~ -11.0 kcal/mol or less. Top 5 screened inhibitors are found to display good pharmacological profiles revealing better absorption, proper permeability across the membrane, uniform distribution, and non-toxic. Further, we conducted 100 ns molecular dynamics simulation of the top 5 docked complexes, and our simulation revealed that one of the ligands dissociates from the binding pocket. However, four molecules, namely ZINC001062406583, ZINC000571366263, ZINC000452260308, and ZINC000680430230, displayed stable binding against $3CL^{pro}$. The molecular mechanics generalized Born surface area (MM/GBSA) method is used to calculate the binding free energy of the four ligands and found that ZINC000452260308 is more potent (ΔG = -14.31 kcal/mol) compared to the other three compounds. For all cases, the binding is mainly driven by the van der Waals interactions. This new compound may have a great potential as a lead molecule for the development of new antiviral drug to fight against COVID-19.

Introduction

The current global crisis due to Coronavirus-2019 (COVID-19) has almost brought healthy life to a standstill in most parts of the world, infecting more than 10 million people in just five months. The outbreak has been declared a global pandemic by the World Health Organization (WHO) on March 11, 2020 (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). Its economic consequences are impacting the whole world with unprecedented speed and severity since many countries have to impose lockdowns and close borders to contain the spread of the virus because, to date, there is no efficient and specific antiviral treatment for COVID-19[2, 3]. So far, more than 0.6 million deaths due to COVID-19 have been reported worldwide. In India, more than 1.2 million people have been infected by the novel coronavirus, including more than 30000 deaths (https://www.mohfw.gov.in/). COVID-19 is caused by a new pathogen, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus, which is from the family of the *beta-corona* virus with a positive-sense single-stranded RNA genome [4-11]. The genome size of SARS-CoV-2 is large, which ranges from approximately 27 to 37 kilobases. SARS-CoV-2, emerged from the Wuhan city of China in late 2019, has drawn considerable attention from the scientific community globally, due to its severity and rapid spread of disease [12-17]. The infection with the new pathogenic SARS-CoV-2 can result in long term reduction in lung function, arrhythmia, and death.

Main protease (M^{pro}), also known as 3-C like protease ($3CL^{pro}$) received significant attention due to its vital function in post-translational processing of replicase polyproteins [9-13, 18-20]. The enzymatic activity of $3CL^{pro}$ leads to the processing of viral new polyproteins, i.e., it digests the specific peptide bonds in ten conserved glutamine residues in the c-terminal region of the same[21-23]. The SARS-COV-2 $3CL^{pro}$ has high structural and sequence similarity (~96%) to that of SARS-CoV $3CL^{pro}$ and ~306 residues long, including three domains, folded into helices and β-strands[4]. The monomer of SARS-COV-2 3CLpro consists of N-terminal domain-I (residues 1-101), and domain-II (residues 102-184) consists of an anti-parallel beta-barrel and C-terminal α-helical domain-III (residues 201-303) is required for enzymatic activity[24]. This protein shares a similar conformation with cysteine protease with an active site lacking the third catalytic residue; it comprises a catalytic dyad, His41 and Cys145 [7, 10, 15] are placed at the junction of domain-I and domain-II.

Recently, Jin and his coworker solved the X-ray crystal structure of COVID-19 $3CL^{pro}$ in complex with an N3 inhibitor (PDB ID 6LU7)[15]. This structure can be used as a potential screening tool for probable inhibitory molecules in the *in-silico* method, which further could

be validated using different computational and experimental methods[1]. As this protease lack enough homology to any human protease, the use of different inhibitors of the same can be safe and very less harmful to the body[14]. Therefore, identification of the inhibitory molecules of this protease is an urgent need to overcome the number of infections and reduce the inflammatory response of infected patients. Currently, remdesivir is one of the most favourable drugs which initially developed to treat Ebola-infected patients[25]. Already several studies have done by the repurposing of different FDA approve drugs in all possible targets like N-protein, 2'-O-ribose methyltransferase, envelope protein ion channel, spike protein as well as the $3CL^{pro}$ protease of COVID-19[26-30]. In case of $3CL^{pro}$, several approved drugs such as disomin, hesperidine, dihydroergocristina, ditercalinium, teniposide, velpastasvir, saquinavir, lopinavir, oseltamivir, ritonavir was analysed using several *in-silico* techniques which can further help to develop potent inhibitors against COVID-19[31, 32]. Apart from the approved drugs, some drug-like molecule shows excellent potential to inhibit the mechanism of the protease like ZINC000000702323 and ZINC000012481889 from the study of Elmezayen *et al.* .[33] as well as the several molecules by the rapid screening method of Ton *et al.* .[1]. Later one studies 1.3 billion drugs like molecules using their newly developed deep learning method, i.e. Deep Docking (DD) and narrowed that number to 1000 potential molecules to further analysed against COVID-19 protease[1].

In this present work, we illustrated an in-silico approach for 1000 drug-like molecules which has been successfully screened across 1.3 billion molecules by Ton *et al.* .[1]. Here we used virtual screening and ADMET profile analyses to screen these molecules in more significant details further. Further, we extended our study by using molecular dynamics simulation in the top few molecules obtained from the screening and also use molecular mechanics generalized Born surface area (MM/GBSA) method to estimate the binding free energies between drug-like molecules and protease. Therefore, our study aimed to summarise the potential pharmacological relationship as well as structural details along with the binding mechanism of selected drug-like molecules against the COVID-19 protease target.

Materials and Method

Selection and the preparation of target protein:

In this study, we desire to design potential inhibitors of the main protease of COVID-19, i.e. $3CL^{pro}$. For that, the crystal structure of $3CL^{pro}$ was retrieved from the protein data bank[34] solvated in complex with the N3 inhibitor (synthetic construct) having a resolution around 2.16 Å (PDB ID: 6LU7)[35]. Preparation of the receptor protein, protein preparation wizard[36, 37] of

Schrödinger software with the help of maestro platform[38] of the same. Missing hydrogen atoms were added, and the crystallographic water beyond 3 Å from the protease was removed. The structure was refined in the presence of sample water orientations and optimized at pH 7.0 with the help of PROPKA software[39]. Finally, under the default settings and the presence of the OPLS3[40] force field, the protein was minimized.

Preparation of the ligand database:

Our study is focused on the further refinement of the top 1000 hit obtained from the 1.3 billion compounds screened by the deep docking method by the Cherkasov group[1]. Those molecules were downloaded in SDF format, which is freely available in their server for further use. Ligands were imported in the maestro portal[38] and subjected to the Ligprep module[41] to generate all the ligands conformer along with their tautomeric combination under pH 7.0 $\pm$ 2.0. The ligands were minimized and optimized using the OPLS3[40] force field after the addition of hydrogens. All ligand conformers were used in the virtual screening method which is discussed in the next section.

Virtual screening:

All 1000 molecules were screened against the $3CL^{pro}$ through the virtual screening workflow under Glide module[42-45] of the Schrödinger suite, which is a grid-based ligand docking procedure. This workflow includes the preparation of ligand (discussed in the previous section), filtering using relevant pharmacological parameters, and three different protocol of docking. Before this, the receptor grid was generated using Glide[45] taking N3 inhibitor of the protease in centre Å with a 12 cubic space around it. Three tiers of virtual screening are high throughput virtual screening (HTVS), standard precession (SP) and extra precession (XP) was used sequentially to achieve a set of potential lead molecules with high accuracy. Top 50 % docked complex from HTVS were used in the SP docking method. Again top 20 % of SP docking were used as input to the XP docking procedure. Finally, 10 % of top lead molecules were kept after XP docking method. The remaining parameter of the workflow was kept on default. Top lead molecules were ranked according to Glide-XP score and selected for the further analysis.

Pharmacokinetic and toxicological profiling:

QikProp[46] module of the Schrödinger suite was used to compute the ADME (Absorption, distribution, metabolism and excretion) properties of our top lead compounds from the virtual screening method. It predicts physically significant and pharmaceutically relevant properties

of organic molecules. This module predicts 35 principal properties such as CNS activity, brain/blood penetration, Madin-Darby Canine Kidney cell permeability, % human oral absorption etc which are used a filter for drug candidate molecules in the drug design program. We also used pkCSM server[47] to validate and estimate various other parameters related to drug-likeness and ADME analysis. Bioactivity of all lead molecules was analysed by using the molinspiration webserver (https://www.molinspiration.com). Estimation of the toxicity of drug-like molecule is one of the critical factors in the field of drug development[48]. Hepatotoxicity, Carcinogenicity, Mutagenicity, and Cytotoxicity of the selected compounds were predicted by using ProTox-II webserver[49]. Predicated results were also revalidated using the pkCSM server toxicity profiling[47]. Finally, comparing all the parameter of ADMET analysis from several servers, we filter the lead molecules of Glide-XP and further processed to the molecular dynamic's simulation.

Molecular dynamics (MD) simulation:

All complex structures which were filter from the virtual screening and ADMET analysis were loaded into the *leap* module of Ambertool19[50], and the adequate number of ions were added to neutralize the systems. A periodic octahedron TIP3P water box[51] was used to solvate the systems which have 10 Å buffer distance from all the directions. A Linux based GPU workstation (GPU-GTX 2080-Ti card) was used to simulated all the systems using *pmemd.cuda* module of Assisted Model Building with Energy Refinement (AMBER) suite[50]. We used amber ff14SB force field[52] and the updated generalized Amber force field (GAFF2)[53] for simulate protein and ligand molecule, respectively. Bond lengths having hydrogen atoms were kept fixed using the SHAKE algorithm[54], and particle mesh Ewald summation (PME)[55] was also used to compute long-range interaction with a non-bonded cut off of 10 Å. Timestep, in all cases, was fixed to 2.0 fs. We followed the same simulation protocol from our previous work on the potency of drug binding of $3CL^{pro}$ of COVID 19[56]. All the simulations were carried out up to 100 ns under the NPT ensemble generating a total of 10,000 snapshots. Further, these trajectories were used to analyze root means square deviation, the radius of gyration, solvent accessible surface area (SASA), protein-ligand hydrogen bonds using *Cpptraj* module[57] of AMBER18[50].

Binding free energy analysis using MM-GBSA scheme:

Determination of binding free energy of the protein-ligand system can be done using molecular mechanics generalized Born surface area (MM-GBSA) scheme, which widely used

techniques[58-63]. Free energy is comprised of three critical factors i.e. gas-phase molecular mechanics energy (ΔE_{MM}), desolvation free energy (ΔG_{solv}) and the configuration entropy ($T\Delta S$) related by the following formula;

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (1)$$

Detail description of the MM-GBSA method was discussed in our previous work[63-70] and here we used the same. Herein, for all the energy estimations, we used 2500 frames from the last 50 ns trajectories and MMPBSA.py script implemented in Ambertool19[50]. Configurational entropy was calculated using normal mode analysis method using the same script and due high computational cost, only 25 frames are considered from the same region which is used in the free energy calculation. Contribution from each amino acid in the binding free energy was also performed using the MM-GBSA pair-wise decomposition scheme.

Result and Discussion

The steps involved in our study are summarized schematically in Figure 2. We used the top 1000 molecules screened by Cherkasov group[1] form 1.3 billion compounds of ZINC15 database[71] and further filtered by virtual screening, several pharmacological studies and molecular dynamics. The best lead was selected based on MM-GBSA calculated using the MD simulation trajectories.

Virtual screening of compounds against COVID-19 protease:

Before the virtual screening, we used the QikProp module[46] and Lipinski rule of 5 for filtering these molecules based on drug-likeness properties. After screening, a total of 938 molecules were obtained and further docked to the protease using HTVS protocol. Subsequently, we reduced the number of molecules using SP and XP docking algorithm in the above-discussed pattern. Finally, we obtained the top nine molecules, results of the XP protocol, which are shown in Figure 1 along with their chemical formula, ZINC IDs and molecular weights. G scores, found from the Glide XP for top nine ligands, ZINC000541677852, ZINC001062406583, ZINC000571366263, ZINC000452260308, ZINC000679651603, ZINC000680430230, ZINC000527019428, ZINC000544491494 are listed in the Table 1, along with the other important terms. XP gives us better accuracy and precision compare to SP as it uses an anchor and grows sampling method[72]. This accuracy comes with a cost in docking time i.e. approximately 2 min/compound which is significantly lower in case of SP algorithm (10 sec/compound). All the nine listed compounds have G score more than equals to -11 kcal/mol. Among these compounds, ZINC000541677852 and ZINC001062406583 have --

11.57 kcal/mol and -11.56 kcal/mol, which are the best lead found through the virtual screening workflow. Through our virtual screening set-up, our selected lead molecules accounts less than 1% percent of the input database.

Physicochemical and pharmacokinetic features of the lead molecules:

We calculated the ADME properties i.e. absorption, distribution, metabolism and excretion of the top nine lead molecules obtained from the virtual screening workflow by the use of QikProp module[46]. Significant properties of the ADME analysis, i.e. octanol/water partition coefficient, aqueous solubility, Caco-2 cell permeability, IC 50 value for the blockage of HERG K+ channels, MDCK cell permeability, and human oral absorption are listed in Table 2. All the parameters obtained from QikPRop are also listed in the supplementary information (Table S1-S3). We further verified the ADMET properties using the pkCSM software by studying absorption, distribution, metabolism, excretion, and toxicity properties individually, which are listed in the supporting information (Table S4 -S7). Molecular weights of all lead molecules are below 500 Dalton, as listed in Table 1, which is an indication of better absorption than the higher molecular weight molecules[73]. Predicted water-octanol partition co-efficient for all the molecules are in between the acceptable range, which indicates an excellent permeability of those across the cell membranes. Caco2 permeability is a well-established in-silico technique to screen oral absorption as well as the transport mechanism of the drugs, where the permeability of a compound is checked across Caco2 monolayer cell[74]. All nine molecules have a good score in this analysis, whereas ZINC000679651603 shows the best permeability.

Analysis of the distribution of all lead molecules involves the volume of distribution (VDss), blood-brain barrier (BBB) permeability, central nervous system (CNS) permeability shows all molecules are in the good pharmacological reference ranges (Table S4). Among all, ZINC001062406583 shows the lowest BBB and CNS permeability, which is a good indication of a functioning drug. Metabolism studies show that ZINC001062406583 is a non-substrate of any possible cytochrome P450 isoform i.e. CYP2D6, CYP3A4, CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4 indicating a proper metabolism (see Table S5). The rest of them also show almost no interaction except for 1-2 isoforms. We also calculated the excretion properties of all inhibitors (Table S6) and classified the data according to Paine et al. i.e. high (>1 mL/min/kg), medium (> 0.1 to <1 mL/min/kg) and low (≤0.1 mL/min/kg)[75]. ZINC001062406583, ZINC000679651603, ZINC000680430230 have high renal clearance compare to other molecules. The toxicity analysis of these molecules is also listed in Table

S7. Except for two molecules, the rest of them are found to be non-Ames toxic i.e., they have less chance to create mutation, which leads to cancer. These are ZINC000541676760, ZINC000 527019428, ZINC000544491494. Inhibition of HERG-1 channel leads to QT syndrome whereas all molecules in our studies are found to be non-inhibitor of the same. Besides, all of them are negative in sensitizing of the skin.

We also determine the bio-activity score of all drug-like molecules in our study using the molinspiration portal and listed in Table 3. Molecules having more than 0, said to be bioactive, score in between 0 to -0.5 terms as moderate bioactive and finally, less than -0.5 are inactive. All ligands molecules pass the bioactive criteria except ZINC000527019428 and ZINC000544491494 which have less than -0.5 against nuclear receptors.

Toxicity assessment of potential inhibitors via in-silico method:

Clinical safety of novel therapeutic agent is one of the key concerns for successful drug development[76]. Carcinogenicity, mutagenicity and cytotoxicity are one of the major concerns in the toxicity profiling after successful delivery of a specific drug[77, 78]. Also, we estimated the toxicity profiles for all our lead molecules by estimating their hepatotoxicity, carcinogenicity, mutagenicity and cytotoxicity with the help of ProTox-II server[49] and shown in Table 4. Predicted toxicity group for all molecules were 4 in 1-5 scale (higher the number lower the toxicity). Among all nine molecules, five of them are non-hepatotoxic, non-carcinogen, non-mutagenic and non-cytotoxic. ZINC000541677852, ZINC000541676760, ZINC000541676760 and ZINC 000544491494 have slightly hepatoxicity and carcinogenic in nature.

Finally, five molecules were selected by analysing all parameters, and then their binding mechanism was studied in detail via molecular dynamic simulation. These are ZINC001062406583, ZINC000571366263, ZINC000452260308, ZINC 000679651603 and ZINC000680430230.

Molecular dynamics studies

Followed the virtual screening and the ADMET screening, we obtained the best five compounds from the library of 1000 molecules. Then to investigate further the binding mechanism and dynamic behaviour of these molecules, we performed the MD simulations of complexes for 100 ns.

Structural stability and flexibility of the complexes

Firstly, we computed the root-mean-square deviations (RMSD) of backbone atoms for all the complexes relative to their initial structures. The time evolution of the RMSD is shown in Figure 3A. RMSD values for $3CL^{pro}$ complexes with ZINC000571366263 and ZINC000452260308 remain stable after 60 ns depicting the convergence of simulations. Similarly, the complex of ZINC001062406583 converged after 70 ns. However, for the complexes of ZINC000680430230 and ZINC000679651603 takes a longer time to converge approximately after 80 ns. Overall, within 100 ns, all systems got converged. The average values of RMSDs for all the complexes are listed in Table 5. The average values vary between 1.44 Å and 3.22 Å. The highest deviation was observed for the $3CL^{pro}$/ZINC001062406583 (3.22 Å), while the lowest was obtained for $3CL^{pro}$/ZINC000452260308 (1.44 Å). It suggests that the complex with ZINC000452260308 is more stable in the simulations. Besides, we computed the centre of the mass (COM) distance between inhibitor and protein and shown in Figure 3B. The average values vary between 17.34 Å and 24.20 Å. For which the highest average value was obtained for the $3CL^{pro}$/ZINC000679651603. In contrast, the lowest value was obtained for the $3CL^{pro}$/ZINC000452260308. The time evolution of the COM from Figure 3B reveals that complexes $3CL^{pro}$/ZINC000452260308 and $3CL^{pro}$/ZINC000680430230 showed less average distance compared with $3CL^{pro}$/ZINC001062406583 and $3CL^{pro}$/ZINC000571366263, depicting the strong interaction of these inhibitors with the protein. However, the time evolution of the COM distance for the $3CL^{pro}$/ZINC000679651603 showed that after 80 ns, the distance increases, depicting that interaction between the inhibitor and $3CL^{pro}$ breaks, suggesting that the same ligand is no longer bonded with the $3CL^{pro}$. Furthermore, the RMSD of all inhibitors were computed and shown in Figure 3C**.** The average values vary between 0.87 Å and 1.63 Å. The lowest was obtained for the complex $3CL^{pro}$/ZINC000452260308, while the highest was obtained for the $3CL^{pro}$/ZINC000679651603. It also agrees that ZINC000679651603 showed more significant deviations compared to other inhibitors and hence could not be considered as a potent inhibitor against $3CL^{pro}$. Besides, the potential of mean force (PMF) w.r.t RMSD of inhibitors were also calculated and shown in Figure 3D. It depicts that the lowest RMSD with a single narrow peak was observed for ZINC000452260308. ZINC000680430230 showed slightly higher RMSD compared to ZINC001062406583 but it showed a narrower peak than ZINC001062406583. Further increased RMSD with broader peak was observed for the ZINC000571366263. However, the highest RMSD peak value was obtained for the ZINC000679651603. Overall, it

suggests that among the screened inhibitors, ZINC000452260308 showed strong and stable binding with the $3CL^{pro}$, while the ZINC000679651603 showed the least. Hence, for further analysis, we had discarded the ZINC000679651603 complex. For rest of the section, we term these top four molecules as lead 1 (ZINC001062406583), lead 2 (ZINC000571366263), lead 3 (ZINC000452260308) and lead 4 (ZINC000680430230) for the ease of the discussion,

Next, we investigated the flexibility of the different regions of all the complexes by calculating the root-mean-square-fluctuations (RMSF) of C_α atoms and shown in Figure 4A. It is evident from the RMSF plot that leads 1 complex exhibit more significant fluctuation in the different regions of $3CL^{pro}$ compared to other complexes. Also, lead 2 exhibits relatively large variations around the 50 residues (domain I). However, the lead 3 and lead 4 complexes showed lesser fluctuations compared to other inhibitors.

Furthermore, the compactness of all the complexes was estimated by calculating the radius of gyration (RoG) from the MD trajectories and shown in Figure 4B. The average values of RoG are reported in Table 5. It varies between 21.71 Å and 22.18 Å. It showed approximately similar RoG for all the systems. Also, the solvent-accessible-surface area (SASA) for all the systems were calculated from the MD trajectories and shown in Figure 4C. The reported average value of SASA (Table 5) varies between 137.52 $Å^2$ and 142.93 $Å^2$. The lead 2 complex showed the highest SASA value compared to other complexes.

Binding free energy of protein-inhibitor complex

To best describe the binding potency of the selected molecules against the $3CL^{pro}$ from the virtual screening, we had computed the binding free energy and its various components using the MM/GBSA scheme. It provides the total binding energy (ΔG_{bind}) components such as van der Waal interactions (ΔE_{vdW}), electrostatic interactions (ΔE_{ele}), polar solvation energy (ΔG_{pol}), non-polar solvation free energy (ΔG_{np}) and configurational entropy ($T\Delta S$). In total, 2500 frames were chosen from the stable regions for the calculation of binding free energy, and 25 frames were selected for the high-computational entropy calculation via nmode method.

The details of the binding free energy and its various components are listed in Table 6 and shown graphically in Figure 5A. MM/GBSA method can also be used to rank the different compounds and as can be seen from Table 6 that the high binding affinity for the $3CL^{pro}$ is observed for lead 3 (-14.31 kcal/mol) followed by lead 3 (-11.45 kcal/mol), lead 2 (-7.21 kcal/mol), and lead 4 (-6.55 kcal/mol). Overall, it suggests that among the screened compounds, ZINC-000452260308 (lead 3) binds strongly with the $3CL^{pro}$. It is evident from

the binding energy calculation that the van der Waal interactions (ΔE_{vdW}) and electrostatic interactions (ΔE_{ele}) favours the binding between all the inhibitors against $3CL^{pro}$. Besides, the non-polar solvation energy (ΔG_{np}) also supports complex formation. In contrast, the polar solvation energy (ΔG_{pol}), as well as the configurational entropy ($T\Delta S$), opposes the complexation. It is noted that ΔE_{vdW} varies between -36.31 kcal/mol and -42.56 kcal/mol, while ΔE_{ele} varies between -14.80 kcal/mol and -24.45 kcal/mol for all the complexes. It suggests that the van der Waal interaction plays a significant role in the complex formation between the inhibitors and $3CL^{pro}$. Among the screened compounds, the most favourable values for both ΔE_{vdW} (-42.56 kcal/mol) and ΔE_{ele} (-24.45 kcal/mol) were found to be for the lead 3 complex. Although the overall net polar contributions ($\Delta E_{ele} + \Delta G_{pol}$) for compounds lead 1, lead 2, lead 3, and lead 4 were found to be nearly close and unfavourable as 13.06 kcal/mol, 13.38 kcal/mol, 13.48 kcal/mol, and 12.22 kcal/mol, respectively. However, the overall net non-polar contributions ($\Delta E_{vdW} + \Delta G_{np}$) for complexes such as compounds lead 1, lead 2, lead 3, and lead 4 were found to be -40.97 kcal/mol, -41.13 kcal/mol, -47.97 kcal/mol, and -44.71 kcal/mol, respectively. This suggests that the higher binding affinity for the complex $3CL^{pro}$/lead 3 was due to its favourable non-polar energy components compared to other complexes. In comparing with our previous study with $3CL^{pro}$ with α-ketomide, ZINC000452260308 or lead 3 has overall higher finding affinity mostly because of the sharp decrease in the entropic contribution[56].

Per-residue contributions to binding free energy

To provide further insights into the binding mechanism of the screened inhibitors and $3CL^{pro}$, the contributions from each residue was calculated using the MM/GBSA scheme and was listed in Table 7. The per-residue decomposition of free energy depicts the hot spot residues involved in the binding of a protein-inhibitor complex. The values higher than 1.0 kcal/mol were considered and shown in Table 7. The interaction spectra of all the protein-inhibitor complexes are also shown in Figure 5B. It is evident from Figure 5B that residues Met49, Met165, and Gln189 were found to be common critical residues for all the screened compounds. Besides, Asn142 and Ser144 were also observed to be crucial for the complexes $3CL^{pro}$/lead 3 and $3CL^{pro}$/lead 4. This result also agrees with a higher binding affinity of lead 3 and lead 4 compared to lead 1 and lead 2. However, the binding energy of these hot spot residues obtained for $3CL^{pro}$/lead 3 was found to be more favourable compared to the complex $3CL^{pro}$/lead 4 Overall, the identification of these critical residues for the $3CL^{pro}$/inhibitors can facilitate the discovery of the new selective inhibitors against $3CL^{pro}$.

Finally, we supplemented the above results by analysing the final conformation of each production simulation with the help of Schrodinger Maestro software, and different h-bonds and hydrophobic interactions were shown in Figure 6. Hydrogen bonds are depicted in a pink single arrow line, while dark lime-green residues are involved in hydrophobic interactions. For the $3CL^{pro}$/lead 1 complex, Figure 6A shows two very stable h-bonds with Gly143 and Asn142 and displayed seven hydrophobic interactions with Leu27, Val42, Cys44, Met49, Met165 and Val186. Lead 2 or ZINC000571366263 formed hydrophobic interactions with Met49, Cys145, Met165, Leu167 and Val186 and formed three H-bonds with Gly143, Asn142 and Gln189 (see in Figure 6B). In the case of lead 3, we noticed that there were four stable H-bonds with Leu141, Gly143, Ser144 and Cys145 and six hydrophobic interactions with Leu27, Met49, Phe140, Leu141, Cys145 and Met165 were formed as revealed by Figure 6C. Finally, Figure 6D shows that $3CL^{pro}$/lead 4 formed hydrophobic interactions with Met49, Phe140, Leu141, Cys145 and Met165 and h-bonds forming residues are Gly143, Cys145 and Leu141. The hydrophobic and lipophilicity of four complexes were shown in Figure 7. Overall, lead 3 has a higher binding affinity toward main protease $3CL^{pro}$ compared to the other lead small molecules due to a more significant number of stable hydrogen bonds and hydrophobic interactions. Also, we have analysed the number of hydrogen bonds, as well as their stability between protein-drug like molecules and are listed in Table 8 which also validate the above results. Lead 2 only has one strong hydrogen bond (Q189@NE2 and O2 atom of ligand) having occupancy around 46 %, which indicates the lesser amount of electrostatic contribution in the binding free energy of the same. On the other hand, there are 4-5 moderately strong H-bonds (above 15 %) for the other three cases show a significant increase in the electrostatic contribution (see Table 6). Overall, all the analysis summarise that the lead 3 or ZINC000452260308 drug-like molecule has a potent candidate to the COVID-19 protease target.

Conclusion

Our study on the top 1000 molecules database constructed by Ton *et al.* .[1] estimated four drug-like molecules which are passed all the criteria of pharmacology and also show a promising binding against the COVID-19 protease. Surprisingly our prediction ruled out the top leads of the same database i.e. ZINC000541677852 as because of the toxicity pattern of it. At first, from the 1000 molecules, a virtual screening protocol was conducted by the glide module of Schrodinger[45] leads to nine molecules. Further analysis of different properties of ADMET using several servers leads to 5 molecules. All these five molecules showed better

pharmacological and pharmacokinetic properties such as cell membrane permeability, BBB permeability, proper metabolism against cytochrome isoforms, non-Ames toxicity and others. Also, these molecules are non-carcinogenic, non-cytotoxic, non-hepatotoxic and non-carcinogenic. We also extended our studies with these molecules through MD simulation and found out one of them i.e. ZINC000679651603, detached from the protease on a long time scale. Our study of the top four molecules up to 100 ns along with the binding free energy calculation mechanism scheme MM/GBSA stated the mechanism of binding. A combination of the non-polar and the van der Waals distinguish the best lead molecule i.e. ZINC000452260308 from rest of the other. Less penalty from the entropic side also has a significant effect on achieving higher binding affinity in all four cases and makes a strong candidate for further drug design. So, this lead molecule can be used for the backbone of further modification of their functional groups and other pharmacological parameters to a potent candidate of the drug development against the novel coronavirus pandemic.

Supporting Information

The Supporting Information is available free of charge on the XXX Publications website:

Acknowledgements

This work was partially supported by the Department of Biotechnology, Govt. of India (grant number BT/RLF/Re-entry/40/2014, DBT-Ramalingaswami Re-entry Fellowship), and Department of Science and Technology (DST), Govt. of India (grant number ECR/2017/000010). RJ thanks the Indian Institute of Technology Indore for financial assistance.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Table 1: Different components of docking scores obtained from Glide-XP docking protocol.

Lead Molecule	Molecular Weight	G-Score[a]	Glide-lipo[b]	Glide-hbond[c]	Glide-evdw[d]
ZINC 000541677852	394.397	-11.5702	-3.0529	-1.648	-44.2723
ZINC 000541676760	380.37	-11.5567	-3.00123	-1.57357	-44.6861
ZINC 001062406583	323.437	-11.2196	-2.21994	-1.31262	-30.7709
ZINC 000571366263	368.399	-11.1854	-3.23713	-1.59783	-44.513
ZINC 000452260308	370.371	-11.1382	-2.96289	-1.52686	-38.3051
ZINC 000679651603	330.428	-11.0647	-3.22382	-1.57569	-39.3433
ZINC 000680430230	344.455	-11.0198	-3.43388	-1.57329	-41.1864
ZINC 000527019428	330.363	-11.0055	-3.19259	-0.50507	-37.7579
ZINC 000544491494	391.269	-10.9977	-3.24318	-1.5716	-44.8975

a-Glide Score (kcal/mol)
b-Lipophilic term derived from hydrophobic grid potential
c-hydrophilic term
d-protein-ligand steric contact information

Table 2: Prediction of essential ADME properties of all lead molecules obtained from the QikProp.

Lead Molecule	logP(o/w)[a]	logS[b]	Caco-2d[c]	logHERG[d]	PMDCK[e]	%Human Oral absorption
ZINC 000541677852	3.001	-4.028	354.857	-3.567	1111.358	90.161
ZINC 000541676760	3.097	-4.877	394.405	-4.225	1933.238	91.54
ZINC 001062406583	0.204	-1.169	256.239	0.246	453.297	71.25
ZINC 000571366263	2.887	-4.193	497.153	-2.082	3360.944	92.111
ZINC 000452260308	2.868	-3.93	422.98	-2.291	3162.588	90.744
ZINC 000679651603	1.978	-3.077	526.344	-1.858	750.711	87.235
ZINC 000680430230	2.377	-3.369	498.185	-2.147	767.939	89.141
ZINC 000527019428	1.959	-3.388	268.696	-3.805	529.453	81.892
ZINC 000544491494	2.695	-4.318	392.973	-4.201	1171.076	89.157

a-Predicted octanol/water partition coefficient log P (acceptable range −2.0 to 6.5)
b—Predicted aqueous solubility log S in mol/L (acceptable range: −6.5 to 0.5)
c—Predicted Caco-2 cell permeability in nm/s (acceptable range: < 25 is poor and > 500 is great)
d—Predicted IC 50 value for blockage of HERG K+ channels (concern below −5.0)
e—Predicted apparent MDCK cell permeability in nm/s (acceptable range: < 25 is poor > 500 is great)
f—Percentage of human oral absorption (acceptable range: < 25% is poor and > 80% is high)

Table 3: Bioactivity of lead molecules using Molinspiration webserver.

Lead Molecule	GPCR ligand	Ion channel modulator	Kinase Inhibitor	Nuclear receptor ligand	Protease Inhibitor	Enzyme Inhibitor
ZINC 000541677852	0.02	-0.28	-0.03	-0.43	0.10	-0.34
ZINC 000541676760	-0.01	-0.32	-0.04	-0.42	0.08	-0.36
ZINC 001062406583	0.31	-0.03	-0.32	-0.27	0.55	0.03
ZINC 000571366263	0.17	0.09	-0.32	-0.26	0.40	-0.02
ZINC 000452260308	0.28	0.27	-0.42	-0.17	0.58	0.03
ZINC 000679651603	0.08	-0.06	-0.45	-0.45	0.33	-0.07
ZINC 000680430230	0.10	-0.04	-0.42	-0.46	0.36	-0.05
ZINC 000527019428	0.09	-0.18	0.13	-0.63	0.07	-0.12
ZINC 000544491494	-0.20	-0.49	-0.18	-0.81	-0.14	-0.48

Table 4: Toxicity prediction of all lead molecules using ProTox-II webserver.

Lead Molecule	Hepato toxicity	Carcino genicity	Muta genicity	Cyto toxicity	Predicted LD50 (mg/kg)	Acute Toxicity
ZINC 000541677852	Active (0.53)	Active (0.59)	Inactive (0.61)	Inactive (0.74)	1000	IV
ZINC 000541676760	Active (0.57)	Active (0.59)	Inactive (0.60)	Inactive (0.74)	1000	IV
ZINC 001062406583	Inactive (0.84)	Inactive (0.61)	Inactive (0.75)	Inactive (0.73)	400	IV
ZINC 000571366263	Inactive (0.84)	Inactive (0.61)	Inactive (0.75)	Inactive (0.73)	680	IV
ZINC 000452260308	Inactive (0.73)	Inactive (0.59)	Inactive (0.62)	Inactive (0.70)	1250	IV
ZINC 000679651603	Inactive (0.82)	Inactive (0.58)	Inactive (0.73)	Inactive (0.74)	1250	IV
ZINC 000680430230	Inactive (0.82)	Inactive (0.61)	Inactive (0.70)	Inactive (0.70)	1250	IV
ZINC 000527019428	Active (0.52)	Active (0.55)	Inactive (0.65)	Inactive (0.70)	1000	IV
ZINC 000544491494	Active (0.52)	Active (0.50)	Inactive (0.61)	Inactive (0.69)	1300	IV

Table 5: Average RMSD (backbone of protease, ligand), radius of gyration and solvent accessible surface area for all protein-drug complexes.

Lead Molecule	Backbone RMSD (Å)	Ligand RMSD (Å)	RoG (Å)	SASA (nm^2)
ZINC 001062406583	3.224 (0.002)	1.263 (0.004)	21.712 (0.002)	137.519 (0.037)
ZINC 000571366263	1.516 (0.002)	1.420 (0.002)	22.183 (0.002)	142.927 (0.033)
ZINC 000452260308	1.435 (0.003)	0.865 (0.003)	22.025 (0.001)	141.344 (0.034)
ZINC 000680430230	2.433 (0.003)	1.202 (0.004)	21.875 (0.001)	139.140 (0.043)
ZINC 000679651603	2.620 (0.005)	1.631 (0.004)		

Table 6: Components of the binding free energy for COVID-19 protein and inhibitors complex using MMGBSA scheme. Standrad errors of the mean are provided in the paraenthesis.

Components	ZINC 001062406583	ZINC 000571366263	ZINC 000452260308	ZINC 000680430230
ΔE_{vdW}	-36.47	-36.31	-42.56	-39.99
	(0.12)	(0.07)	(0.06)	(0.09)
ΔE_{elec}	-20.50	-14.80	-24.45	-24.03
	(0.17)	(0.12)	(0.08)	(0.08)
ΔG_{pol}	33.56	28.18	37.93	36.25
	(0.14)	(0.10)	(0.06)	(0.07)
ΔG_{np}	-4.50	-4.82	-5.41	-4.72
	(0.01)	(0.00)	(0.00)	(0.01)
ΔG_{solv}[a]	29.06	23.36	32.52	31.53
	(0.13)	(0.10)	(0.05)	(0.07)
ΔE_{MM}[b]	-56.97	-51.11	-67.01	-64.02
	(0.25)	(0.15)	(0.10)	(0.12)
$-T\Delta S$	21.36	20.54	20.18	21.04
	(0.88)	(0.74)	(0.88)	(0.65)
ΔG_{Total}[c]	-27.91	-27.75	-34.49	-32.49
	(0.14)	(0.07)	(0.07)	(0.08)
$\Delta G_{Bind}^{sim\text{-}d}$	-6.55	-7.21	-14.31	-11.45
	(0.89)	(0.74)	(0.88)	(0.65)

a- $\Delta G_{np}+ \Delta G_{pol}$
b- $\Delta E_{vdW}+ \Delta E_{elec}$
c- $\Delta E_{vdW}+ \Delta E_{elec}+ \Delta G_{np}+ \Delta G_{pol}$
d- $\Delta E_{vdW}+ \Delta E_{elec}+ \Delta G_{np}+ \Delta G_{pol}-T\Delta S$

Table 7: Decomposition of binding free energy of all four complexes by each amino acids of protease.

Residue	E_{vdW}	E_{elec}	G_{pol}	G_{np}	G_{side_chain}	$G_{backbone}$	G_{total}
3CL^{PRO}/ *ZINC001062406583*							
M49	-2.08	-0.92	1.12	-0.21	-1.92	-0.17	-2.09
Q189	-1.45	-1.54	1.91	-0.24	-1.05	-0.27	-1.32
M165	-1.87	1.17	-0.42	-0.20	-1.46	0.14	-1.32
S46	-0.83	-1.01	0.96	-0.19	-0.64	-0.43	-1.07
H41	-1.57	-1.09	1.82	-0.16	-0.95	-0.05	-1.00
3CL^{PRO}/ *ZINC000571366263*							
Q189	-2.22	-3.03	2.90	-0.36	-2.26	-0.45	-2.71
M165	-2.34	0.19	0.06	-0.27	-2.18	-0.18	-2.36
M49	-1.44	-0.19	0.38	-0.11	-1.24	-0.12	-1.36
3CL^{PRO}/ *ZINC000452260308*							
M165	-2.61	0.12	0.58	-0.29	-2.11	-0.09	-2.20
S144	-0.58	-2.64	1.28	-0.01	-0.67	-1.28	-1.95
C145	-1.15	-0.73	0.42	-0.14	-0.84	-0.76	-1.60
N142	-1.80	-1.25	1.85	-0.25	-0.56	0.89	-1.45
G143	-0.26	-2.31	1.30	-0.03	-0.10	-1.20	-1.30
M49	-1.19	0.07	0.06	-0.20	-1.18	-0.08	-1.26
Q189	-1.81	-0.49	1.39	-0.27	-0.90	-0.28	-1.18
3CL^{PRO}/ *ZINC000680430230*							
S144	-0.79	-2.24	1.20	-0.02	-0.70	-1.15	-1.85
C145	-1.01	-1.24	0.68	-0.13	-0.82	-0.88	-1.70
M49	-1.42	-0.42	0.52	-0.16	-1.35	-0.12	-1.48
M165	-2.14	0.83	0.06	-0.23	-1.58	0.10	-1.48
G143	-0.30	-2.53	1.45	-0.03	-0.10	-1.31	-1.41
N142	-1.83	-1.19	1.97	-0.26	-0.48	-0.83	-1.31
Q189	-2.01	-0.28	1.48	-0.34	-0.91	-0.24	-1.15

Energetics contributions from the van der Waals (E_{vdW}) and electrostatic interactions (E_{elec}) as well as polar (G_{pol}) and nonpolar solvation energy (G_{np}) and the total contribution of given residue (G_{total}) for 3CL^{PRO}-inhibitor complexes are listed. G_{side_chain} and $G_{backbone}$ represent the side chain and backbone contributions. Only residues with $|\Delta G| \geq 1.0$ kcal/mol are shown. All values are given in kcal/mol.

Table 8: Occupancy of hydrogen bonds between protein and ligand complexes in each case during MD simulation.

Acceptor	Donor	Distance (Å)	Angle (˚)	Occupancy (%)
Lig: *ZINC001062406583*				
Lig@O1	C145@N	2.88	151.41	22.19
Lig@O1	G143@N	2.85	149.15	21.82
Lig@O1	S144@N	2.86	147.23	15.88
Lig@O3	Q189@NE21	2.84	159.75	10.04
Lig@O3	S46@OG	2.75	162.95	6.72
N142@OD1	LIg@N1	2.86	162.52	27.96
L141@O	Lig@N1	2.88	162.55	10.77
Lig: *ZINC000571366263*				
Lig@O2	Q189@NE2	2.84	160.14	46.42
Lig@O2	Q189@NE2	2.86	155.56	5.91
Lig@F3	Q192@NE2	2.88	155.89	5.41
Lig@F1	Q192@NE2	2.88	155.33	5.37
Lig@F2	Q192@NE2	2.87	155.98	5.27
N142@OD1	Lig@N1	2.85	158.06	6.46
Lig: *ZINC000452260308*				
Lig@O1	S144@N	2.85	149.40	32.61
Lig@O1	C145@N	2.88	150.77	26.81
Lig@O1	G143@N	2.84	146.19	26.04
H164@O	Lig@N2	2.87	160.09	24.83
S144@OG	Lig@N1	2.91	160.74	16.61
S144@HG	Lig@N1	2.91	156.37	8.27
Lig: *ZINC000680430230*				
Lig@O1	C145@N	2.88	150.49	32.39
Lig@O1	S144@N	2.86	147.58	25.29
Lig@O1	G143@N	2.36	146.15	23.56
H164@O	Lig@N2	2.87	159.46	20.11
S144@OG	Lig@N1	2.92	159.64	10.52
S144@HG	Lig@N1	2.89	155.75	9.49

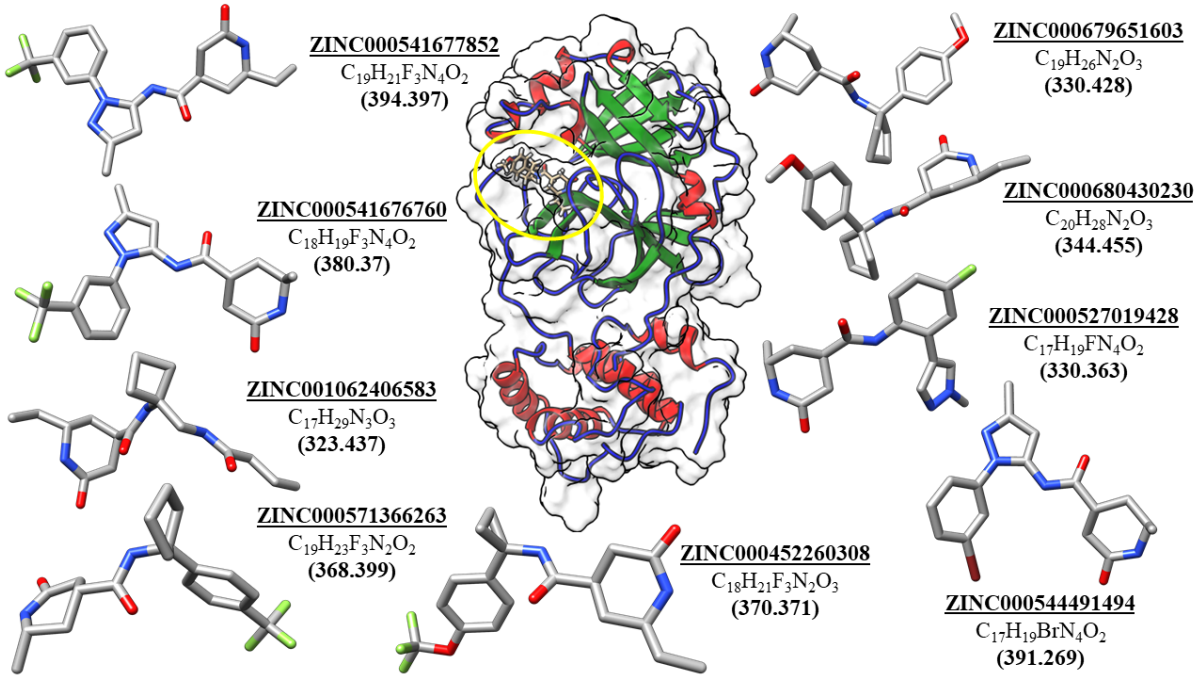

Figure 1: Cartoon representation of COVID-19 $3CL^{pro}$ along with inhibitor as shown in ball and stick. Top nine molecules that are used in our study also shown in ball and stick model along with their ZINC ID, chemical formula and molecular weight.

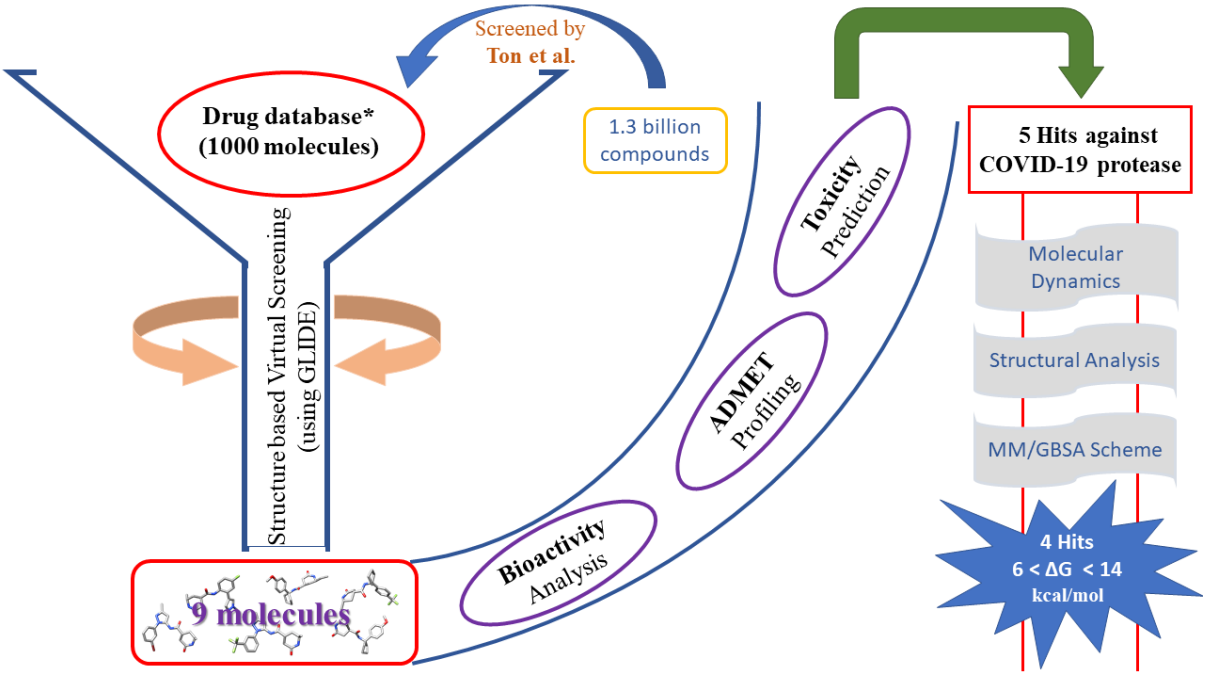

Figure 2: Workflow of indentification of lead molecules against of COVID-19 protease via strucuture based virtual screening. * Intial databases is downloaded from the Ton et al study[1].

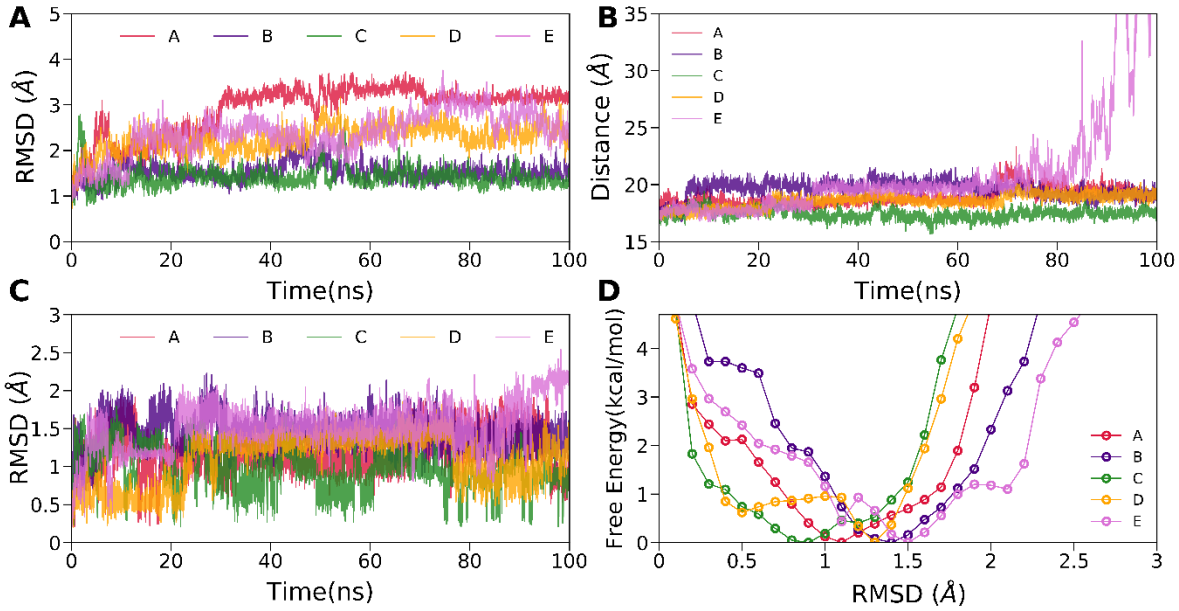

Figure 3: A) Time evolution of root mean square deviation (RMSD) backbone atoms of COVID-19 $3CL^{pro}$ B) Centre of mass (COM) distance between $3CL^{pro}$ and the inhibitors throughout the time scale C) time evolution of RMSD of inhibitors D) Potential mean force of ligand RMSD. PMF were calculated at 300 K. Inhibitors in the complexes as in legends are as follow: A) ZINC001062406583, B) ZINC000571366263, C) ZINC000452260308, D) ZINC000680430230, E) ZINC000679651603.

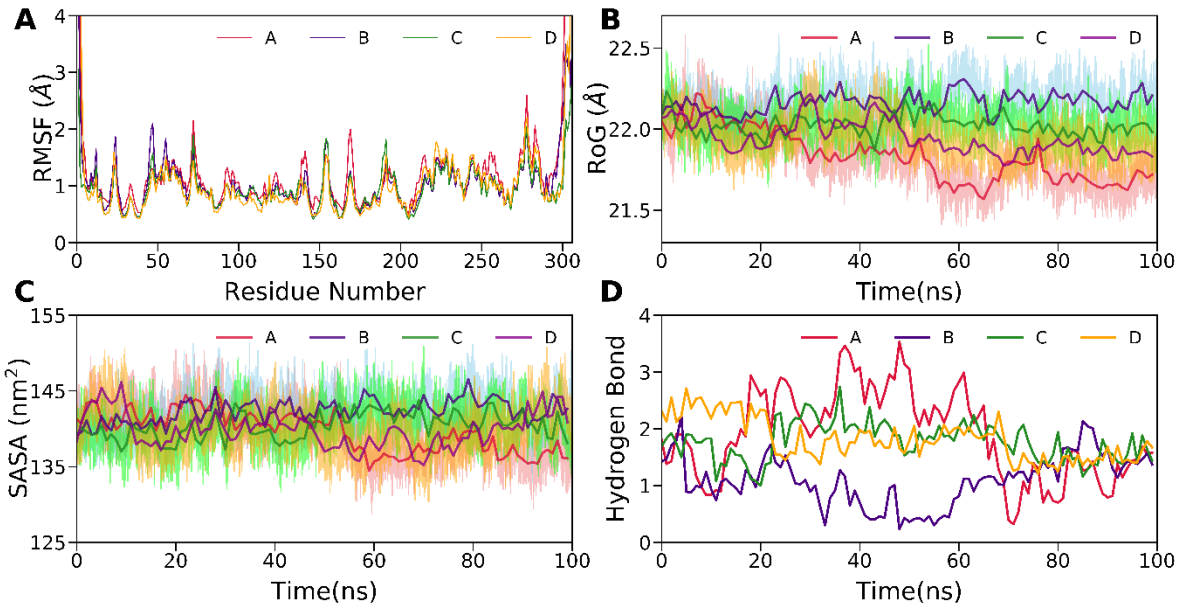

Figure 4: A) Root meand squre of fluctuations (RMSF) of Cα atoms of 3CLl^{pro}, B) Radius of gyration of 3CL^{pro}, C) solvent accessible surface area of 3CL^{pro} and D) numbrer of hydrogen bond between 3CL^{pro} and inhibitos for all four protein-inhibitor complexes. Inhibitors in the complexes as in legends are as follow: A) ZINC001062406583, B) ZINC000571366263, C) ZINC000452260308, D) ZINC000680430230.

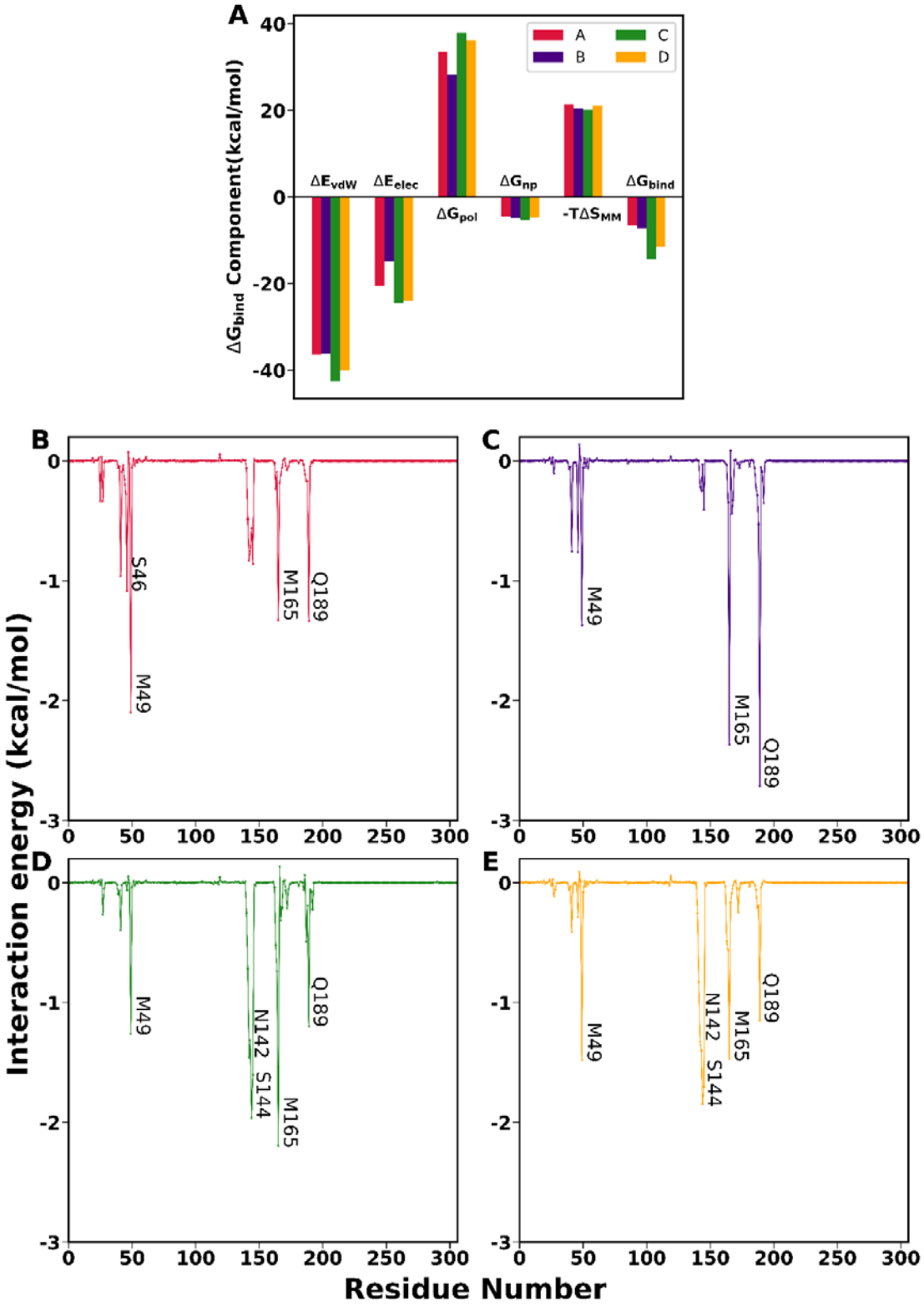

Figure 5: A) Components of binding free energy (kcal/mol) for protien-inhibitors complex. Each drug-protein complexes are presented as follows: ZINC001062406583 (red (A)), ZINC000571366263 (blue (B)), ZINC000452260308 (green (C)), ZINC000680430230 (yellow (D)) Decomposition of total binding free energy from individual amino acids B) ZINC001062406583, C) ZINC000571366263, D) ZINC000452260308, E) ZINC000680430230

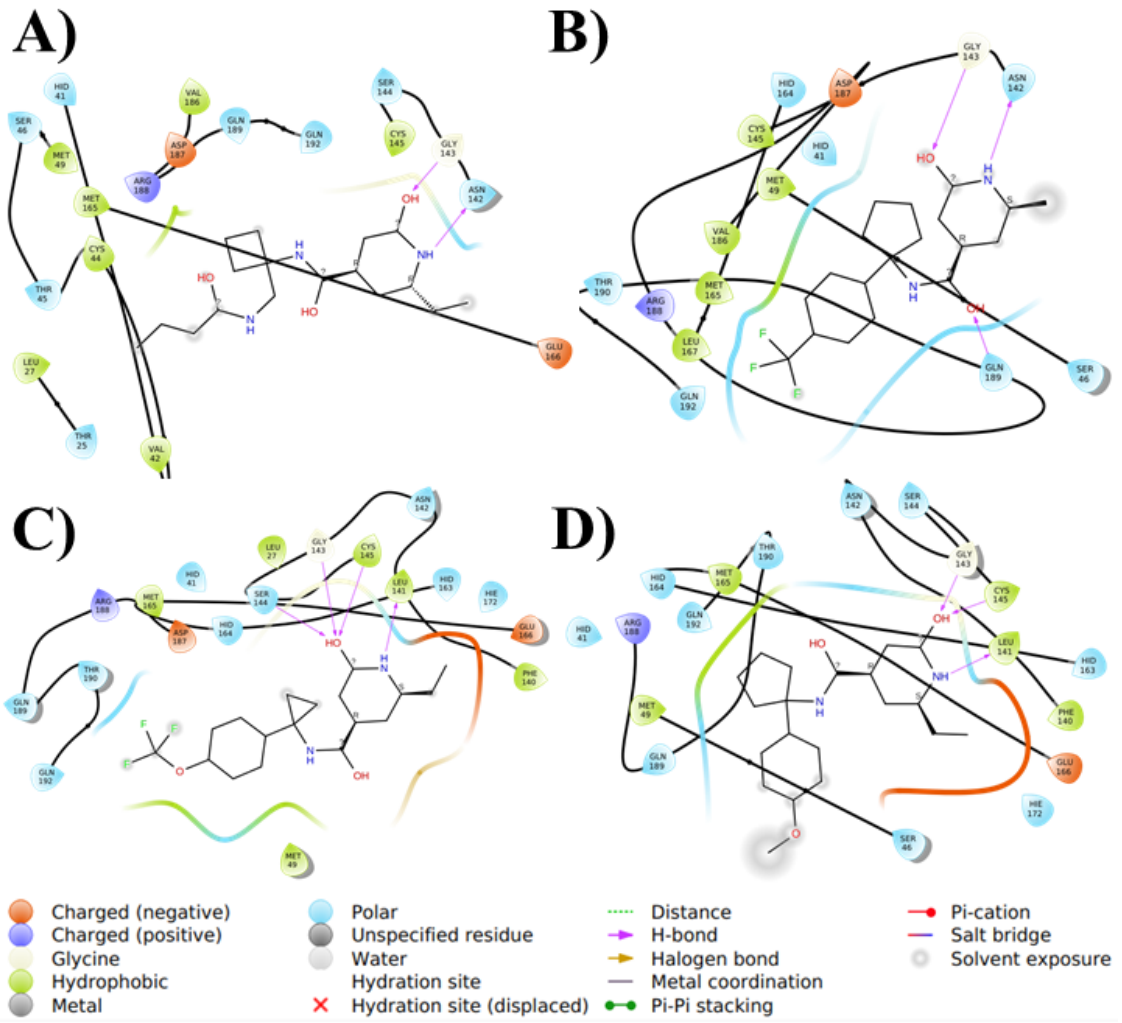

Figure 6: 2D interaction image for top four protein-inhibitor complexes. A) ZINC001062406583, B) ZINC000571366263, C) ZINC000452260308, D) ZINC000680430230.

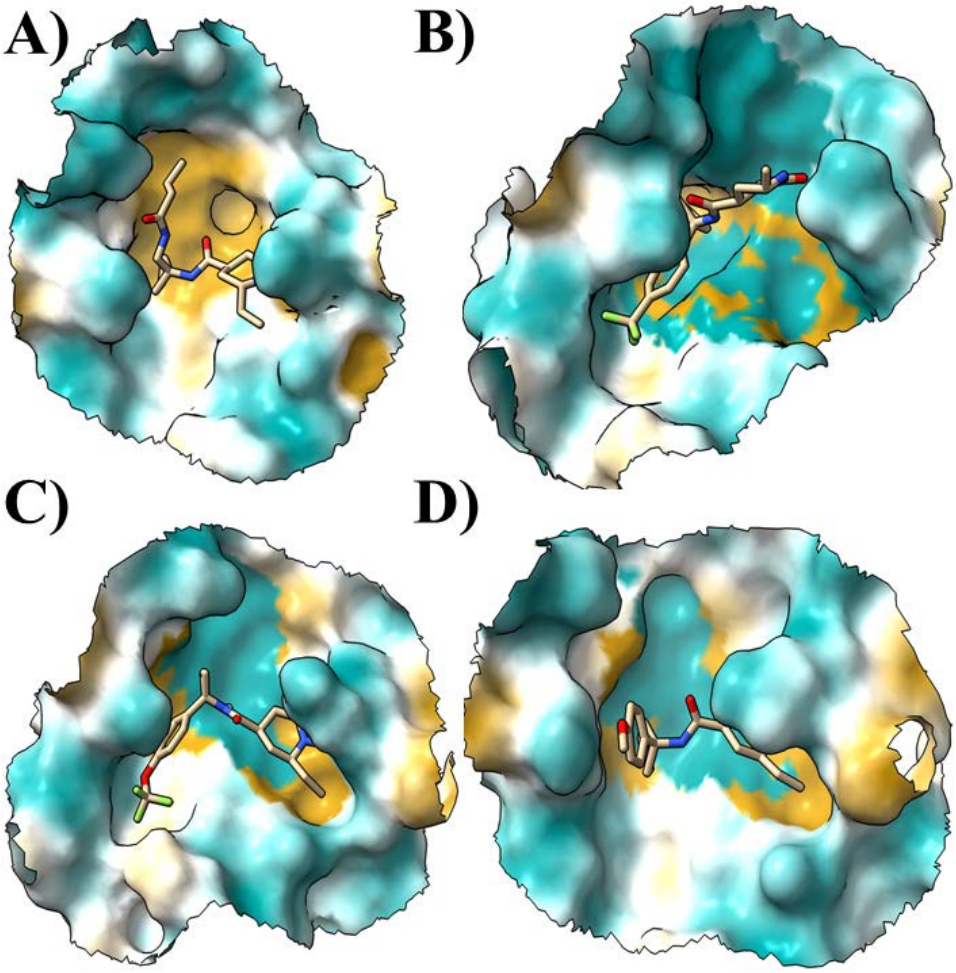

Figure 7: Surface representation of the drug bindiing portion of COVID-19 $3CL^{pro}$. Here molecular lipophilicity potential of the protein is mapped to the surface ignoring the non-protein atom in the calculation. Representation of the color-scheme is as follows: dark cyan (hydrophilic) to white to dark goldenrod (lipophilic). A) ZINC001062406583, B) ZINC000571366263, C) ZINC000452260308, D) ZINC000680430230.

References

1. Ton AT, Gentile F, Hsing M, Ban F, Cherkasov A. Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds. Molecular informatics 2020.
2. Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Science China Life Sciences 2020; 63:457-60.
3. Guarner J. Three emerging coronaviruses in two decades: the story of SARS, MERS, and now COVID-19. Oxford University Press US, 2020.
4. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. The Lancet 2020; 395:565-74.
5. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020; 579:265-9.
6. Chang C-k, Hou M-H, Chang C-F, Hsiao C-D, Huang T-h. The SARS coronavirus nucleocapsid protein–forms and functions. Antiviral Res 2014; 103:39-50.
7. Das S, Sarmah S, Lyndem S, Singha Roy A. An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1763201.
8. Elfiky AA. SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: An in silico perspective. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1761882.
9. Islam R, Parves R, Paul AS, Uddin N, Rahman MS, Mamun AA, et al. A Molecular Modeling Approach to Identify Effective Antiviral Phytochemicals against the Main Protease of SARS-CoV-2. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1761883.
10. Khan RJ, Jha RK, Amera G, Jain M, Singh E, Pathak A, et al. Targeting SARS-CoV-2: A Systematic Drug Repurposing Approach to Identify Promising Inhibitors Against 3C-like Proteinase and 2'-O-RiboseMethyltransferase. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1753577.
11. 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:doi:10.1080/07391102.2020.1751298.
12. Muralidharan N, Sakthivel R, Velmurugan D, Gromiha MM. Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-CoV-2 Protease against COVID-19. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1752802.
13. Sk MF, Roy R, Jonniya NA, Poddar S, Kar P. Elucidating biophysical basis of binding of inhibitors to SARS-CoV-2 main protease by using molecular dynamics simulations and free energy calculations. J Biomol Struct Dyn 2020:https://doi.org/10.1080/07391102.2020.1768149.
14. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Science 2020; 368:409-12.
15. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of Mpro from COVID-19 virus and discovery of its inhibitors. Nature 2020:doi:10.1038/s41586-020-2223-y.
16. Pant S, Singh M, Ravichandiran V, Murty U, Srivastava HK. Peptide-like and small-molecule inhibitors against Covid-19. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1757510.
17. Wang J. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) Through Computational Drug Repurposing Study. J Chem Inf Model 2020:doi:10.1021/acs.jcim.0c00179.
18. 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:doi:10.1080/07391102.2020.1758788.
19. Elmezayen AD, Al-Obaidi A, Şahin AT, Yelekçi K. Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. J Biomol Struct Dyn 2020:doi:10.1080/07391102.2020.1758791.

20. Enmozhi SK, Raja K, Sebastine I, Joseph J. Andrographolide As a Potential Inhibitor of SARS-CoV-2 Main Protease: An In Silico Approach. J Biomol Struct Dyn 2020:1-10.
21. Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. Journal of general virology 2002; 83:595-9.
22. Liu J, Cao R, Xu M, Wang X, Zhang H, Hu H, et al. Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. Cell discovery 2020; 6:1-4.
23. Woo PC, Huang Y, Lau SK, Tsoi Hw, Yuen Ky. In silico analysis of ORF1ab in coronavirus HKU1 genome reveals a unique putative cleavage site of coronavirus HKU1 3C-like protease. Microbiology and immunology 2005; 49:899-908.
24. Umesh DK, Chandrabose Selvaraj, Sanjeev Kumar Singh & Vikash Kumar Dubey. Identification of new anti-nCoV drug chemical compounds from Indian spices exploiting SARS-CoV-2 main protease as target. Journal of Biomolecular Structure and Dynamics 2020:doi:10.1080/07391102.2020.1763202.
25. Yuen K-S, Ye Z-W, Fung S-Y, Chan C-P, Jin D-Y. SARS-CoV-2 and COVID-19: The most important research questions. Cell & bioscience 2020; 10:1-5.
26. Boopathi S, Poma AB, Kolandaivel P. Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. Journal of Biomolecular Structure and Dynamics 2020:1-10.
27. Khan RJ, Jha RK, Amera GM, Jain M, Singh E, Pathak A, et al. Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2'-O-ribose methyltransferase. Journal of Biomolecular Structure and Dynamics 2020:1-14.
28. Aanouz I, Belhassan A, El-Khatabi K, Lakhlifi T, El-Ldrissi M, Bouachrine M. Moroccan Medicinal plants as inhibitors against SARS-CoV-2 main protease: Computational investigations. Journal of Biomolecular Structure and Dynamics 2020:1-9.
29. Gupta MK, Vemula S, Donde R, Gouda G, Behera L, Vadde R. In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel. Journal of Biomolecular Structure and Dynamics 2020:1-11.
30. 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. Journal of Biomolecular Structure and Dynamics 2020:1-10.
31. Muralidharan N, Sakthivel R, Velmurugan D, Gromiha MM. Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-CoV-2 Protease against COVID-19. Journal of Biomolecular Structure and Dynamics 2020:1-6.
32. Chen YW, Yiu C-PB, Wong K-Y. Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL pro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. F1000Research 2020; 9.
33. Elmezayen AD, Al-Obaidi A, Şahin AT, Yelekçi K. Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. Journal of Biomolecular Structure and Dynamics 2020:1-13.
34. Berman H, Henrick K, Nakamura H. Announcing the worldwide protein data bank. Nature Structural & Molecular Biology 2003; 10:980-.
35. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. Nature 2020:1-5.
36. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. Journal of computer-aided molecular design 2013; 27:221-34.
37. Schrödinger Release 2020-2: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2016; Impact, Schrödinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2020.
38. Schrödinger Release 2020-2: Maestro, Schrödinger, LLC, New York, NY, 2020.

39. Olsson MH, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical p K a predictions. Journal of chemical theory and computation 2011; 7:525-37.
40. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, et al. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. Journal of chemical theory and computation 2016; 12:281-96.
41. Schrödinger Release 2020-2: LigPrep, Schrödinger, LLC, New York, NY, 2020.
42. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. Journal of medicinal chemistry 2006; 49:6177-96.
43. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. Journal of medicinal chemistry 2004; 47:1750-9.
44. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of medicinal chemistry 2004; 47:1739-49.
45. Schrödinger Release 2020-2: Glide, Schrödinger, LLC, New York, NY, 2020.
46. Schrödinger Release 2020-2: QikProp, Schrödinger, LLC, New York, NY, 2020.
47. Pires DE, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. Journal of medicinal chemistry 2015; 58:4066-72.
48. Zhang L, McHale CM, Greene N, Snyder RD, Rich IN, Aardema MJ, et al. Emerging approaches in predictive toxicology. Environmental and molecular mutagenesis 2014; 55:679-88.
49. Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: a webserver for the prediction of toxicity of chemicals. Nucleic acids research 2018; 46:W257-W63.
50. Case DA, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham I, V.W.D. Cruzeiro, et al. AMBER 2018. 2018; University of California, San Francisco.
51. Price DJ, Brooks III CL. A modified TIP3P water potential for simulation with Ewald summation. The Journal of chemical physics 2004; 121:10096-103.
52. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of Chemical Theory and Computation 2015; 11:3696-713.
53. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. Journal of computational chemistry 2004; 25:1157-74.
54. Kräutler V, Van Gunsteren WF, Hünenberger PH. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. Journal of computational chemistry 2001; 22:501-8.
55. Darden T, York D, Pedersen L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. Journal of chemical physics 1993; 98:10089-92.
56. Sk MF, Roy R, Jonniya NA, Poddar S, Kar P. Elucidating biophysical basis of binding of inhibitors to SARS-CoV-2 main protease by using molecular dynamics simulations and free energy calculations. Journal of Biomolecular Structure and Dynamics 2020:1-21.
57. Roe DR, Cheatham III TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. Journal of Chemical Theory and Computation 2013; 9:3084-95.
58. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Accounts of chemical research 2000; 33:889-97.
59. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. Journal of molecular graphics and modelling 2006; 25:247-60.
60. Kar P, Seel M, Hansmann UH, Höfinger S. Dispersion Terms and Analysis of Size-and Charge Dependence in an Enhanced Poisson− Boltzmann Approach. Journal of Physical Chemistry B 2007; 111:8910-8.

61. Kar P, Wei Y, Hansmann UH, Höfinger S. Systematic study of the boundary composition in Poisson Boltzmann calculations. Journal of computational chemistry 2007; 28:2538-44.
62. Kar P, Lipowsky R, Knecht V. Importance of polar solvation for cross-reactivity of antibody and its variants with steroids. Journal of Physical Chemistry B 2011; 115:7661-9.
63. Kar P, Lipowsky R, Knecht V. Importance of polar solvation and configurational entropy for design of antiretroviral drugs targeting HIV-1 protease. The Journal of Physical Chemistry B 2013; 117:5793-805.
64. Kar P, Knecht V. Origin of decrease in potency of darunavir and two related antiviral inhibitors against HIV-2 compared to HIV-1 protease. Journal of Physical Chemistry B 2012; 116:2605-14.
65. Kar P, Knecht V. Energetic basis for drug resistance of HIV-1 protease mutants against amprenavir. Journal of computer-aided molecular design 2012; 26:215-32.
66. Kar P, Knecht V. Mutation-induced loop opening and energetics for binding of tamiflu to influenza N8 neuraminidase. Journal of Physical Chemistry B 2012; 116:6137-49.
67. Jonniya NA, Sk MF, Kar P. Investigating Phosphorylation-Induced Conformational Changes in WNK1 Kinase by Molecular Dynamics Simulations. ACS omega 2019.
68. Jonniya NA, Kar P. Investigating specificity of the anti-hypertensive inhibitor WNK463 against With-No-Lysine kinase family isoforms via multiscale simulations. Journal of Biomolecular Structure and Dynamics 2019:DOI:10.1080/07391102.2019.1602079.
69. Roy R, Ghosh B, Kar P. Investigating Conformational Dynamics of Lewis Y Oligosaccharides and Elucidating Blood Group Dependency of Cholera Using Molecular Dynamics. ACS omega 2020; 5:3932-42.
70. Sk MF, Roy R, Kar P. Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. Journal of Biomolecular Structure and Dynamics 2020:DOI: 10.1080/07391102.2020.1724196.
71. Sterling T, Irwin JJ. ZINC 15–ligand discovery for everyone. Journal of chemical information and modeling 2015; 55:2324-37.
72. Vass M, Tarcsay Á, Keserű GM. Multiple ligand docking by Glide: implications for virtual second-site screening. Journal of computer-aided molecular design 2012; 26:821-34.
73. Srimai V, Ramesh M, Parameshwar KS, Parthasarathy T. Computer-aided design of selective Cytochrome P450 inhibitors and docking studies of alkyl resorcinol derivatives. Medicinal Chemistry Research 2013; 22:5314-23.
74. Wang Z, Hop CE, Leung KH, Pang J. Determination of in vitro permeability of drug candidates through a Caco-2 cell monolayer by liquid chromatography/tandem mass spectrometry. Journal of mass spectrometry 2000; 35:71-6.
75. Paine SW, Barton P, Bird J, Denton R, Menochet K, Smith A, et al. A rapid computational filter for predicting the rate of human renal clearance. Journal of Molecular Graphics and Modelling 2010; 29:529-37.
76. Salonen JS, Nyman L, Boobis AR, Edwards RJ, Watts P, Lake BG, et al. Comparative studies on the cytochrome p450-associated metabolism and interaction potential of selegiline between human liver-derived in vitro systems. Drug metabolism and disposition 2003; 31:1093-102.
77. Liebler DC, Guengerich FP. Elucidating mechanisms of drug-induced toxicity. Nature reviews Drug discovery 2005; 4:410-20.
78. Benigni R, Bossa C. Mechanisms of chemical carcinogenicity and mutagenicity: a review with implications for predictive toxicology. Chemical reviews 2011; 111:2507-36.