Molecular Diversity
https://doi.org/10.1007/s11030-022-10513-6

ORIGINAL ARTICLE

Computational studies indicated the effectiveness of human metabolites against SARS-Cov-2 main protease

Rajarshi Roy[1] · Md Fulbabu Sk[1] · Omprakash Tanwar[2] · Parimal Kar[1]

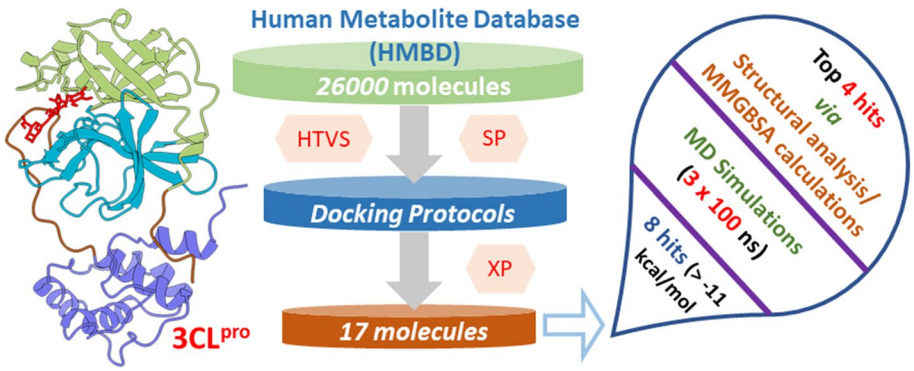
Received: 8 April 2022 / Accepted: 8 August 2022

Abstract

To fight against the devastating coronavirus disease 2019 (COVID-19), identifying robust anti-SARS-CoV-2 therapeutics from all possible directions is necessary. To contribute to this effort, we selected a human metabolites database containing waters and lipid-soluble metabolites to screen against the 3-chymotrypsin-like proteases ($3CL^{pro}$) protein of SARS-CoV-2. The top 8 hits from virtual screening displayed a docking score varying between ~ − 11 and ~ − 14 kcal/mol. Molecular dynamics simulations complement the virtual screening study in conjunction with the molecular mechanics generalized Born surface area (MM/GBSA) scheme. Our analyses revealed that (HMDB0132640) has the best glide docking score, − 14.06 kcal/mol, and MM-GBSA binding free energy, − 18.08 kcal/mol. The other three lead molecules are also selected along with the top molecule through a critical inspection of their pharmacokinetic properties. HMDB0132640 displayed a better binding affinity than the other three compounds (HMDB0127868, HMDB0134119, and HMDB0125821) due to increased favorable contributions from the intermolecular electrostatic and van der Waals interactions. Further, we have investigated the ligand-induced structural dynamics of the main protease. Overall, we have identified new compounds that can serve as potential leads for developing novel antiviral drugs against SARS-CoV-2 and elucidated molecular mechanisms of their binding to the main protease.

Graphical abstract

Identification of probable hits from human metabolites against SARS-CoV-2 using integrated computational approaches-Missed against MS



Keywords SARS-COV-2 · Main protease · Human metabolites · Virtual screening · Molecular dynamics · Free energy

Introduction

In China, a very contagious and severe viral disease was reported at the end of 2019 [1, 2]. This causative agent was later detected as a novel coronavirus (SARS-CoV-2). The

Rajarshi Roy and Md Fulbabu Sk have contributed equally.

✉ Parimal Kar
parimal@iiti.ac.in

Extended author information available on the last page of the article

Published online: 18 August 2022

disease was subsequently referred to as coronavirus disease 2019 (COVID-19) [3, 4]. This COVID-19 threatened human life and the economy across the globe and affected the lives of millions of people, including more than 6.4 million deaths as of August 5, 2022. It emerged as an epidemic and eventually became a pandemic. As a result, the current COVID-19 pandemic has triggered a medical emergency that is unprecedented in recent global history. There is an urgent need to discover effective therapeutics for treating COVID-19 and find some right immune booster supplements to stay healthy and safe [5].

Novel coronaviruses (SARS-CoV-2) are non-segmented, enveloped, positive-sense RNA, single-stranded viruses and are widely distributed in mammals, including humans [6]. Once the cell is infected with SARS-CoV-2, the host cell's ongoing molecular mechanisms are taken over by the virus to translate its RNA into long protein chains and generate more viral replicas. These viral polyproteins (PP1A) are activated when cut into smaller individual functional components by protease (3-chymotrypsin-like cysteine protease, $3CL^{pro}$) [7–9]. Thus, viral proteases ($3CL^{pro}$) are essential in SARS-CoV-2 virus propagation. Consequently, $3CL^{pro}$ is regarded as a principal druggable target for SARS-CoV-2. $3CL^{pro}$ is a heart-shaped protein also known as the main protease ($3CL^{pro}$). It contains 13 β-strands and nine α-helices and is divided into three domains [10, 11], domain I (residues 8–101), domain II (residues 102–184), a long loop (residues 185–200), and domain III (residues 201–306), shown in Fig. 1. Domain I and II contain primary substrate-binding catalytic dyads, His41 and Cys145, respectively. The C-terminal domain III is mainly involved in the main protease's dimerization [12]. Several recent computational studies have identified possible small molecules [13–16], plant-derived polyphenols [17, 18], and natural substances [19] to reduce COVID-19 infections by targeting several structural and non-structural proteins of SARS-CoV-2.

In recent investigations, metabolomics has been shown to be a promising technique for preventing viral diseases [20, 21]. Metabolomics studies different metabolites in living systems such as bio-fluids, cells, tissues, organisms, etc., which can be small molecules, drug metabolites, protein fragments, intermediates of any processes, etc. Metabolomics is an emerging technology, and very soon, it will be an essential part of precision medicine [22]. The Human Metabolome Database (HMDB) is a freely downloadable human metabolite database and contains various details about different metabolites found in the human body [23]. The HMDB contains 114,214 metabolite entries, including lipid-soluble and water-soluble metabolites [24]. Several medications, such as omeprazole, clopidogrel, cyclophosphamide, and diazepam, are present in the market and may become real pharmaceuticals following metabolism. Thus, metabolites can be an important lead for drug development, for example; (i) 6-Mercaptopurine (HMDB ID-HMDB0015167), azathioprine metabolite, which is an FDA-approved drug used to treat cancer and autoimmune; (ii) Oxazepam (HMDB ID-HMDB0014980), a benzodiazepine derivative which is used to treat anxiety; and (iii) Canrenone (HMDB ID- HMDB0003033), an important active metabolite of spironolactone, which is used as a diuretic agent and treatment of hirsutism [25].

Herein, we have screened the HMDB molecules against SARS-CoV-2 $3CL^{pro}$ to find some potential small molecules that might be useful to combat such deadly diseases. This study may be the first to use in-silico investigations to give an idea of the blocking function of $3CL^{pro}$ using human metabolites. Finally, we reported promising human metabolites against $3CL^{pro}$, which may be further developed as a therapeutic agent against COVID-19 based on their molecular mechanics generalized Born surface area (MM-GBSA) results.

Materials and methods

Preparation of protein

The primary atomic coordinates of $3CL^{pro}$ (PDB ID- 6LU7, resolution ~ 2.16 Å) [7] were downloaded from the protein data bank and prepared using the protein preparation wizard of the Schrödinger suite [26]. The water molecules and co-factors were removed before minimization, while only the co-crystalized ligand was kept for final energy minimization. The energy minimization was restricted to an RMSD cut-off value of 0.3 Å with the original structure under the OPLS3 force field [27]. The prepared protein was then used for the grid box generation, and a 20 Å box was created, keeping the crystallized ligand in the center. The same gird was used for our entire virtual screening protocol.

Ligand selection, preparation, and virtual screening protocol

We downloaded metabolites from the HMDB database (HMDB 4.0, accessed in May 2021) [24] and converted them to 3D structures using the Schrödinger software. HMDB contains various metabolites such as small molecules, peptides, triglycerides, etc. For curating potential leads from the metabolites, we selected compounds with molecular weights between 70 and 600 Da (~ 26,000 molecules). Since we are interested in studying all the metabolites in this range, we did not filter this database further using Lipinski's and related filters. Next, we performed the ligand preparation using the *LigPrep* module of the same suit, which resulted in ~ 36,801 entries. This module helps assign each ligand molecule's protonation and ionization

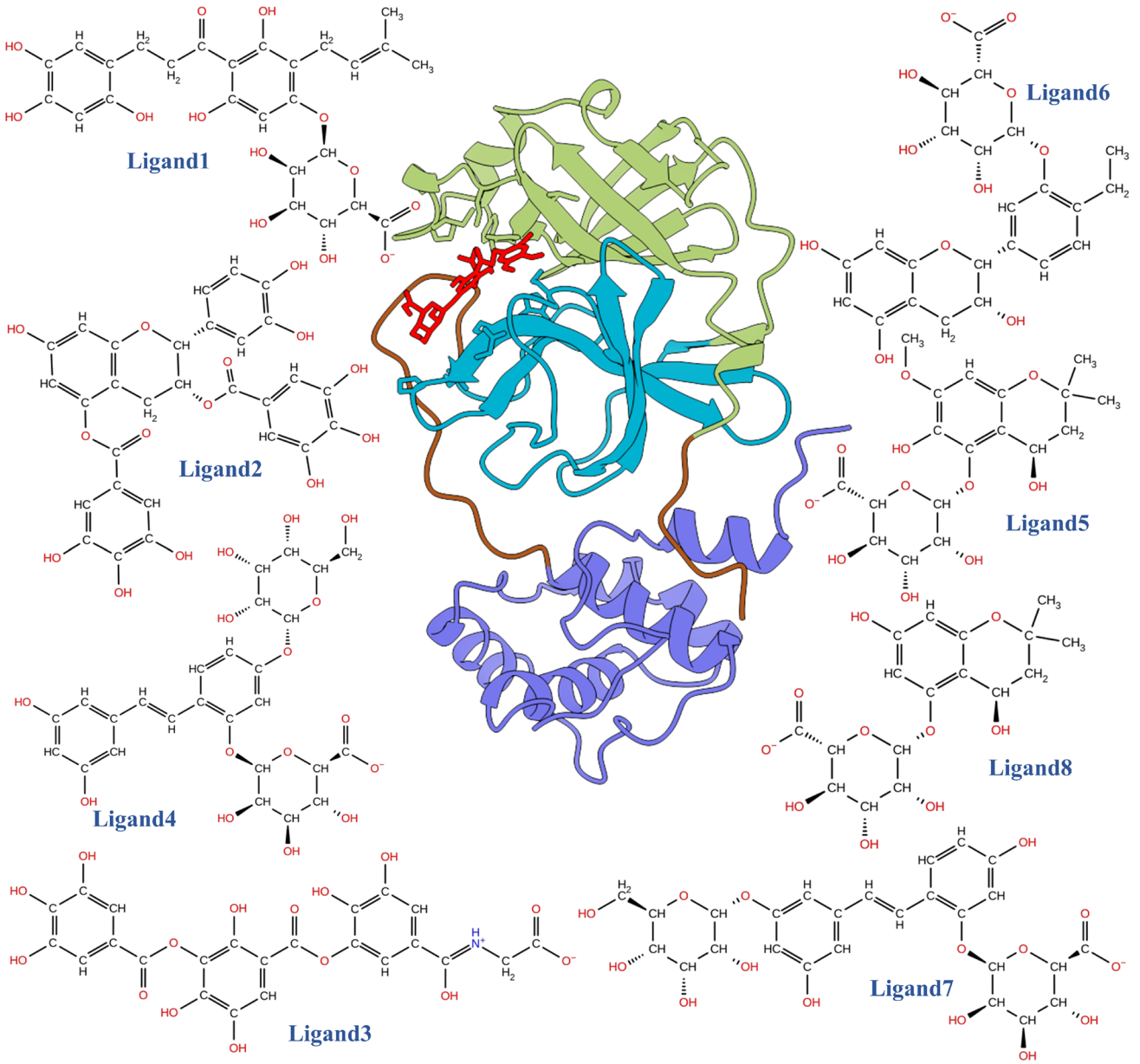


Fig. 1 Ribbon representation of COVID-19 $3CL^{pro}$ complexed with the inhibitor, shown in ball and stick representation. The different part of $3CL^{pro}$ is shown in different color, i.e., Green: Domain I, Cyan: Domain II, Blue: Domain III, Brown: Inter-domain connecting loop, and Red: Ligand molecules. The top 8 molecules which are screened by the virtual screening workflow are shown in 2D illustration

states. Several structures from each ligand were generated with different ionization states, tautomers, stereochemical information, and ring conformations. Finally, ligands were optimized, yielding low-energy isomers.

The virtual screening was done against $3CL^{pro}$ with the help of the virtual screening workflow (VSW) of the GLIDE module of Schrödinger [28–30]. The three tiers of virtual screening protocol, namely HTVS (high throughput virtual screening), SP (standard precision), and extra precision (XP), were sequentially employed to obtain potential lead molecules. To reduce the size of the database, we selected the top 30% of the docked complexes obtained from HTVS and used them for screening via Glide-SP. Finally, the top 20% of the SP docking results were considered for the Glide-XP (extra precision) docking. Moreover, the final compound selection was based on a visual inspection of the docking poses and the corresponding XP-docking scores, resulting in 17 unique molecules that were subjected to further analyses and

investigations. A similar protocol for the virtual screening study was used in our previous work on COVID-19 $3CL^{pro}$ [15].

Molecular dynamics (MD) simulation

Selected docked structures obtained from the virtual screening were used as an input to the *LEaP* module of AmberTools19 [31] to generate the input structure for MD simulations. The TIP3P [32] water model was used to solvate the structure in an octahedron water box with a 10 Å buffer region from all directions. An adequate number of ions were also added to neutralize the system. All the complexes were simulated using the Amber ff14SB forcefield [33] for protein and generalized Amber force field (GAFF2) [34] for ligand molecules. A 10 Å cut-off was fixed for calculating the long-range interaction with the help of the particle-mesh Ewald (PME) scheme [35]. The SHAKE [36] algorithm involved the bond length having hydrogen atoms to keep its motion constant. All the systems were simulated at a constant temperature of 300 K, which was maintained using a Langevin thermostat with a collision frequency of 2 ps^{-1}. A detailed description of the simulation protocol was discussed in our previous studies on COVID-19 [11, 37]. All complex structures were subjected to 3×100 ns production runs under the NPT ensemble. Trajectories were analyzed using the *cpptraj* module of AmberTools19, and the last 50 ns trajectories were used for the binding free energy calculation.

Protein–ligand-binding free energy calculation

The molecular mechanics Poisson–Boltzmann (generalized Born) surface area (MM-PB(GB)SA) scheme [38–42] is widely used to determine the binding free energy between protein-inhibitor [14, 43], protein-nucleic acid [44–46] as well as protein-carbohydrate [47, 48] complexes. The total binding free energy (ΔG_{bind})(ΔG_{bind}) comprises internal energy ($\Delta E_{\text{internal}}$), ($\Delta E_{\text{internal}}$), desolvation free energy (ΔG_{solv})(ΔG_{solv}), and configurational entropy ($T\Delta S$)($T\Delta S$) which are related by the following equation [16, 49–53],

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

To estimate the binding free energy, 2500 frames were selected uniformly from the last 50 ns trajectories, and calculation was done with the help of the *MMPBSA.py* script available on AmberTools19. The entropic contribution was estimated using the normal mode analysis, and the MM-GBSA pair-wise decomposition scheme also assessed the contribution from each amino acid.

ADMET studies of top ligands

To compute the top lead molecules' absorption, distribution, metabolism, and excretion properties, the *QikProp* module of the Schrödinger suite was used. Thirty-five significant pharmaceutical properties were monitored, such as CNS activity, % of human oral absorption, blood–brain barrier prediction, cell permeability, Lipinski rule, etc. The toxicity-related parameters, such as hepatotoxicity, carcinogenicity, mutagenicity, and cytotoxicity, were calculated using the ProTox-II web server [54].

Results and discussions

Virtual screening of metabolites against $3CL^{pro}$

To conduct the virtual screening of the human metabolite database, we started with ~ 36,801 entries out of 114,214 HMDB metabolites, which have been further screened using HTVS, Glide-SP, and Glide-XP protocol to get the top 17 lead compounds. These compounds were ranked according to their glide score, and the top 8 ligands were selected by keeping a cut-off value of − 11 kcal/mol. To verify the docking poses, the top 8 ligand molecules were re-docked using Autodock Vina [55] and Glide-XP separately. Superimposition of the docked structures using both methods suggested similar binding poses, as shown in Supporting Information Fig. S1. We also investigated similarities of our principal metabolites by estimating the Tanimoto coefficient.

Further, this estimation was also carried out for five FDA-approved drugs: Remdesivir, Ritonavir, Favipiravir, Indinavir, and Beclabuvir, as shown in Supporting Information, Table S1). Most of the metabolites are chemically distinct except *Ligand*5 and *Ligand*8. All other combinations showed a very low similarity like 0.14 (*Ligand*3 vs. *Ligand*4) to moderate similarity like 0.58 (*Ligand*4 vs. *Ligand*7). On the other hand, FDA-approved drugs showed significantly less similarity, ranging from 0.06 to 0.15. So, this suggests our metabolites are very much exclusive from the commonly used drugs showing a good binding.

The best lead molecule results in a Glide score of − 14 kcal/mol, indicating a promising candidate for drug design. The docking score of the top 8 lead compounds is shown in Table 1. We provided the SMILES of all top 8 ligands in Supporting Information Table S2. One of the best hits in the present study is **HMDB0132640**, i.e., *Ligand1,* which is a predicted metabolite of 1-(2,4,6-trihydroxy-3-(3-methyl-2-en-1-yl)phenyl)-3-(2,4,5-trihydroxy phenyl)propan-1-one, a non-cyclic derivative of 2-phenylchromen-4-one flavonoids. It belongs to the class of organic compounds known as flavonoid o-glycosides. It can also be classified as 2′-hydroxy-dihydrochalcones,

Table 1 Different components of docking scores obtained from the Glide-XP-docking scheme

Lead molecule	Molecular weight	G-score[a]	Glide-lipo[b]	Glide-hbond[c]	Glide-evdw[d]
*Ligand*1 (HMDB0132640)	568.657	− 14.060	− 4.160	0.000	− 39.543
*Ligand*2 (HMDB0030665)	622.706	− 12.399	− 2.859	− 0.339	− 61.108
*Ligand*3 (HMDB0128347)	573.59	− 12.151	− 2.359	− 0.800	− 48.830
*Ligand*4 (HMDB0134117)	598.64	− 11.724	− 2.228	− 0.769	− 36.536
*Ligand*5 (HMDB0125819)	424.444	− 11.534	− 3.702	− 0.160	− 47.989
*Ligand*6 (HMDB0127868)	492.564	− 11.134	− 2.904	− 0.887	− 29.007
*Ligand*7 (HMDB0134119)	598.64	− 11.119	− 3.060	− 0.430	− 40.603
*Ligand*8 (HMDB0125821)	394.418	− 11.051	− 3.369	− 0.480	− 44.484

[a]Glide score (kcal/mol)

[b]Lipophilic term derived from hydrophobic grid potential

[c]Hydrogen bonding term in GlideScore

[d]Van der Waals energy

which is biosynthesized by reducing α,β-unsaturated ketone (chalcone) in the Human gut and is considered to be a flavonoid type of molecule. This type of flavonoid is present in many plants and has shown antiviral properties. The flavonoids have been reported to have a complementary therapeutical role in the treatment of COVID-19 [56]. *Ligand*1 is one of the predicted metabolites of Isobavachalcone, which is obtained from the seeds of Psoralea corylifolia. Isobavachalcone has proven effective against papain-like protease (PL^{pro}) of SARS-CoV [57]. Isobavachalcone has also been found active against 3C-like protease/main protease ($3CL^{pro}/M^{pro}$) of the Middle East respiratory syndrome coronavirus (MERS-CoV) [58]. Thus, HMDB0132640 can be a potential compound for interventional therapy for COVID-19.

Another hit from HMDB is **HMDB0127868** (*Ligand*6) which belongs to the 5,7-Dihydroxy flavonoid class. It is a predicted metabolite of 2-(4-ethyl-3-hydroxyphenyl)-3,4-dihydro-2 h-1-benzopyran-3,5,7-triol. It is a polyphenolic compound and can have a potential antiviral property. Its binding energy and docking score are also good (− 10.33 kcal/mol, − 11.134, respectively). **HMDB0134119** (*Ligand*7) is our next hit and is a stilbene glycoside. It is a predicted metabolite of (E)-4-(3,5-dihydroxystyryl) benzene-1,3-diol. Some studies support the potential use of stilbene derivatives in treating SARC-COV infection [19, 59]. Its binding energy and docking score are good (− 9.53 kcal/mol, − 11.119, respectively). Another hit (**HMDB0125821/***Ligand*8) belongs to the 1-benzopyran class of compounds and are polyphenolic compounds**.** It is a predicted metabolite of 2,2-dimethyl-3,4-dihydro-2 h-1-benzopyran-4,5,7-triol. These benzopyran compounds are found in mushrooms and are the biomarker for consuming these foods. Its binding energy and docking score are good (− 10.28 kcal/mol, − 11.05, respectively).

Structural stability and flexibility of complexes

Overall $3CL^{pro}$ structure

About eight human metabolites were found to be potential after the virtual screening competition. These compounds were identified as interacting with the binding cavity and catalytic dyads of the SARS-CoV-2 major protease. To further validate the thermodynamics stability and flexibility of the complexes, 100 ns molecular dynamics simulations were carried out, and we monitored each system's structural and energetic properties during the production simulations.

To verify the convergence of simulations, we calculate the time evolution of our system's receptor backbone atoms root mean square deviation (RMSD) concerning their corresponding initial coordinates. The temporal distribution of RMSD of each system is shown in Fig. 2A, B, and their average values for the last 50 ns are listed in Table 2. We also ran two more replicas for all eight complexes, and the RMSD profile of the other two replica runs is shown in Supporting Information Fig. S2. It is evident from Fig. S2 that the overall RMSD profile was the same in all replica runs, suggesting a converged simulation for all but *complex*3. In the case of *complex*3, relatively large fluctuations were observed compared to other complexes. The other seven complex simulations attained stability in the last 50 ns.

Figure 2A, B shows that each system reached a good equilibrium state after 50 ns and maintained it throughout the last 50 ns of production simulations. The average RMSD value was found to vary between (1.5 ± 0.2) Å and (3.0 ± 0.1) Å for all cases. The highest average RMSD value (3.0 ± 0.1) Å was obtained for *complex*1 and the lowest value $(1.5 \pm 0.2$ Å) for *complex*4. For the comparison purpose, we

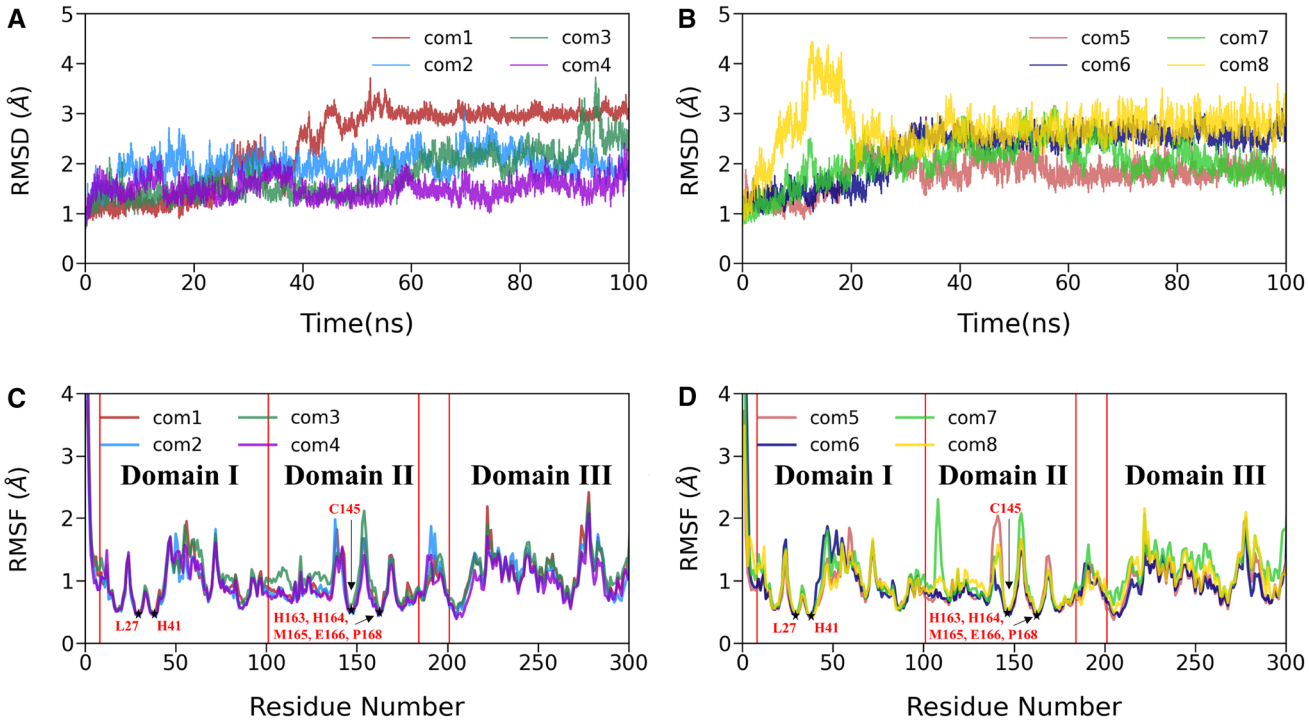


Fig. 2 **A**, **B** The time evolution of root means square deviation (RMSD) of backbone atoms of $3CL^{pro}$-complex relative to their respective initial coordinates. **C**, **D** The root mean square fluctuations (RMSFs) of C_α atoms for all eight $3CL^{pro}$-ligand complexes

Table 2 The average backbone RMSD, the radius of gyration (RoG), and solvent-accessible surface area (SASA) for all eight complexes

System	RMSD (Å)	RoG (Å)	SASA (nm^2)
*Complex*1	3.0 ± 0.1	21.9 ± 0.1	141.9 ± 2.6
*Complex*2	2.1 ± 0.2	22.0 ± 0.1	143.9 ± 2.5
*Complex*3	2.1 ± 0.4	22.1 ± 0.1	144.1 ± 2.7
*Complex*4	1.5 ± 0.2	22.0 ± 0.1	140.8 ± 2.3
*Complex*5	1.8 ± 0.2	21.9 ± 0.1	141.6 ± 2.4
*Complex*6	2.6 ± 0.1	21.8 ± 0.1	137.1 ± 2.3
*Complex*7	2.1 ± 0.3	22.0 ± 0.1	142.1 ± 3.0
*Complex*8	2.8 ± 0.2	21.8 ± 0.1	138.2 ± 2.4

The data are reported as average ± standard error of the mean (SEM)

also estimated the RMSD of the apo structure from our previous work, which was simulated under the same condition and shown in Supporting Information (Fig. S3A). The average RMSD for the apo form was 2.0 ± 0.01 Å, comparable to the RMSD value of complexes.

We extended our residual fluctuation study, and the root means square fluctuations (RMSFs) of C_α atoms in each protein complex were explored throughout the simulations and displayed in Fig. 2C, D. Figure 2C, D shows that all complexes' overall atomic fluctuations pattern is the same. Due to the inhibitor binding, we observed lower fluctuations for domain I (residues 8–101) and domain II (residues 102–184). On the other hand, domain III (residues 201–306) showed relatively higher fluctuations than domains I and II. RMSF values rarely crossed 2 Å for most of the C_α atoms, except terminal residues, which is a usual phenomenon. Further, it is evident from Fig. 2C, D that the ligand-binding sites, including Leu27, His41, Cys145, His163, His164, Met165, Glu166, and Pro168, displayed the lowest fluctuation. It can further be observed from Fig. 2C, D that the off-binding site residues like Ser46, Glu47, Leu50, and Pro52 from domain I; Asn151, Ile152, Asp153, Tyr154, and Asp155 from domain II; Met276, Asn277, Gly278, Arg279, Thr280, and Gly302 from domain II showed higher fluctuations compared to other residues. However, our previous study on SARS-CoV-2 $3CL^{pro}$ [60] suggested that the apo $3CL^{pro}$ protein structure had less atomic fluctuations than the ligand-bound protease, indicating that the apo $3CL^{pro}$ is less flexible (see Fig. S3B).

Ligand dynamics and binding pocket stability

After investigating the overall $3CL^{pro}$ structure, we also explored the conformation of all the ligands and the respective binding pocket stability in terms of its heavy atoms and backbone atoms RMSD, respectively, as shown in Fig. 3. As shown in Fig. 3A, B, the ligand RMSD values fluctuated within 3 Å, and the average RMSD values ranged from

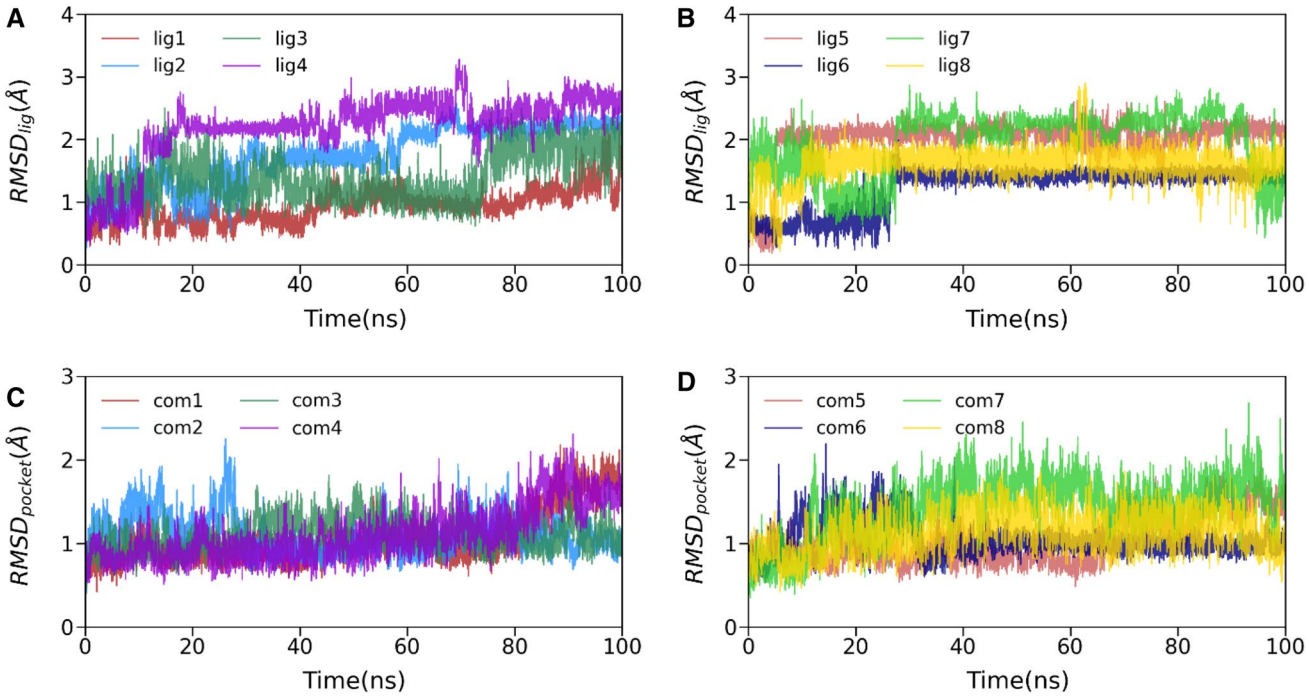


Fig. 3 **A**, **B** Time evolution of inhibitors' heavy atoms root mean square deviations (RMSDs) concerning initial conformations. **C**, **D** temporal RMSD variations of backbone atoms around the 5 Å of each ligand (binding pocket)

1.1 to 2.5 Å. Initially, all ligand RMSDs except for *ligand*3 increased up to 30 ns and stabilized afterward. In the case of *ligand*3, we observed more fluctuations than other ligands bound to $3CL^{pro}$, suggesting that it may make unstable interactions with the binding pocket of $3CL^{pro}$. This may weaken its affinity toward $3CL^{pro}$. Overall, the flexibility observed from the ligand RMSD profile is evident as these are small molecules. We also estimated the ligand stability by calculating the ligand–protein distance, which is discussed in the subsequent section. However, in the last 30 ns, all ligand RMSDs attain a steady state, which signifies a continuous binding.

Similarly, in Fig. 3C, D, we explore the time evolution of backbone atoms deviation at residues of 5 Å around each ligand. As the RMSD plot suggests, the fluctuations of all system's ligand-binding pockets are lower. The average RMSD values range from 1.0 to 2.0 Å. If we see the fluctuation pattern, all the complexes except (*complex*1 and *complex*4) reached stability after 20 ns. The pocket RMSD of *complex*1 and *complex*4 revealed the two complexes have almost similar behavior during the entire production simulation and achieved structural stabilities during the initial 80 ns and the final 15 ns. These results suggest that the binding pocket of $3CL^{pro}$ bound to human metabolites is relatively rigid and compact, which is suitable for better affinity.

Protein compactness and solvent exposure of binding sites

The residual compactness of the protein–ligand structure during molecular dynamics simulation is best described by the radius of gyration (RoG) and the solvent exposure measure in terms of solvent-accessible surface area (SASA). We computed RoG for all the complexes, shown in supporting information, Fig S4A, B, and the last 50 ns trajectories' average values are listed in Table 2. It is evident from Fig. S4A, B that all the complexes were stable and compact throughout the 100 ns simulations, and the average RoG values range from 21.8 to 22.1 Å for all cases. The initial RoG value for all is high compared to the last 50 ns trajectory. It may indicate that all complexes become more compact due to the binding of human metabolites.

In order to know the solvent exposure of the $3CL^{pro}$ binding cavity for all the eight simulated systems, we estimated the solvent-accessible surface area (SASA) of protein structures as shown in Fig. S4C, D. The average SASA values of the last 50 ns simulations are listed in Table 2. The initial surface area occupied by each complex is relatively high compared to the final 50 ns. It is evident from Table 2 that SASA values vary between 137.1 nm^2 and 144.1 nm^2 for all systems. On the other hand, the SASA value for the apo $3CL^{pro}$ was estimated as 146.5 nm^2 [60]. The lower SASA value signifies strong ligand-binding inside the cavity, suggesting the water molecule's displacement from it. A similar observation was found in earlier studies [61–63].

Human metabolites and 3CL^{pro} domains distance analysis

The 3CL^{pro} has divided into three different domains, see Fig. 1. If we talk about the binding of any small molecules on the 3CL^{pro}, then domain I and domain II are mainly responsible for that, and domain III is involved in the dimerization of 3CL^{pro}. Therefore, to explore the structural displacement of the binding pocket and ligand behavior inside the binding cavity, we calculate the center of mass (CoM) distance among human metabolites (ligands) and domains (mainly domain I and domain II) see Fig. 4. As suggested in Fig. 4, except for *complex*3, *complex*4, and *complex*5, the rest have shown no change in their distance plots. It is evident from Fig. 4D that the distance between ligand and domains initially decreases up to 37 ns; afterward, it increases up to 80 ns and finally reaches stable equilibrium and fluctuates around 20 Å for the last 20 ns of simulations. Similarly, we also see some drifting in the distance plot of *complex*3 and *complex*5. The binding of human metabolites to 3CL^{pro} is directly affected by these distance deflections.

Energetics of human metabolites affinity

To further elucidate the recognition, binding affinity, and specificity of human metabolites against SARS-CoV-2 main protease 3CL^{pro}, we have estimated the total binding affinity (ΔG_{bind})(ΔG_{bind}) and energetic components using molecular mechanics generalized Born surface area

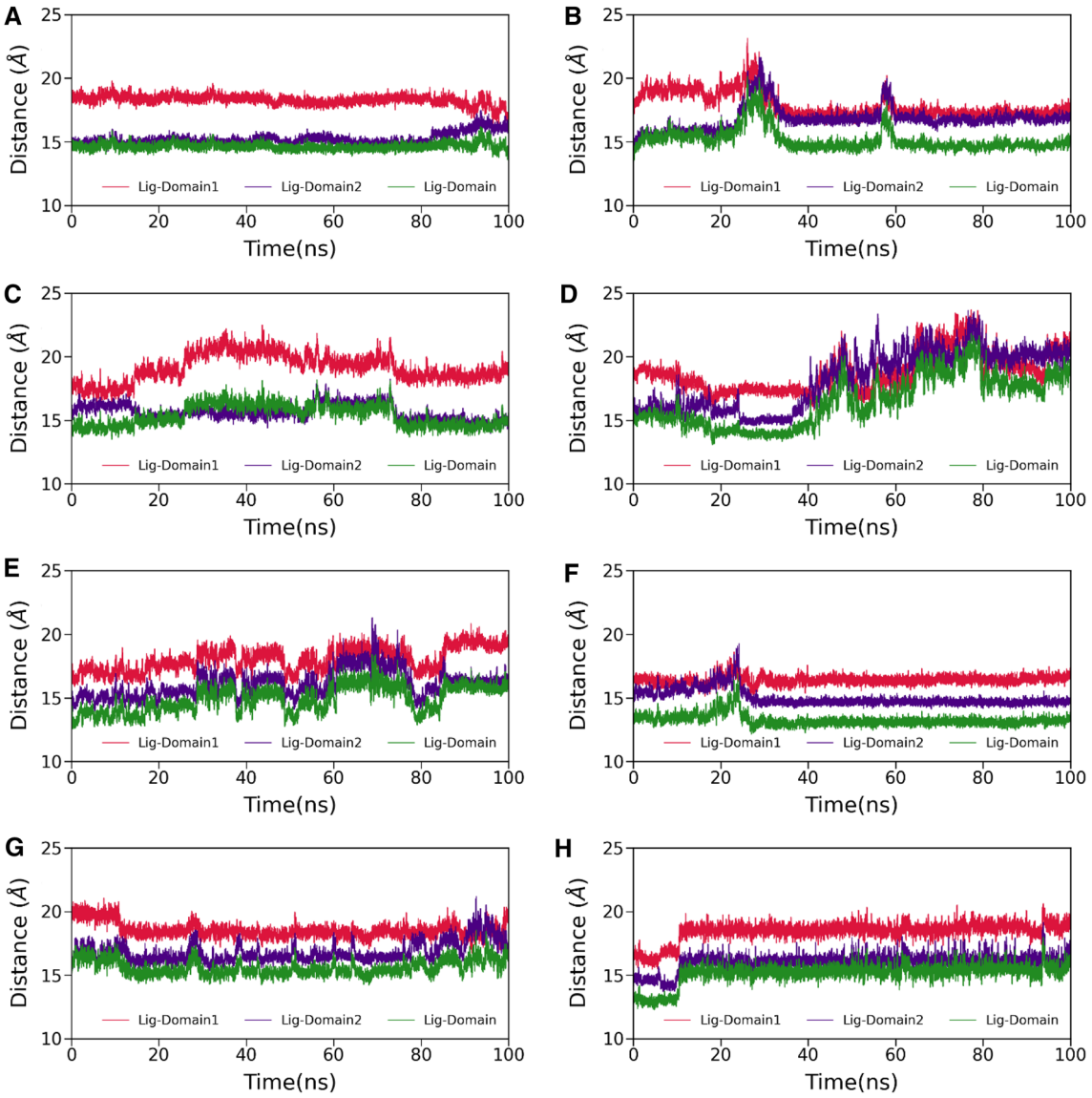


Fig. 4 The time evolution of distance between ligand and domain I (in red), ligand and domain II (in violet), ligand and domain I & II (in green). **A** *complex1*, **B** *complex2*, **C** *complex3*, **D** *complex4*, **E** *complex5*, **F** *complex6*, **G** *complex7* and **H** *complex8*, respectively

Table 3 Energetic components of the binding free energy for $3CL^{pro}$ of SARS-CoV-2 complexed with human metabolites, estimated using the MM/GBSA (kcal/mol) method

Systems	ΔE_{vdW}	ΔE_{elec}	ΔG_{pol}	ΔG_{np}	ΔE_{MM}[a]	ΔG_{solv}[b]	ΔH[c]	$-T\Delta S$	ΔG_{bind}[d]
*Complex*1	− 52.91 (0.10)	− 57.22 (0.35)	70.40 (0.30)	− 6.79 (0.01)	− 110.13 (0.39)	63.61 (0.29)	− 46.51 (0.11)	28.43 (0.86)	− 18.08 (0.87)
*Complex*2	− 44.26 (0.11)	− 45.62 (0.30)	62.87 (0.22)	− 6.34 (0.01)	− 89.88 (0.32)	56.53 (0.21)	− 33.35 (0.14)	28.01 (0.67)	− 5.34 (0.68)
*Complex*3	− 32.39 (0.10)	− 79.24 (0.32)	87.79 (0.29)	− 5.40 (0.01)	− 111.64 (0.35)	82.79 (0.29)	− 29.25 (0.10)	30.43 (0.70)	1.18 (0.71)
*Complex*4	− 29.93 (0.09)	− 23.86 (0.20)	38.44 (0.16)	− 4.46 (0.01)	− 53.80 (0.23)	33.98 (0.15)	− 19.82 (0.11)	23.99 (0.61)	4.17 (0.62)
*Complex*5	− 37.50 (0.11)	− 20.02 (0.30)	33.29 (0.26)	− 4.59 (0.01)	− 57.52 (0.34)	28.70 (0.25)	− 28.82 (0.13)	20.07 (0.57)	− 8.75 (0.58)
*Complex*6	− 48.53 (0.05)	− 27.72 (0.10)	48.09 (0.07)	− 5.89 (0.00)	− 76.25 (0.10)	42.20 (0.07)	− 34.05 (0.05)	23.67 (0.46)	− 10.33 (0.46)
*Complex*7	− 44.14 (0.11)	− 32.75 (0.22)	47.63 (0.16)	− 5.95 (0.01)	− 76.89 (0.28)	41.68 (0.15)	− 35.21 (0.16)	25.68 (0.84)	− 9.53 (0.85)
*Complex*8	− 33.95 (0.06)	− 17.63 (0.17)	28.26 (0.15)	− 4.18 (0.01)	− 51.58 (0.17)	24.09 (0.15)	− 27.49 (0.06)	17.21 (0.58)	− 10.28 (0.58)

Standard errors of the mean (SEM) are provided in parentheses

[a] $\Delta E_{vdW} + \Delta E_{elec}$

[b] $\Delta G_{pol+} \Delta G_{np}$

[c] $\Delta E_{vdW} + \Delta E_{elec} + \Delta G_{pol+} \Delta G_{np}$

[d] $\Delta E_{vdW} + \Delta E_{elec} + \Delta G_{pol+} \Delta G_{np} - T\Delta S$

(MM-GBSA) approach. A summary of the binding affinity and its components are shown in Supporting Information Fig. S5 and Table 3. As suggested in Fig. S5, the polar solvation (ΔG_{pol})(ΔG_{pol}) and configurational entropic ($T\Delta S$)($T\Delta S$) components disfavor the binding of human metabolites while the rest of the components ($\Delta E_{vdW}\Delta E_{vdW}$, ΔE_{elec} and $\Delta G_{np}\Delta E_{elec}$ and ΔG_{np}) are favorable toward the protein–ligand complexation. Moreover, Table 3 reveals that the calculated binding affinity of *complex*1 is the highest (− 18.08 ± 0.87 kcal/mol) followed by *complex*6 (− 10.33 ± 0.46 kcal/mol), *complex*8 (− 10.28 ± 0.58 kcal/mol), *complex*7 (− 9.53 ± 0.85 kcal/mol), *complex*5 (− 8.75 ± 0.58 kcal/mol) and *complex2* (− 5.34 ± 0.68 kcal/mol). On the other hand, the worst binding affinity (due to their high configurational entropic contributions, $-T\Delta S\Delta S$) of *complex*3 and *complex*4 suggest that these complexes' human metabolites are not good binder against $3CL^{pro}$. However, the static docking results are promising. It is worth noting that ligand3 displayed an increased configurational entropic contribution as it experienced more fluctuations than other ligands (see Fig. 3A). These results indicate that the *complex*1's human metabolite has the strongest binding affinity with $3CL^{pro}$.

Han et al. computationally explored the binding affinity of some clinically approved drugs (Chloroquine, Hydroxychloroquine, Remdesivir, Ritonavir, Favipiravir, Beclabuvir, and Indinavir) and recently designed α-ketoamide (13b) inhibitor against SARS-CoV-2 $3CL^{pro}$ [64]. Han et al. suggested that α-ketoamide (13b) is a highly potent inhibitor, and the binding affinity is − 17.75 kcal/mol. Overall, our estimations indicate that human metabolites could be more potent than α-ketoamide (13b), as we get better affinity (− 18.08 kcal/mol) against $3CL^{pro}$ due to its low value of configuration entropy (28.43 kcal/mol).

Critical residues involved in human metabolites binding with $3CL^{pro}$

To gain insights into the contribution of each amino acid to the binding, we decomposed the total binding free energy at the residual level using the molecular mechanics generalized Born surface area (MM-GBSA) approach. The protein–ligand interactions spectra are shown in Fig. 5. The binding hotspot residues whose contributions are more than − 1.5 kcal/mol are listed in Supporting Information Table S3. It is clear from Fig. 5 that the favorable binding residues are Met165, His41, Cys145, Glu166, Gln189, Asn142, Leu141, and Met49, and the interaction spectra for all are the same. This agrees with other computational studies [65, 66]. Our RMSF analysis also shows that these binding site residues display low fluctuations due to strong interactions with human metabolites (see Fig. 2C, D). Figure 5 and some recent studies suggest domains I and II are involved favorably in human metabolites like small molecules, and catalytic dyads His41 and Cys145 play an essential role. Moreover, for all cases, the assistance from the backbone and side-chain atoms are also favorable toward the binding, which is a good thing for an inhibitor's stronger affinity.

As Table S3 suggested that in the case of *complex*1, Gln189 (− 2.68 kcal/mol), Met165 (− 2.51 kcal/mol), Glu166 (− 2.10 kcal/mol), and Thr190 (− 1.51 kcal/mol)

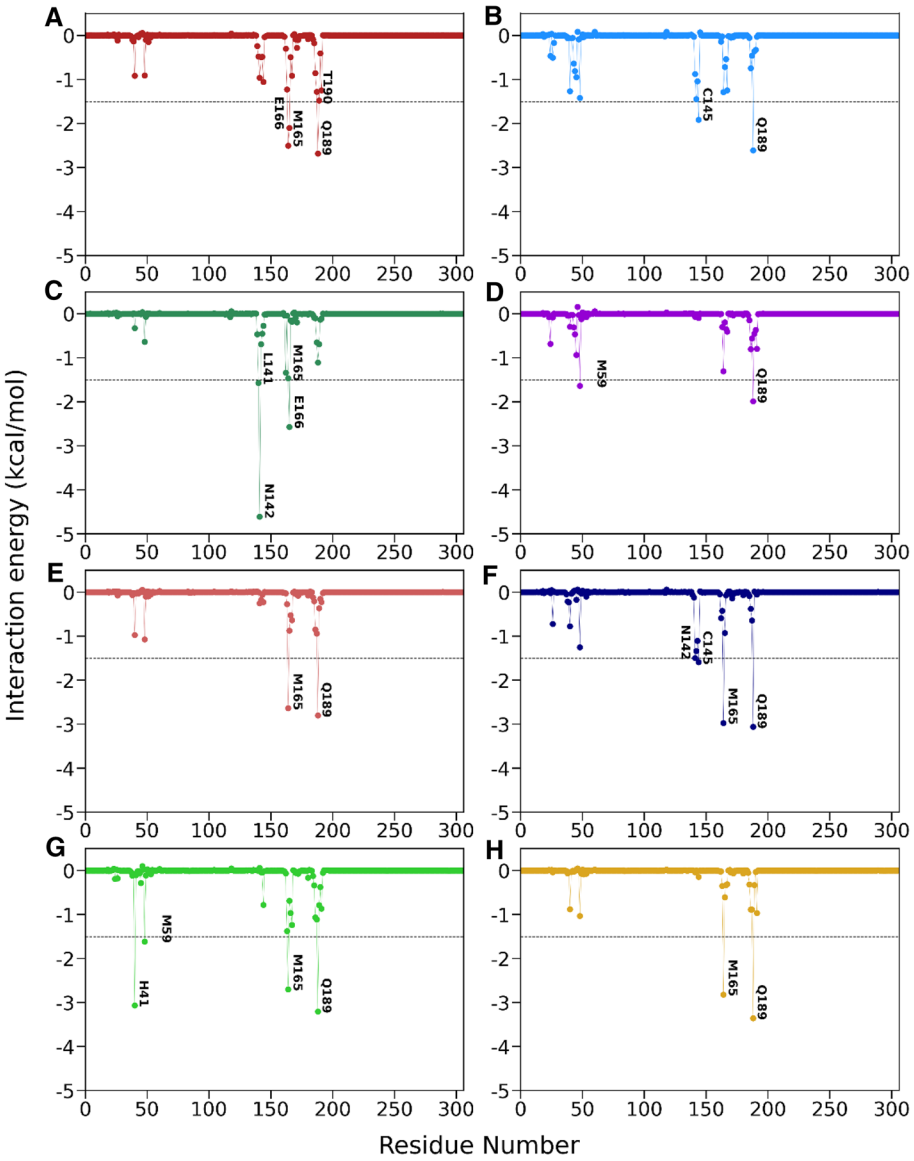


Fig. 5 Per-residue decomposition of the binding free energy for **A** *complex*1, **B** *complex*2, **C** *complex*3, **D** *complex*4, **E** *complex*5, **F** *complex*6, **G** *complex*7 and **H** *complex*8, respectively. The energy contribution ≥ − 1.5 kcal/mol is shown in each plot

are the hotspot residues. The van der Waals interactions with these residues have essential contributions to the overall binding affinity of the metabolite. This result implies that the human metabolite in *complex1* depicted a tendency to develop hydrophobic interactions with most of these binding site residues. This interaction is thought to reduce binding site residues' flexibility, as seen in the RMSF analysis (Fig. 2C, D). Similarly, residue Gln189 in *complex*4 (− 1.99 kcal/mol), *complex*5 (− 2.80 kcal/mol), *complex*6 (− 3.06 kcal/mol), *complex*7 (− 3.20 kcal/mol), and *complex*8 (− 3.35 kcal/mol) contributed more favorably than the catalytic dyads.

Hydrogen bonds and hydrophobic interactions stability analysis

To complement the energetic analysis and understand the complex stability, we determined the time evolution of the center of the mass distance between the $3CL^{pro}$ and eight human metabolites into their respective complex, shown in Fig. 6A, B. As suggested in Fig. 6A, B, the CoM distance of these complexes is at a stable equilibrium. Although we see some deflection in *complex*4 up to 80 ns, after that, it fluctuated around 23 Å.

Next, we estimated the temporal evolution of the number of hydrogen bonds between $3CL^{pro}$ and metabolites throughout the simulation (see Fig. 6C, D). Detailed profiling of the prominent hydrogen bonds is listed in Supporting Information Table S4. It is evident from Fig. 6C, D that *complex*3

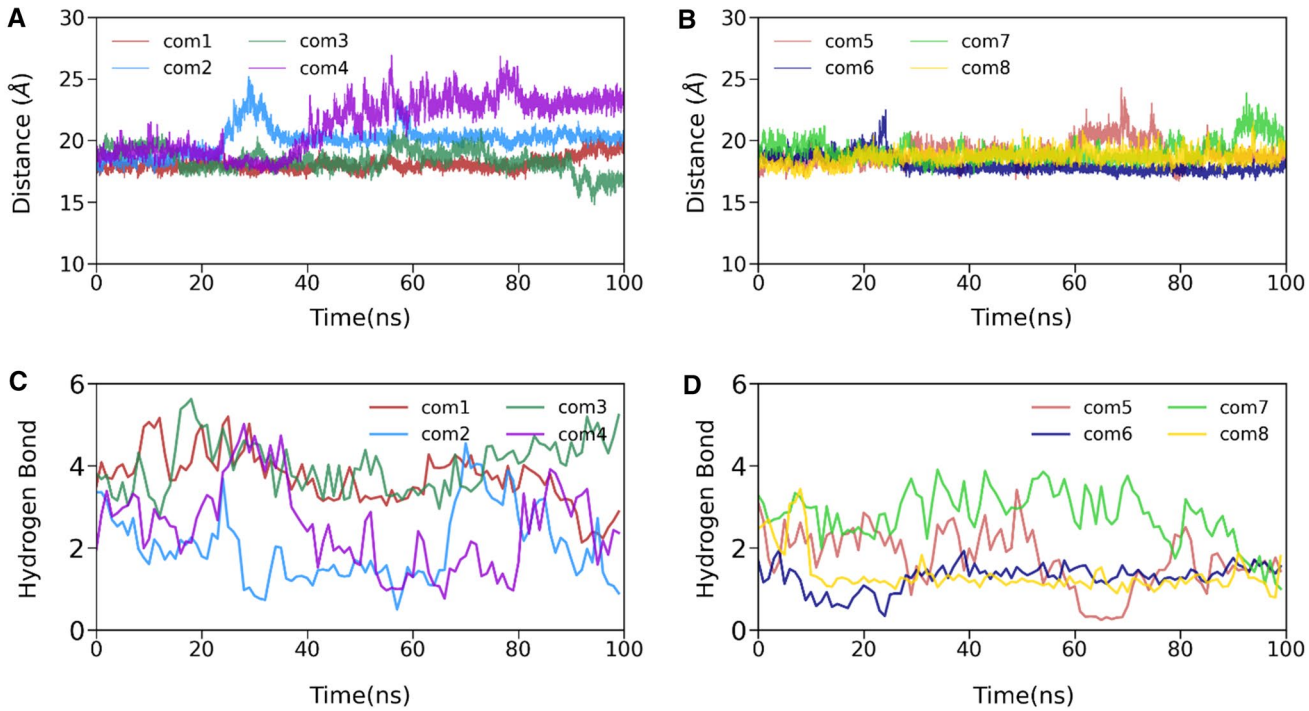


Fig. 6 The time evolution of the center of the mass distance (**A**, **B**) and the number of hydrogen bonds between $3CL^{pro}$ and human metabolites (**C**, **D**). In figure legend, the complex is termed as com in all cases

shows the highest number of hydrogen bonds compared to the rest of the complexes. This high number of hydrogen bonds increases the electrostatic interactions between $3CL^{pro}$ and human metabolites. However, in *complex*3, this electrostatic interaction is compensated by polar solvation energy, which is the highest among all complexes. Thus, the number of hydrogen bonds alone is insufficient for obtaining high affinity, as also shown in an earlier study [67]. However, *ligand*1 has very stable hydrogen between Gln192 and the oxygen (O3) atom (80.39%), which is missing in the case of *complex*3. The *complex*2, *complex*4, *complex*6, and *complex*8 have less than two hydrogen bonds in the 100 ns simulation. These results suggest that non-polar solvation and hydrophobic van der Waals interactions are the primary binding force.

Furthermore, we supplemented the above findings for the top four molecules by exploring various hydrogen bonding and hydrophobic interaction profile between the main protease, $3CL^{pro}$, and human metabolites for the final production simulation structure, as shown in Fig. 7. The interaction profiles for *complex*2, *complex*3, *complex*4, and *complex*5 are shown in Supplementary Information, Fig. S6. The 3D conformation of ligands in the binding site shows the possible orientation of the different ring structures, which leads to the possible set hydrogen bonding pattern (Fig. 7A–D). The detailed interaction profiles were estimated using $Ligplot^{+}$ (Fig. 7E–H) [68]. For the *complex*1, Fig. 7E displayed ten hydrophobic interactions with Cys44, Gly143, Tyr54, Gln189, Met49, Asp187, Glu166, Met165, Glm192, and Leu167. These extensive hydrophobic interactions account for the high affinity and stability of human metabolites in *complex1*. Similarly, residues, including Thr26, His41, Arg188, and Thr190, form strong hydrogen bonding with metabolites. In *complex*6, seven hydrophobic interactions with residues His163, Arg188, His41, His164, Asp187, Met165, and Gln189 were found, as revealed by Fig. 7F. Figure 7G, H shows that Ala191, Arg188, Met165, Gln189, His41, Met49, and Asn142 are strongly involved in hydrophobic interactions. His164, Gln189, Gln166, and Gln192 participate in the hydrogen bonding interactions for *complex*7 and *complex*8. Due to extensive hydrophobic interactions and strong hydrogen bonding, *complex1* may be more potent against SARS-CoV-2 $3CL^{pro}$.

Prediction of pharmacological and toxicological properties

Different physicochemical properties like CNS, Molecular weight, acceptor and donor hydrogen bond cut-off, octanol/water partition coefficient, brain–blood barrier coefficient, and others were calculated and listed in Supporting Information (Table S5). *Ligand8* is the least violated Lipinski's rule among all other top leads. The best lead ligand obtained from the MD simulation studies shows poor druggability

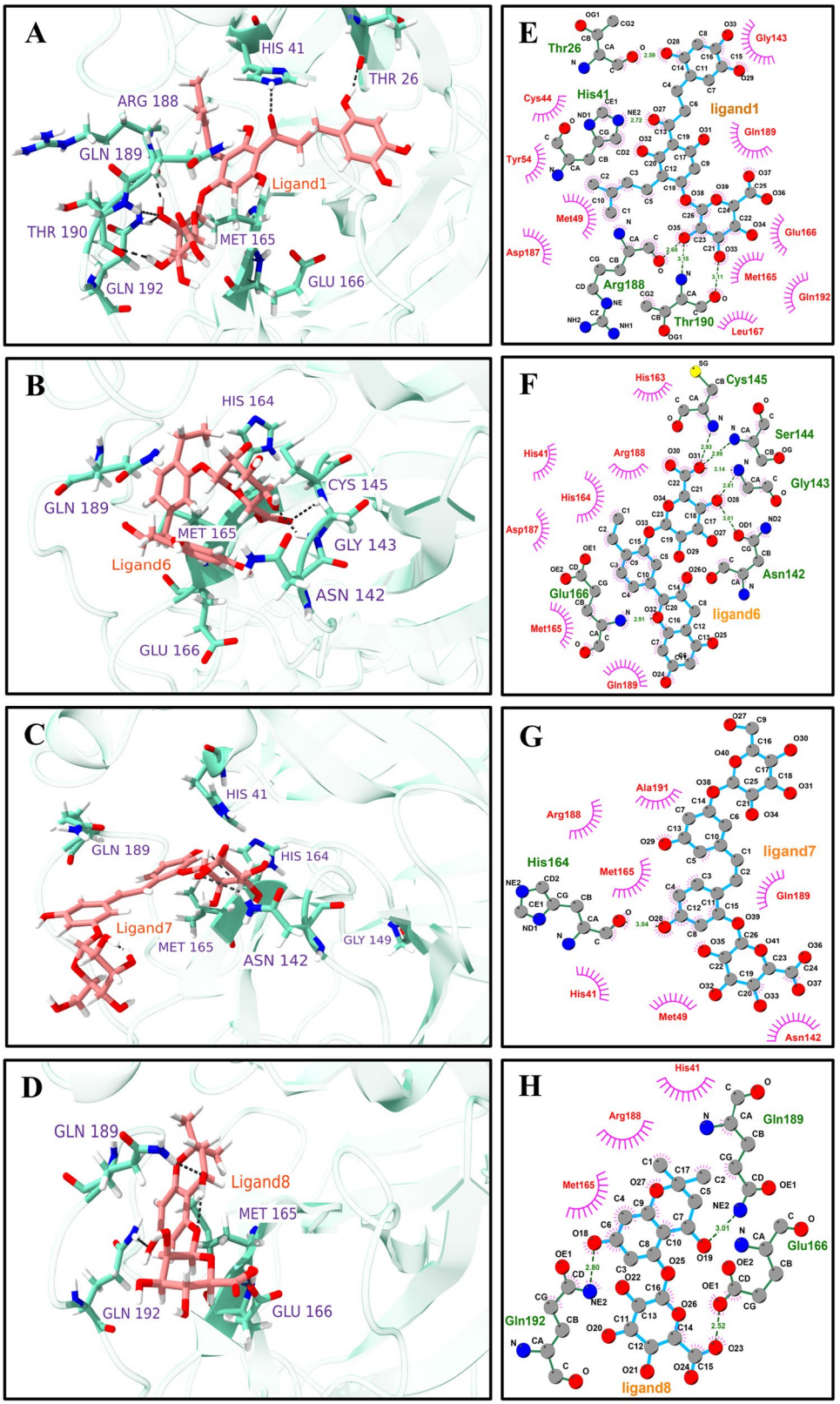


Fig. 7 The main protease, $3CL^{pro}$, and human metabolites interaction profiles (affinity more than − 9 kcal/mol). 3D interaction profile for selected ligands where key residues are shown ball and stick model **A** *Ligand*1, **B** *Ligand*6, **C** *Ligand*7, and **D** *Ligand*8. 2D interaction plots are generated by using Ligplot+, **E** *Ligand*1, **F** *Ligand*6, **G** *Ligand*7, and **H** *Ligand*8. Pink spike semicircles show the residues involved in the hydrophobic interactions, and hydrogen bonds are shown as green dotted lines

as it has the highest number of rule violations. However, recently, it has been found that many drugs do not follow Lipinski's rule and are presently found on the FDA-approved drug list [69]. Instead of being employed as a therapeutic candidate, these compounds will be used as lead molecules.

*Ligand*6 and *Ligand*8 have less than 500 Dalton molecular weights, indicating a promising lead molecule for drug design. Only in the case of human oral absorption, all four ligands were out of the acceptable ranges. Compared with all the parameters, *Ligand*8 stands out to be the best choice for a lead molecule. Toxicity profiling is one of the critical parameters for successful drug development, which was done in our lead molecules and shown in Supporting Information, Table S6. *Ligand*1 and *Ligand*7 were estimated as toxicity group 4, whereas *Ligand*6 and *Ligand*8 were estimated as group 5 on a scale of 1 to 5 (higher the number, lower the toxicity). All four lead ligands were evaluated as non-hepatotoxic, non-carcinogen, non-mutagenic, and non-cytotoxic. In comparison to *Ligand1*, which has the highest binding free energy of all the molecules, *Ligand8* exhibits several interesting features.

Conclusion

This study concludes that the natural compounds and their metabolites can play a promising role in managing the SARS-COV-2 infection. The current study is a reference model in which we recommend adjusting the body's metabolites or eating enough food to develop appropriate metabolites to combat viral infections. We screened the human metabolite database and selected the top 8 lead molecules for further validation. The chosen molecules were further screened using 3×100 ns long conventional molecular dynamics simulations and the free energy calculation using the MM/GBSA scheme. Based on our computational study, we identified top four lead compounds to become possible drug candidates. We found that one of the metabolites of isobavachalcone (*Ligand*1: HMDB0132640) binds very well (− 18.08 kcal/mol) to $3CL^{pro}$ of SARS-COV-2. Although it does not follow all the druggability rules, it is worth trying a molecule that can bind very well to $3CL^{pro}$ and may serve as the best lead for this target. On the other hand, *Ligand*8 is another suitable candidate based on suitable ADMET properties. Therefore, despite some deviations from drug-likeness, four of the eight hits can be taken further as they have ideal interactions with $3CL^{pro}$, excellent binding free energy, and good pharmacokinetic properties. Overall, these metabolites have a good chance of being developed as possible COVID-19 protease inhibitors.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11030-022-10513-6.

Acknowledgements This work was partially supported by the Department of Science and Technology, Govt. of India (grant number DST/NSM/R&D_HPC_Applications/2021/03.18). RR thanks the Indian Institute of Technology Indore for financial assistance. MFS would like to thank the Department of Science and Technology, Govt. of India, for providing a doctoral fellowship under the INSPIRE fellowship scheme (DST/INSPIRE Fellowship/2017/IF170145).

Data availability The data supporting this study's findings are available from the corresponding author (PK) upon reasonable request.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

Research involving human and animal rights There is no human or animal experiment in this study.

References

1. Wu JT, Leung K, Leung GM (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 395(10225):689–697. https://doi.org/10.1016/S0140-6736(20)30260-9
2. Xu X et al (2020) Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Sci China Life Sci 63(3):457–460. https://doi.org/10.1007/s11427-020-1637-5
3. Hui DS et al (2020) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 91:264–266. https://doi.org/10.1016/j.ijid.2020.01.009
4. Wu F et al (2020) A new coronavirus associated with human respiratory disease in China. Nature 579(7798):265–269. https://doi.org/10.1038/s41586-020-2008-3
5. Madboly WE, Shehata MG, Nashed MSM, Abu-Dief AM (2022) Using safe calculated low power of electrons to cut, analyze and exterminate the outer and inner biological elements of SARS-CoV-2, MERS-CoV-2 and influenza viruses in vitro. J Sci Res Rep 28(1):1–15. https://doi.org/10.9734/jsrr/2022/v28i130482
6. Woo PCY, Huang Y, Lau SKP, Yuen K-Y (2010) Coronavirus genomics and bioinformatics analysis. Viruses 2(8):1804–1820. https://doi.org/10.3390/v2081803
7. Jin Z et al (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature 582(7811):289–293. https://doi.org/10.1038/s41586-020-2223-y
8. Choudhary MI, Shaikh M, tul-Wahab A, ur-Rahman A (2020) In silico identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA predictive binding energy calculations, and molecular dynamics simulation. PLoS ONE 15(7):e0235030. https://doi.org/10.1371/journal.pone.0235030
9. Alamri MA et al (2021) Pharmacoinformatics and molecular dynamics simulation studies reveal potential covalent and FDA-approved inhibitors of SARS-CoV-2 main protease 3CLpro. J Biomol Struct Dyn 39(13):4936–4948. https://doi.org/10.1080/07391102.2020.1782768
10. Khan RJ et al (2021) Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2′-O-ribose methyltransferase. J Biomol Struct Dyn 39(8):2679–2692. https://doi.org/10.1080/07391102.2020.1753577

11. Sk MF, Roy R, Jonniya NA, Poddar S, Kar P (2021) 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 39(10):3649–3661. https://doi.org/10.1080/07391102.2020.1768149
12. Shi J, Wei Z, Song J (2004) Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. J Biol Chem 279(23):24765–24773. https://doi.org/10.1074/jbc.M311744200
13. Bzówka M, Mitusińska K, Raczyńska A, Samol A, Tuszyński JA, Góra A (2020) Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. Int J Mol Med 21(9):3099. https://doi.org/10.3390/ijms21093099
14. Thurakkal L, Singh S, Roy R, Kar P, Sadhukhan S, Porel M (2021) An in-silico study on selected organosulfur compounds as potential drugs for SARS-CoV-2 infection via binding multiple drug targets. Chem Phys Lett 763:138193. https://doi.org/10.1016/j.cplett.2020.138193
15. Roy R, Sk MF, Jonniya NA, Poddar S, Kar P (2021) Finding potent inhibitors against SARS-CoV-2 main protease through virtual screening, ADMET, and molecular dynamics simulation studies. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2021.1897680
16. Abu-Dief AM, El-Sagher HM, Shehata MR (2019) Fabrication, spectroscopic characterization, calf thymus DNA binding investigation, antioxidant and anticancer activities of some antibiotic azomethine Cu(II), Pd(II), Zn(II) and Cr(III) complexes. Appl Organomet Chem 33(8):e4943. https://doi.org/10.1002/aoc.4943
17. Singh S, Sk MF, Sonawane A, Kar P, Sadhukhan S (2021) Plant-derived natural polyphenols as potential antiviral drugs against SARS-CoV-2 *via* RNA-dependent RNA polymerase (RdRp) inhibition: an in-silico analysis. J Biomol Struct Dyn 16:6249–6264. https://doi.org/10.1080/07391102.2020.1796810
18. Ghosh R, Chakraborty A, Biswas A, Chowdhuri S (2021) Evaluation of green tea polyphenols as novel corona virus (SARS CoV-2) main protease (Mpro) inhibitors—an in silico docking and molecular dynamics simulation study. J Biomol Struct Dyn 39(12):4362–4374. https://doi.org/10.1080/07391102.2020.1779818
19. Wahedi HM, Ahmad S, Abbasi SW (2021) Stilbene-based natural compounds as promising drug candidates against COVID-19. J Biomol Struct Dyn 39(9):3225–3234. https://doi.org/10.1080/07391102.2020.1762743
20. Decuypere S et al (2012) Metabolomics based biomarker discovery for infectious diseases, the case of melioidosis. Int J Infect Dis 16:e216. https://doi.org/10.1016/j.ijid.2012.05.810
21. Kumar R, Ghosh M, Kumar S, Prasad M (2020) Single cell metabolomics: a future tool to unmask cellular heterogeneity and virus-host interaction in context of emerging viral diseases. Front Microbiol 11:1152. https://doi.org/10.3389/fmicb.2020.01152
22. Clish CB (2015) Metabolomics: an emerging but powerful tool for precision medicine. Cold Spring Harb Mol Case Stud 1(1):e000588. https://doi.org/10.1101/mcs.a000588
23. Wishart DS et al (2007) HMDB: the human metabolome database. Nucleic Acids Res 35:D521-526. https://doi.org/10.1093/nar/gkl923
24. Wishart DS et al (2018) HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 46(D1):D608–D617. https://doi.org/10.1093/nar/gkx1089
25. Gardiner P et al (1989) Spironolactone metabolism: steady-state serum levels of the sulfur-containing metabolites. J Clin Pharmacol 29(4):342–347. https://doi.org/10.1002/j.1552-4604.1989.tb03339.x
26. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 27(3):221–234. https://doi.org/10.1007/s10822-013-9644-8
27. Harder E et al (2016) OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. J Chem Theory Comput 12(1):281–296. https://doi.org/10.1021/acs.jctc.5b00864
28. Friesner RA et al (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47(7):1739–1749. https://doi.org/10.1021/jm0306430
29. Friesner RA et al (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. J Med Chem 49(21):6177–6196. https://doi.org/10.1021/jm051256o
30. Halgren TA et al (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47(7):1750–1759. https://doi.org/10.1021/jm030644s
31. Case DA et al (2018) AMBER 2018. University of California, San Francisco
32. Price DJ, Brooks CL (2004) A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys 121(20):10096–10103. https://doi.org/10.1063/1.1808117
33. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 11(8):3696–3713. https://doi.org/10.1021/acs.jctc.5b00255
34. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25(9):1157–1174. https://doi.org/10.1002/jcc.20035
35. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N ·log(N) method for Ewald sums in large systems. J Chem Phys 98(12):10089–10092. https://doi.org/10.1063/1.464397
36. Kräutler V, van Gunsteren WF, Hünenberger PH (2001) A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. J Comput Chem 22(5):501–508. https://doi.org/10.1002/1096-987X(20010415)22:5%3c501::AID-JCC1021%3e3.0.CO;2-V
37. Sk MF, Jonniya NA, Roy R, Poddar S, Kar P (2020) Computational investigation of structural dynamics of SARS-CoV-2 methyltransferase-stimulatory factor heterodimer nsp16/nsp10 bound to the cofactor SAM. Front Mol Biosci 7:590165. https://doi.org/10.3389/fmolb.2020.590165
38. Massova I, Kollman PA (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov Des 18(1):113–135. https://doi.org/10.1023/A:1008763014207
39. Kar P, Seel M, Hansmann UHE, Höfinger S (2007) Dispersion terms and analysis of size- and charge dependence in an enhanced Poisson−Boltzmann approach. J Phys Chem B 111(30):8910–8918. https://doi.org/10.1021/jp072302u
40. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10(5):449–461. https://doi.org/10.1517/17460441.2015.1032936
41. Wang C, Greene D, Xiao L, Qi R, Luo R (2018) Recent developments and applications of the MMPBSA method. Front Mol Biosci 4:87. https://doi.org/10.3389/fmolb.2017.00087
42. Wang E et al (2019) End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in

drug design. Chem Rev 119(16):9478–9508. https://doi.org/10.1021/acs.chemrev.9b00055
43. Sk MF, Haridev S, Roy R, Kar P (2021) Investigating potency of TMC-126 against wild-type and mutant variants of HIV-1 protease: a molecular dynamics and free energy study. SAR QSAR Environ Res 32(11):941–962. https://doi.org/10.1080/1062936X.2021.1999318
44. Chang S, Zhang D-W, Xu L, Wan H, Hou T-J, Kong R (2016) Exploring the molecular basis of RNA recognition by the dimeric RNA-binding protein via molecular simulation methods. RNA Biol 13(11):1133–1143. https://doi.org/10.1080/15476286.2016.1223007
45. Peng Y, Sun L, Jia Z, Li L, Alexov E (2018) Predicting protein-DNA binding free energy change upon missense mutations using modified MM/PBSA approach: SAMPDI webserver. Bioinformatics 34(5):779–786. https://doi.org/10.1093/bioinformatics/btx698
46. Roy R, Mishra A, Poddar S, Nayak D, Kar P (2022) Investigating the mechanism of recognition and structural dynamics of nucleoprotein-RNA complex from *Peste des petits ruminants virus* via Gaussian accelerated molecular dynamics simulations. J Biomol Struct Dyn 40(5):2302–2315. https://doi.org/10.1080/07391102.2020.1838327
47. Roy R, Ghosh B, Kar P (2020) Investigating conformational dynamics of Lewis Y oligosaccharides and elucidating blood group dependency of cholera using molecular dynamics. ACS Omega 5(8):3932–3942. https://doi.org/10.1021/acsomega.9b03398
48. Roy R, Jonniya NA, Sk MF, Kar P (2022) Comparative structural dynamics of isoforms of helicobacter pylori adhesin BabA bound to Lewis b hexasaccharide via multiple replica molecular dynamics simulations. Front Mol Biosci 9:852895. https://doi.org/10.3389/fmolb.2022.852895
49. Kollman PA et al (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33(12):889–897. https://doi.org/10.1021/ar000033j
50. Abu-Dief AM et al (2022) Synthesis, structural elucidation, DFT calculation, biological studies and DNA interaction of some aryl hydrazone Cr^{3+}, Fe^{3+}, and Cu^{2+} chelates. Comput Biol Chem 97:107643. https://doi.org/10.1016/j.compbiolchem.2022.107643
51. Abu-Dief AM et al (2021) Synthesis and intensive characterization for novel Zn(II), Pd(II), Cr(III) and VO(II)-Schiff base complexes; DNA-interaction, DFT, drug-likeness and molecular docking studies. J Mol Struct 1242:130693. https://doi.org/10.1016/j.molstruc.2021.130693
52. Abu-Dief AM, Abdel-Rahman LH, Abdel-Mawgoud AAH (2020) A robust in vitro anticancer, antioxidant and antimicrobial agents based on new metal-azomethine chelates incorporating Ag(I), Pd (II) and VO (II) cations: probing the aspects of DNA interaction. Appl Organomet Chem 34(2):e5373. https://doi.org/10.1002/aoc.5373
53. AbdelRahman LH, Abu-Dief AM, El-Khatib RM, Abdel-Fatah SM, Adam AM, Ibrahim EMM (2018) Sonochemical synthesis, structural inspection and semiconductor behavior of three new nano sized Cu(II), Co(II) and Ni(II) chelates based on tri-dentate NOO imine ligand as precursors for metal oxides. Appl Organomet Chem 32(3):e4174. https://doi.org/10.1002/aoc.4174
54. Banerjee P, Eckert AO, Schrey AK, Preissner R (2018) ProTox-II: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res 46(W1):W257–W263. https://doi.org/10.1093/nar/gky318
55. Trott O, Olson AJ (2009) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31(2):455–461. https://doi.org/10.1002/jcc.21334
56. Solnier J, Fladerer J-P (2020) Flavonoids: a complementary approach to conventional therapy of COVID-19? Phytochem Rev 20:773–795. https://doi.org/10.1007/s11101-020-09720-6
57. Kim DW et al (2014) Phenolic phytochemical displaying SARS-CoV papain-like protease inhibition from the seeds of Psoralea corylifolia. J Enzym Inhib Med Chem 29(1):59–63. https://doi.org/10.3109/14756366.2012.753591
58. Jo S, Kim H, Kim S, Shin DH, Kim M-S (2019) Characteristics of flavonoids as potent MERS-CoV 3C-like protease inhibitors. Chem Biol Drug Des 94(6):2023–2030. https://doi.org/10.1111/cbdd.13604
59. Li Y-Q, Li Z-L, Zhao W-J, Wen R-X, Meng Q-W, Zeng Y (2006) Synthesis of stilbene derivatives with inhibition of SARS coronavirus replication. Eur J Med Chem 41(9):1084–1089. https://doi.org/10.1016/j.ejmech.2006.03.024
60. Sk MF, Roy R, Kar P (2021) Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. J Biomol Struct Dyn 39(3):988–1003. https://doi.org/10.1080/07391102.2020.1724196
61. Nand M et al (2020) Virtual screening of anti-HIV1 compounds against SARS-CoV-2: machine learning modeling, chemoinformatics and molecular dynamics simulation based analysis. Sci Rep 10(1):20397. https://doi.org/10.1038/s41598-020-77524-x
62. Kumar S et al (2020) Discovery of new hydroxyethylamine analogs against $3CL^{pro}$ protein target of SARS-CoV-2: molecular docking, molecular dynamics simulation, and structure-activity relationship studies. J Chem Inf Model 60(12):5754–5770. https://doi.org/10.1021/acs.jcim.0c00326
63. Gupta S et al (2021) Identification of potential natural inhibitors of SARS-CoV2 main protease by molecular docking and simulation studies. J Biomol Struct Dyn 39(12):4334–4345. https://doi.org/10.1080/07391102.2020.1776157
64. Han Y, Wang Z, Ren J, Wei Z, Li J (2021) Potential inhibitors for the novel coronavirus (SARS-CoV-2). Briefings Bioinform 22(2):1225–1231. https://doi.org/10.1093/bib/bbaa209
65. Fiorucci D, Milletti E, Orofino F, Brizzi A, Mugnaini C, Corelli F (2020) Computational drug repurposing for the identification of SARS-CoV-2 main protease inhibitors. J Biomol Struct Dyn 39(16):6242–6248. https://doi.org/10.1080/07391102.2020.1796805
66. Pant S, Singh M, Ravichandiran V, Murty USN, Srivastava HK (2021) Peptide-like and small-molecule inhibitors against Covid-19. J Biomol Struct Dyn 39(8):2904–2913. https://doi.org/10.1080/07391102.2020.1757510
67. Kar P, Lipowsky R, Knecht V (2013) Importance of polar solvation and configurational entropy for design of antiretroviral drugs targeting HIV-1 protease. J Phys Chem B 117(19):5793–5805. https://doi.org/10.1021/jp3085292
68. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 51(10):2778–2786. https://doi.org/10.1021/ci200227u
69. Mullard A (2018) Re-assessing the rule of 5, two decades on. Nat Rev Drug Discov 17(11):777–777. https://doi.org/10.1038/nrd.2018.197

Authors and Affiliations

Rajarshi Roy[1] · Md Fulbabu Sk[1] · Omprakash Tanwar[2] · Parimal Kar[1]

[1] Department of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore, Madhya Pradesh 453552, India

[2] Department of Pharmacy, Shri G. S. Institute of Technology and Science, Indore, Madhya Pradesh 452003, India