



Meticulous assessment of natural compounds from NPASS database for identifying analogue of GRL0617, the only known inhibitor for SARS-CoV2 papain-like protease (PLpro) using rigorous computational workflow

Paritosh Parmar¹ · Priyashi Rao² · Abhilasha Sharma³ · Arpit Shukla^{1,4} · Rakesh M. Rawal^{2,3} · Meenu Saraf¹ · Baldev V. Patel¹ · Dweipayan Goswami¹

Received: 18 April 2021 / Accepted: 5 May 2021 / Published online: 18 May 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

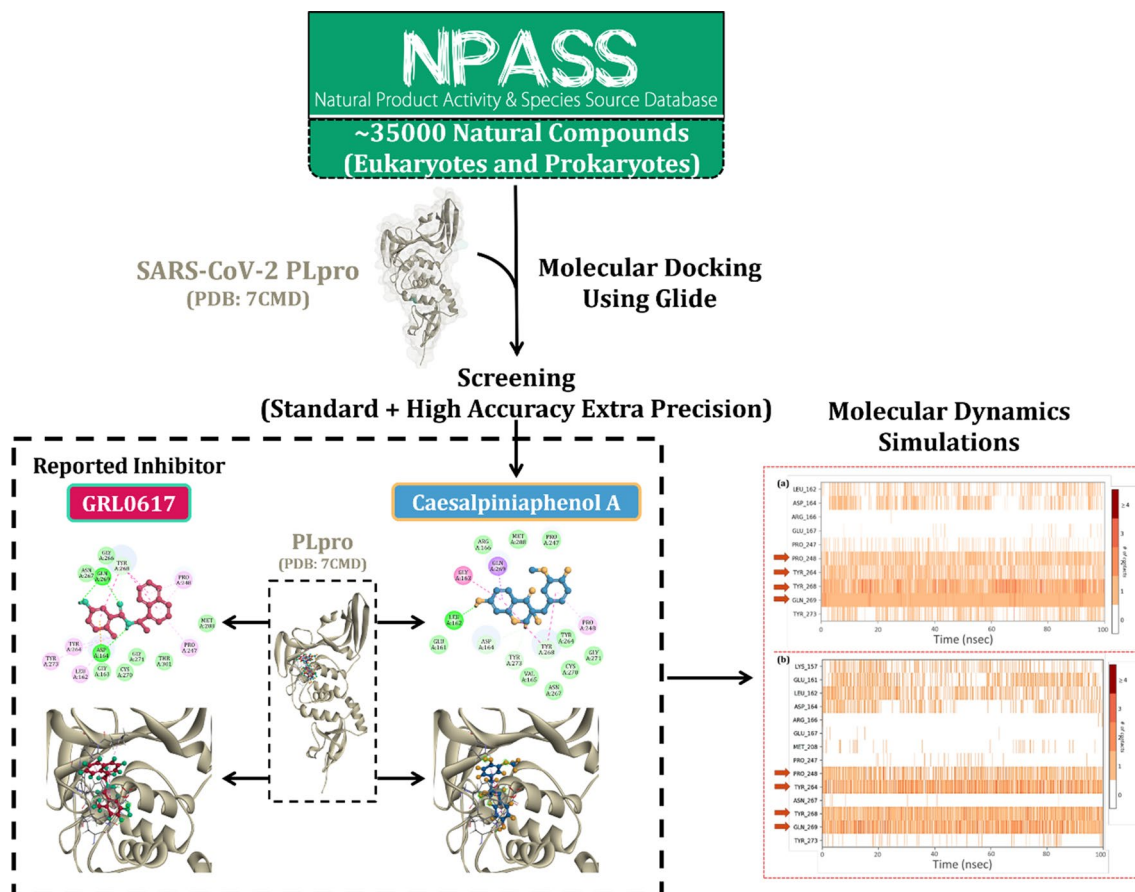
Abstract

The latest global outbreak of 2019 respiratory coronavirus disease (COVID-19) is triggered by the inception of novel coronavirus SARS-CoV2. If recent events are of any indicators of the epidemics of past, it is undeniable to state a fact that the SARS-CoV2 viral infection is highly transmissible with respect to its previously related SARS-CoV's. Papain-like protease (PLpro) is an enzyme that is required by the virus itself for replicating into the host system; and it does so by processing its polyproteins into a functional replicase complex. PLpro is also known for downregulating the genes responsible for producing interferons, an essential family of molecules produced in response to viral infection, thus making this protein an indispensable drug target. In this study, PLpro inhibitors were identified through high throughput structure-based virtual screening approach from NPASS natural product library possessing ~35,000 compounds. Top five hits were scrutinised based on structural aromaticity and ability to interact with a key active site residue of PLpro, Tyr268. For second level of screening, the MM-GBSA End-Point Binding Free Energy Calculation of the docked complexes was performed, which identified Caesalpinaphenol A as the best hit. Caesalpinaphenol A not only possess a double ring aromatic moiety but also has lowest minimum binding energy, which is at par with the control GRL0617, the only known inhibitor of SARS-CoV2 PLpro. Details of the Molecular Dynamics (MD) simulation and ADMET analysis helped to conclusively determine Caesalpinaphenol A as potentially an inhibitor of SARS-CoV2 PLpro.

✉ Dweipayan Goswami
dweipayan.goswami@gujaratuniversity.ac.in

- ¹ Department of Microbiology & Biotechnology, University School of Sciences, Gujarat University, Ahmedabad, Gujarat 380009, India
- ² Department of Biochemistry & Forensic Science, University School of Sciences, Gujarat University, Ahmedabad, Gujarat 380009, India
- ³ Department of Life Science, University School of Sciences, Gujarat University, Ahmedabad, Gujarat 380009, India
- ⁴ Department of Biological Sciences and Biotechnology, University of Innovation, Institute of Advanced Research, Koba Institutional Area, Gandhinagar, Gujarat 382426, India

Graphic abstract



Keywords SARS-CoV2 novel corona virus · Papain-like protease (PLpro) · Bioactive metabolites · High throughput virtual screening · Computational drug prediction

Introduction

SARS-CoV2 the causative agent of twenty-first century pandemic has resulted in incalculable amount of damage since its inception [1–3]. Also known as COVID-19; continues to expand worldwide and as World Health Organization’s (WHO) most recent situation report, dated March 30, 2021, it has accounted for > 128 million cases and > 2.79 million fatalities so far (<https://covid19.who.int/>). Severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV) and SARS-CoV-2 are three of the highly pathogenic and lethal human coronaviruses that have arisen in the last two decades belonging to the same family. The small re-emergence of SARS in late 2003, another deathful emergence of MERS-CoV in 2012 and presently the exposure to SARS-CoV-2 of 2019 has awoken the scientific community across the globe to intently get involved

in discovering a form of immunisation or a candidate drug to put an end to this remorseful condition [4].

SARS-CoV-2 as named by the International Committee on Taxonomy of Viruses (ICTV), is a type of evolutionary enveloped virus with a single stranded, positive sense (+ ssRNA) genomic RNA of type Beta coronaviruses belonging to the Coronaviridae family, known to have the largest genome for an RNA virus of approximately 26–32 kb in size [5, 6]. The symptoms usually are identical to other respiratory virus infections like influenza or pneumonia however, the viral load and spread is far greater as what has been observed with previous coronaviruses. This particular virus enters the cell by lodging itself on an angiotensin-converting enzyme 2 (ACE2) receptor. Spike protein, a special surface glycoprotein that binds the virus to human ACE2, is used to penetrate the host cell. Entry of virion into the human host cell, allows the translation of 5’ open reading frames which involves ORF1a and ORF1ab

to yield two large polyproteins, pp1a and pp1ab [6]. These polyproteins are then processed by the papain-like protease (PLpro), and 3-chymotrypsin-like main protease (3CLpro/Mpro) [5, 7, 8]. PLpro further cleaves nsp1, nsp2 and nsp3 while, 3CLpro processes the remaining 13 non-structural proteins, required for formation of a replicase complex on the host membrane to carry out initiation of replication and transcription of the viral genome. Both of these proteolytic enzymes have an added function where they tend to dysregulate the innate immune response for the host cell, by inhibiting the hosts initial inflammatory response and subsequent interferon response (Fig. 1). Upon this dysregulation, the interferon system is unable to generate an antiviral state as an innate immune response through transcriptional upregulation of more than 300 Interferon-Stimulated Genes (ISGs), which normally can efficiently detect and generate a strong response to viral threats [9–13]. This regulation impairment often leads to generation of a cytokine storm that has been correlated with multiple organ damages resulting in comorbidities. Most significantly, PLpro has ability to efficiently remove ISG15 (called as deISGylation) and other ubiquitin-like modifications. Ubiquitination is a process where several

Ubiquitin (Ub) chains are added to lysine residues of a protein as part of post-translational modification for proteasomal degradation. While ISGylation is a process where ISG's are conjugated to a target protein induced by IFN-I to regulates its function and is known to inhibit viral replication. Both ubiquitination and ISGylation play crucial roles in the regulation of innate immune responses to viral infection, weakening the overall inflammation and antiviral signalling mechanism of the human host [14]. Suggesting that these proteolytic enzymes, specifically PLpro is an important target for viral replication suppression and inhibition of PLpro would ultimately cause the SARS-CoV-2 proteases to lose their ability to thrive in the host cells. As a result, PLpro has been recognized and concerted in this study as a target protein [14].

In-vitro studies on IFN-I stimulated lung cancer cell line has brought out the conclusion that GRL0617 can effectively inhibit the deubiquitination and deISGylation activities of SARS-CoV-2-PLpro and also is able to restore the IFN-I response [15]. Another study based on multiplicity of transfections on Vero E6 cell lines revealed GRL0617 to completely inhibit viral replication with no apparent cytotoxicity to Vero E6 cells. Proving that a 100 μ M aliquot can strongly

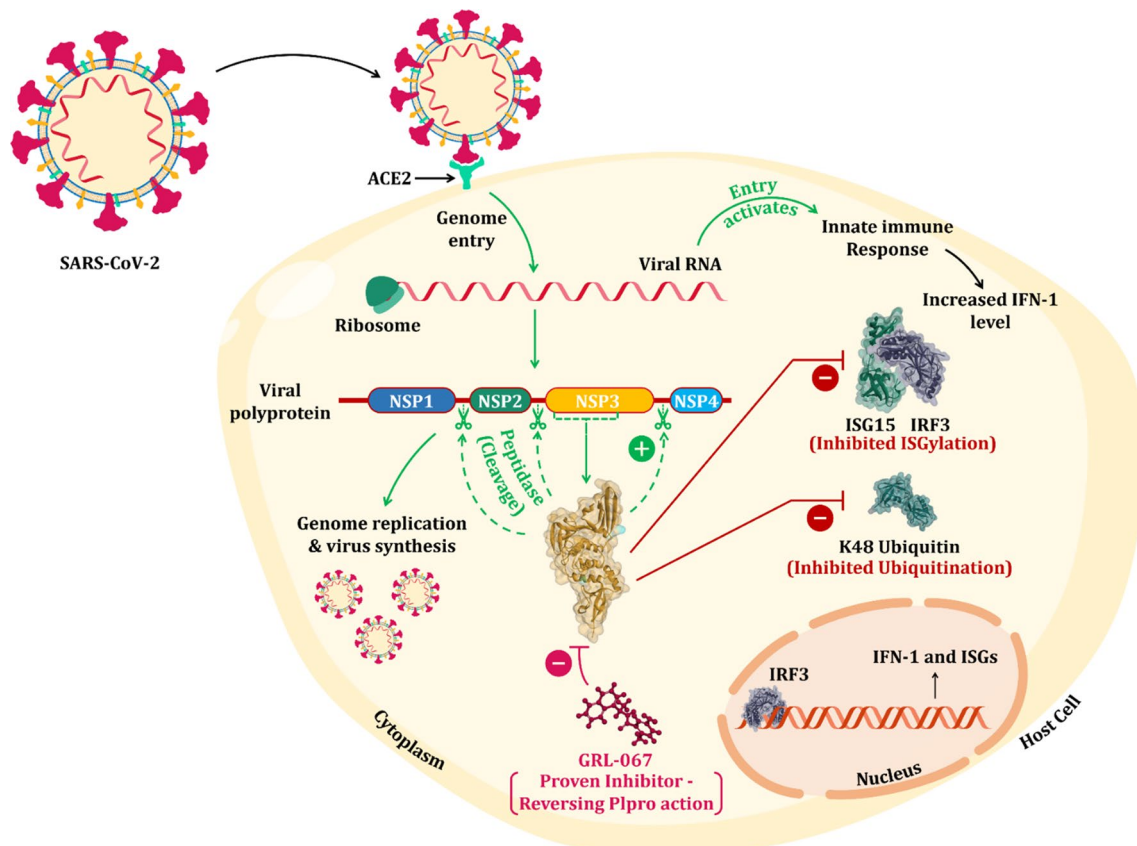


Fig. 1 Outline portraying functionality and mechanism of increase of innate immune response by IFN1 and IRF3 response. Figure also depicts the inhibition of ISGylation and Ubiquitination by PLpro. Image is prepared using Office 365 PowerPoint

inhibit the viral replication up to 50%. Furthermore, in-silico studies have revealed that the naphthalene-based GRL0617 interacts with Tyr269 and Tyr268 of SARS-CoV-PLpro and SARS-CoV2-PLpro, respectively [14]. Although the same inhibitor is not quite effective against MERS-CoV- PLpro, which involves a Thr amino acid residue instead of a Tyr residue in SARS at its conserved location. The near Tyr residue similarity in between the two SARS may be owned to the fact that both SARS-CoV's and SARS-CoV-2's PLpro share 82.9% sequence identity and a 99.9% sequence identity for their binding site allowing to incorporate small compounds. This qualitative similarity suggests similar assessment of activity of GRL0617 towards either substrate of SARS [15]. Recently we have postulated Fonseca, a fungal metabolite to be the structural–functional analogue of GRL0617 [16]. However, the current manuscript is the extension of our previous study and the detailed rationale for undertaking current theoretical study is expressed as follows.

Due to urgent need to overcome the pandemic of COVID-19, advancements in Structure based virtual screening have become a crucial component in modern day drug discovery. Accordingly, in the present study we describe our efforts to screen a 35,000-compound library of natural products retrieved from NPASS (Natural Product Activity and Species Source) to screen an analogue of naphthalene based GRL0617, an only known inhibitor of SARS-CoV2-PLpro. NPASS is an open-source database (available at <http://bidd2.nus.edu.sg/NPASS/>) encompassing a total of 35,032 bioactive molecules from 25,041 species (includes both prokaryotes and eukaryotes) known to interact with 5863 targets (that includes, 2946 proteins, 1352 microbial species and 1227 cell-lines). Further, NPASS also contains 446,552 records of biologically quantitative activity (e.g., IC₅₀, Ki, EC₅₀, GI₅₀ or MIC mainly in units of nM) of 222,092 bioactive compound-target pairs and 288,002 bioactive compound-species pairs [17]. The computational strategy employed in this study includes molecular docking to perform a high throughput structure-based virtual screening of all the 35,032 bioactive molecules from NPASS with SARS-CoV2-PLpro. Based on molecular docking scores, molecular interactions with the active site catalytic residue Tyr268 and structural aromaticity; the lead compounds were assigned. Moreover, while screening, the compounds with molecular weight and number of cyclic rings identical to GRL0617 were chosen with a goal to identify a natural substitute as a structural and functional analogue. Further, the End-Point Binding Free Energy Calculation through MM-GBSA, Molecular Dynamic (MD) simulation and pharmacokinetic ADMET (Adsorption, Distribution, Metabolism, Excretion and Toxicological) prediction aided in identifying a top hit compound. Further, the respective values of lead compound obtained during each computational assessment

were compared with GRL0617 to get an accurate prediction of the potency of identified lead with reference to GRL0617. The workflow of the tasks performed to fulfil the rationale of the current manuscript is presented in Fig. 2.

Materials and methods

Preparation of proteins and ligands

The structure of SARS-CoV-2 viral polyprotein PLpro, co-crystalized with inhibitor GRL0617 was retrieved from Protein Data Bank (PDB) (<https://www.rcsb.org/>); accession ID:7CMD. Protein preparation involved removing unwanted molecules which adhered to the protein structure during X-ray Crystallography (e.g., waters, other impurities); correcting charges, assigning bond orders, creating disulphide bonds, filing missing side chain residues and converting selenomethionine to methionine. After pre-processing the protein and inspecting the protein reports, the structure was optimised and minimised with default parameters under OPLS 2005 (Optimized Kanhesia for Liquid Simulations) force field [18–20]. These tasks were all performed using the Protein Preparation Wizard of Schrödinger Maestro [21, 22].

All the 35,032 natural compound structures retrieved from the NPASS database were downloaded in 3D format for virtual screening. Ligand preparation helps in generating the low energy structures and allow the option to expand each input's structure according to its desired stereochemistry by generating variations on ionisation state tautomer's ad ring confirmations. LigPrep wizard in Schrödinger Maestro was used to generate ionization states for each ligand structure with Epik [20, 21] at a physiological pH of 7.2 ± 0.2 unit. Rest other options were kept as default and the ligands were minimized under OPLS2005 force field.

Structure based virtual screening

Virtual screening of 35,032 natural compounds from the NPASS database was carried out in three phases: (a) High throughput virtual screening (HTVS) to screen vast number of ligands in a rapid way. Then, the top 25 of the best docked poses were subjected to (b) Standard Precision (SP) and (c) Extra precision (XP) which are known to be robust, discriminate and allow more time for docking in comparison to HTVS with higher torsional refinement and extensive sampling. These tasks were all performed Glide [21, 22]. Top 5 screened compounds were screened further on the basis of docking score range, molecular interactions with the key residue Tyr268 and structural aromaticity.

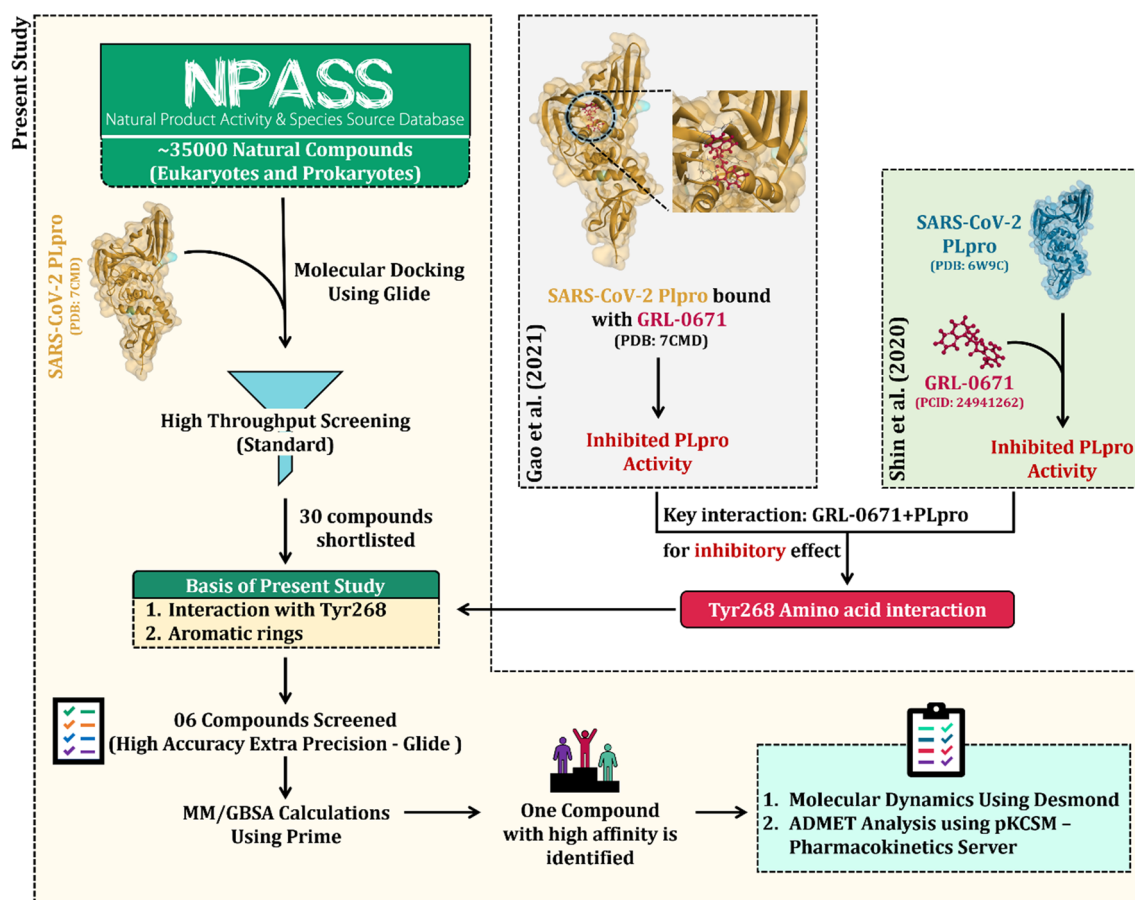


Fig. 2 Pictorial workflow of the tasks performed to fulfil the rationale

End-point binding free energy change calculation using MM-GBSA

The end point binding free energy change was calculated using Molecular Mechanics Generalized Born Surface Area (MM-GBSA) [23–26]. The docked complexes were optimized using the Prime wizard of Schrödinger Maestro’s local optimization function [27]. The binding free energy change for a group of receptors and ligands was calculated using the OPLS-2005 force field. The binding free energy transition was calculated using the following equation:

$$\Delta G_{\text{Bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{Solv}} + \Delta G_{\text{SA}} \quad (1)$$

Here, ΔG_{Bind} stands for the binding of receptor and ligand molecules in solution as the molar Gibbs free energy change. ΔE_{MM} is the variance between the minimized energy of the protein–ligand complexes, while ΔG_{Solv} is the sum of the solvation energies for the protein and ligand and the variation between the GBSA solvation energy of the same. ΔG_{SA} is the difference in the surface area energies for the complexes.

Molecular dynamic simulation

The simulations for SARS-CoV2-PLpro in presence of the known inhibitor GRL0617 and simulations for SARS-CoV2-PLpro-top hit compound from NPASS were performed for a period of 100 ns each for the docked complex using Desmond package (Schrödinger Release 2018-4) [28]. The 100 ns MD simulation of SARS-CoV2-PLpro-GRL0617 docked complex, was performed to consider as control.

The system was built using TIP3P solvent model which specifies a 3-site rigid water molecule with charges and Lennard–Jones parameters assigned to each of the 3 atoms. Periodic boundary conditions (PBC) were setup by selecting the orthorhombic shape simulation box with the dimension of $10 \text{ \AA} \times 10 \text{ \AA} \times 10 \text{ \AA}$ with default box angles and box volume (\AA^3). Followed by neutralisation with placement of 3 Na^+ ions and salt concentration of 0.15 M Na^+ and Cl^- counter ions to simulate the background salt and physiological conditions using OPLS2005 force fields. Once the system gets incorporated, it is minimized with restraints using Steepest descent energy minimization with NPT (constant Number of particles, Pressure, and Temperature), 300 K temperature

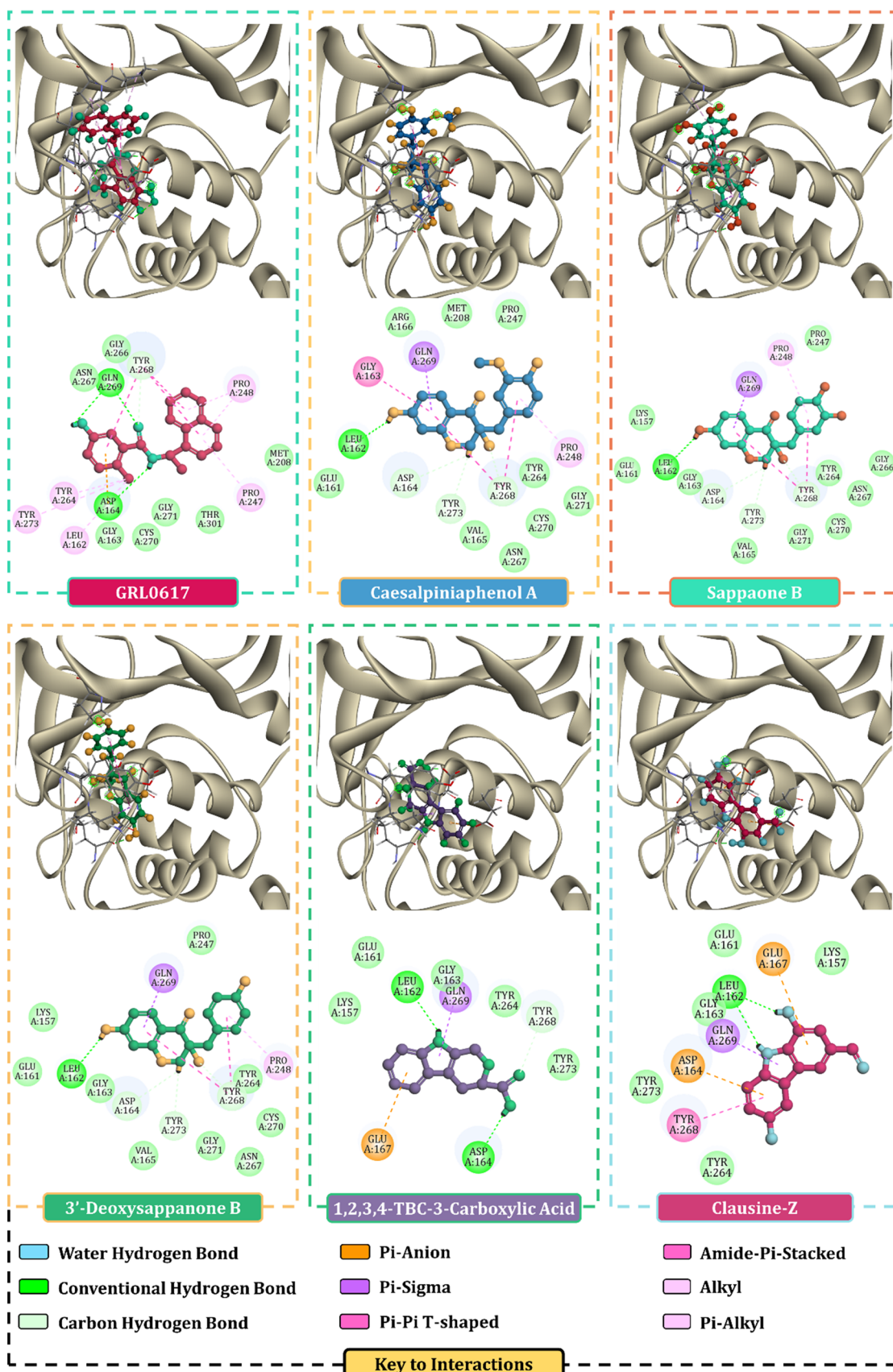


Fig. 3 Interaction profile for top five natural compounds from NPASS database, along with established inhibitor-GRL0617 in docked complex with SARS-CoV2-PLpro (PDB: 7CMD)

and 1.013 bar atomic pressure and default surface tension using Smooth Particle Mesh Ewald (PME) method to neutralise the electrostatic interactions. The MD simulation was executed for a period of 100 ns for each of the complexes with energy recording interval of 1.2 ps and recording Trajectories after every 4.8 ps. On completion of simulation, each trajectory was analysed in Simulation Interaction Diagram wizard which computes trajectories for Root Mean Square Deviation (RMSD) and Root Means Square Fluctuation (RMSF). Protein–ligand contact profiles for crucial interacting amino acid residues and timeline of these specific interactions are also computed with respect to 100 ns simulation.

Pharmacokinetics assessment using ADMET prediction

While a potential drug's optimum binding properties to the therapeutic target are vital, it's also critical to ensure that it can access the target site in adequate quantities to achieve the physiological result without causing any toxicity. Most often the probable drug compound falls out of the clinical trial due to poor pharmacokinetic properties. The experimental evaluation of these ADMET properties is also both time-consuming and expensive to scale in human and animal modes hence it becomes relevant to circumspect the physicochemical properties through computational approaches. The pharmacokinetic properties for GRL0617 and top five screened lead compounds were predicted in-silico using pkCSM-pharmacokinetics server of [29].

Both pharmacokinetic and toxicity properties were predicted using SMILES (Simplified Molecule Input Line Entry Specification) retrieved from PubChem for the lead compounds. It computed in vivo Absorption parameters like; Water solubility in buffer system (SK atomic types, mg/L), in vivo Caco2 cell permeability (Human colorectal carcinoma), Human intestinal absorption (HIA, %), in vivo P-glycoprotein substrate, inhibitor I and II; followed by assessment of in vivo skin permeability (logKp, cm/hour). Distribution property included tests like, Volume of Distribution of drug in the human system (VDss (human)), Plasma Protein fraction unbound feasibility, Blood–Brain Barrier (BBB) permeability and Central Nervous System (CNS) penetration. Metabolic parameters were determined using in-vivo Cytochrome P450 2C19, Cytochrome P450 2C9, Cytochrome P450 2D6, Cytochrome P450 3A4 inhibition, along with in-vivo Cytochrome P450 2D6 and 3A4 substrate inhibition. In this study we included Total Renal clearance and Renal OCT2 Substrate to identify Excretion efficacy

for the proposed natural products. To access the toxicity of our lead compounds, various important parameters including Acute algae toxicity, Ames test of mutagenicity, two years carcinogenicity bioassay in mouse, two years carcinogenicity bioassay in rat, in-vivo Ames test result in TA100 strain (Metabolic activation by rat liver homogenate) were computed.

Results

Redocking of GRL0617 with PLpro

Mostly the hydrophobic interactions drive the binding of GRL0617 with PLpro of SARS-CoV2 that imparts inhibition of the later. The aromatic rings 1-naphthyl moiety forms the most important interaction in the form of Pi-pi interaction with Tyr264 and Tyr268, is relatively less exposed to solvent and fits in to the cavity at the locus that accommodate the leucine at the P4 position. Moreover, the 1-naphthyl moiety also interacts with Pro247 and Pro248 side chains. The (R)-methyl group located at the stereocenter of GRL0617, positions itself into the protein interior in the small space between Tyr264 and Thr301, where it fits in the small polar space. Apart from 1-naphthyl moiety, GRL0617 also contains the lone aromatic ring containing $-NH_2$ at R3 position of GRL0617, positions itself at the opening of the cavity which has a polar nature due to the presence of multiple polar groups such as, the side chain oxygens of Gln269 and the hydroxyl group of Tyr268, where they mainly participate in the interaction by serving as hydrogen bond acceptor (Fig. 3).

Structure based virtual screening

A total of 35,032 natural compounds obtained from NPASS libraries were docked into the predicted active site of SARS-CoV2-PLpro. A step wise filtering protocol was used, in the first stage, the compounds were docked through HTVS where a total of 25 best hits were obtained. These 25 compounds were further docked with Glide SP and XP docking protocol and only one pose per ligand was retained. Finally, a total of 5 lead compounds were refined as shown in Table 1. Based on molecular docking scores, molecular interactions with the active site catalytic residue Tyr268 and structural aromaticity identical to GRL0617 were chosen (Fig. 3); the 5 lead compounds are namely, Caesalpiniaephenol A (−9.258 kcal/mol), Sappanone B (−9.531 kcal/mol), 3'-Deoxysappanone B (−8.476 kcal/mol), 1,2,3,4-tetrahydro-beta-carboline-3-carboxylic acid (−5.542 kcal/mol) and Clausine Z (−6.011 kcal/mol). While the docking score for the native ligand GRL0617 is −6.915 kcal/mol. Moreover,

Table 1 Binding energies and amino acid interaction profile of the top five hits obtained on performing molecular docking

Ligand	Docking score	Binding energy (Kcal/mol)	Compound ranking based on binding energy	Amino acid interactions
GRL-0617	−6.915	−46.308	Native inhibitor	LEU162, GLY163, ASP164, MET208, PRO247, PRO248, TYR264, GLYS266, ASN267, TYR268, GLN269, CYS270, GLY271, TYR273, THR301
Caesalpiniafenol A	−9.258	−45.319	1	GLU161, LEU162, GLY163, ASP164, VAL165, ARG166, MET208, PRO247, PRO248, TYR264, ASN267, TYR268, GLN269, CYS270, GLY271, TYR273
Sappone B	−9.531	−40.429	2	LEU162, ASP164, PRO248, TYR268, GLN269, TYR273
3'-Deoxysappone B	−8.476	−38.357	3	LYS157, GLU161, LEU162, GLY163, ASP164, VAL165, PRO247, PRO248, TYR264, ASN267, TYR268, GLN269, CYS270, GLY271, TYR273,
1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid	−5.542	−33.075	4	LYS157, GLU161, LEU162, GLY163, ASP164, GLU167, TYR264, TYR268, GLN269, TYR273
Clausine-Z	−6.011	−30.261	5	LYS157, GLU161, LEU162, GLY163, ASP164, GLU167, TYR264, TYR268, GLN269, TYR273

the structural features and chemical features of these compounds are represented in Table 2.

End-point binding free energy change calculation using MM-GBSA

Post docking MM-GBSA evaluation turns out to be significantly more dependent on values of ionic, hydrophilic and hydrophobic attraction of the protein–ligand intricate. The energy delivered (ΔG_{bind}) because of bond development, or rather communication of the ligand with protein is through restricting free energy and it decides the dependability of protein–ligand complex under study. The overall free energy

for the whole protein–ligand so obtains, if found negative, then the interaction is considered to be occurring spontaneously and thus such a scenario would make the reaction to occur more favourable. The binding free energy change profiles in forms of MM-GBSA values of all the 5 lead compounds in correlation with reference GRL0617 is addressed in Table 3. The ligand–protein association of GRL0617 with SARS-CoV2-PLpro as SARS-CoV2-PLpro-GRL0617 complex is predicted to happen profoundly immediately as the ΔG_{bind} is -66.235 kcal/mol.

When it comes to screened natural products from NPASS library, out of the top 5 lead compounds, in terms of binding free energy change Caesalpiniafenol A with ΔG_{bind}

Table 2 Structure and chemical properties of shortlisted natural aromatic compounds

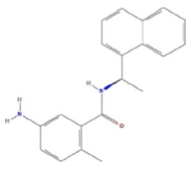
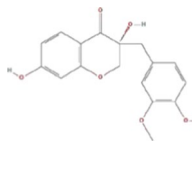
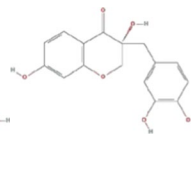
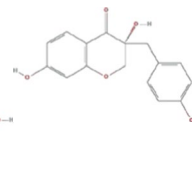
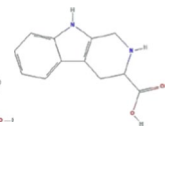
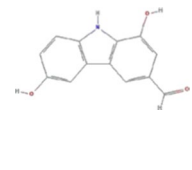
Descriptor	Compound name and values					
	GRL-0617	Caesalpiniafenol A	Sappone B	3'-Deoxysappone B	1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid	Clausine-Z
2D Structure						
Molecular weight	304.393	316.309	302.282	286.283	216.24	227.219
LogP	4.22142	1.6553	1.3523	1.6467	1.2668	2.5448
#Rotatable bonds	3	3	2	2	1	1
#Acceptors	2	6	6	5	2	3
#Donors	2	3	4	3	3	3
Surface area	135.681	132.079	125.394	120.600	92.282	96.119

Table 3 MM-GBSA binding free energy change profiles of ligands with PLpro of SARS-CoV2 for docked compounds

Ligand	ΔG_{Bind} (Kcal/mol)	$\Delta G_{\text{Coulomb}}$ (Kcal/mol)	ΔG_{Hbond} (Kcal/mol)	ΔG_{Lipo} (Kcal/mol)	$\Delta G_{\text{Packing}}$ (Kcal/mol)	ΔG_{vdW} (Kcal/mol)
GRL-0617	−66.235	−19.279	−2.37	−26.1	−3.064	−46.057
Caesalpinaphenol A	−60.297	−26.503	−1.614	−20.966	−2.187	−33.475
Sappone B	−56.638	−26.669	−1.654	−20.012	−2.294	−31.72
3'-Deoxysappone B	−56.993	−27.047	−1.608	−19.968	−2.245	−30.657
1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid	−31.214	−11.18	−1.123	−12.302	−2.219	−32.201
Clausine-Z	−30.634	−21.314	−1.594	−10.23	−4.294	−25.255

ΔG_{Bind} –Binding energy, $\Delta G_{\text{Coulomb}}$ –Coulomb energy, ΔG_{Hbond} –Hydrogen-bonding correction, ΔG_{Lipo} –Lipophilic energy, $\Delta G_{\text{Packing}}$ –Pi-Pi packing correction, ΔG_{vdW} –Van der Waals energy

−60.297 kcal/mol is considered the best ligand. The next best ligand to interact with SARS-CoV2-PLpro is Sappanone B with −56.638 kcal/mol, followed by 3'-Deoxysappone B with −56.993 kcal/mol; 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid having −31.214 kcal/mol and the lowest Clausine Z with −30.634 kcal/mol binding free energy change. In addition to the total energy, the contributions of the total energy from different components such as Hydrogen-bonding correction, Coulomb energy, Pi-pi packing correction, Van der Waals energy and Lipophilic energy are provided in Table 3. All these parameters conclusively help to determine Caesalpinaphenol A as a top hit compound and was further validated as such by performing simulations.

Molecular dynamic simulation

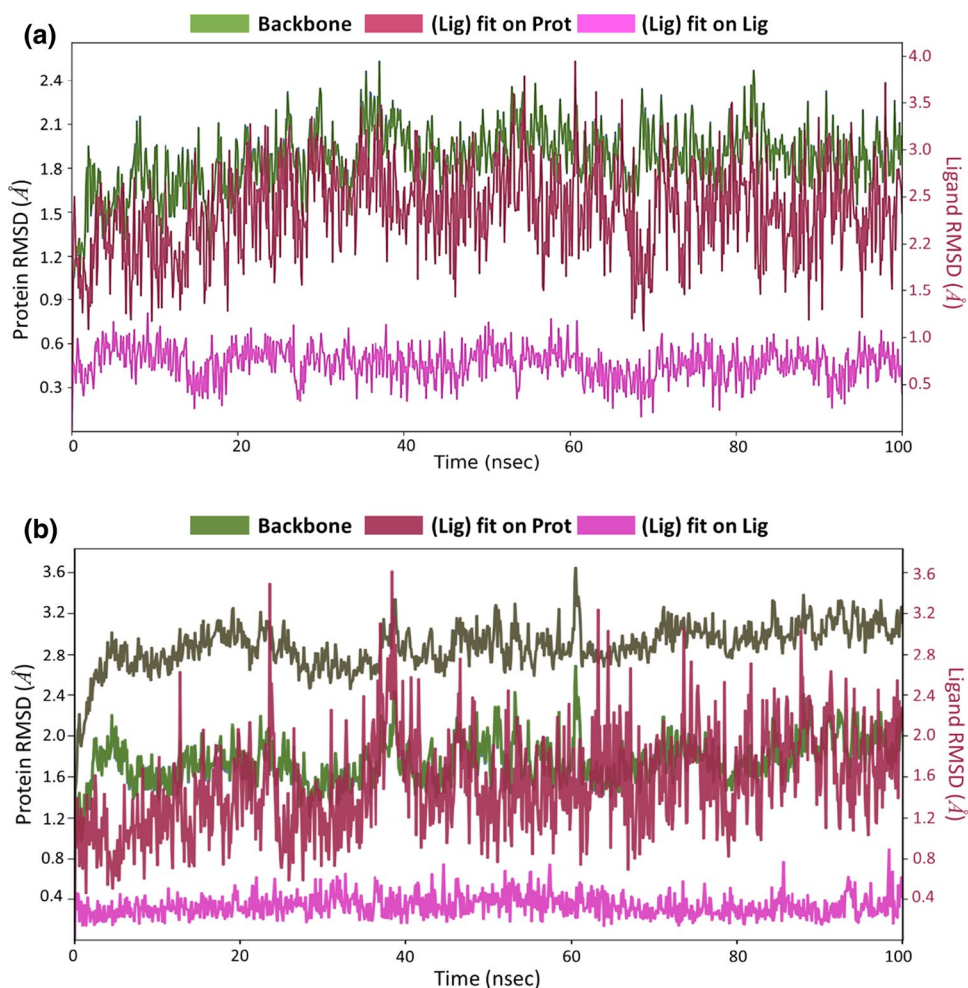
After carrying out virtual screening and post-docking analysis of the natural compounds from NPASS database, Caesalpinaphenol A was chosen as an ideal candidate compound to interact with SARS-CoV-2-PLpro; therefore, to investigate the validity of the docking data, the MD simulations were performed. The docked complexes of SARS-CoV2-PLpro-Caesalpinaphenol A and SARS-CoV2-PLpro-GRL0617 were subjected to a 100 ns MD simulation, where the docked complex with GRL0617 was taken as a control.

Figure 4 is the graph portraying RMSD deviations in the segments of the protein during simulation, part of which is also protein equilibration (left Y-pivot). During assessing the trajectories of MD reproductions, the RMSD examination is needed to be performed for evaluating the typical change in movement of the of structural atoms with respect to the reference starting frame. The posture of docked ligand with protein in the complex is considered as the static initial reference orientation and afterward the changes taking place in the posture of entire complex for this unique arrangement during MD simulation is checked by overlapping all the protein movement frames obtained. Appraisal of the RMSD

of a protein can give knowledge regarding its movements during the MD reproduction is simplified and represented in 2D chart. Besides, RMSD assessment can also determine whether the simulation has equilibrated—its progressions towards the completion of the simulation is driven to achieve stable conformity. Notwithstanding, this range of RMSD value enlarges with the increments in the size of protein. For the complex of SARS-CoV-2-PLpro-GRL0617 (Fig. 4a) the protein's RMSD value does not surpass 2.5 Å; for the complex of SARS-CoV-2-PLpro-Caesalpinaphenol A (Fig. 4b), the RMSD value peaks till 3.6 Å, which is worthy since for globular proteins changes of 1–4 Å are totally agreeable. Ligand RMSD (right Y-axis, plots of Fig. 4) recommends the solidness of ligand orientation concerning the docked position of the ligand in protein cavity. 'Lig fit Prot' proposes the RMSD of a ligand for with reference to protein. For this, the RMSD values marginally greater than the protein's RMSD are viewed as agreeable however in the event where the qualities noticed are fundamentally bigger than the RMSD of the protein, at that point almost certainly, the ligand acquires an unexpected stable orientation in comparison to its orientation in reference pose. The 'Lig fit Prot' RMSD values for both the docked complexes stays in the range of 1.0–3.5 Å and comparatively 1.0–3.2 Å all through simulation for SARS-CoV-2-PLpro-GRL0617 and SARS-CoV2-PLpro-Caesalpinaphenol A, separately. 'Lig Fit Lig' proposes how much the ligands reorients from its actual position determined during docking, Here for Caesalpinaphenol A (0.4–0.8 Å) this worth is little lower than appeared by GRL0617 (0.5–1.0 Å), recommending that the Caesalpinaphenol A remains very steady, with lesser vibrations for the given its docked pose with respect to its control, GRL0617.

The value of RMSF gives data about the dynamic behaviour of protein in aqueous simulated system and is helpful for depicting localized minute changes along the length of protein chains. In the diagram, the pinnacles show districts of the protein that shift the most all through the simulation

Fig. 4 A 100 ns simulation profile of Protein–ligand interaction, root-mean-square deviation (RMSD) for **a** SARS-CoV2-PLpro-GRL0617 **b** SARS-CoV2-PLpro-Caesalpin-iaphenol A

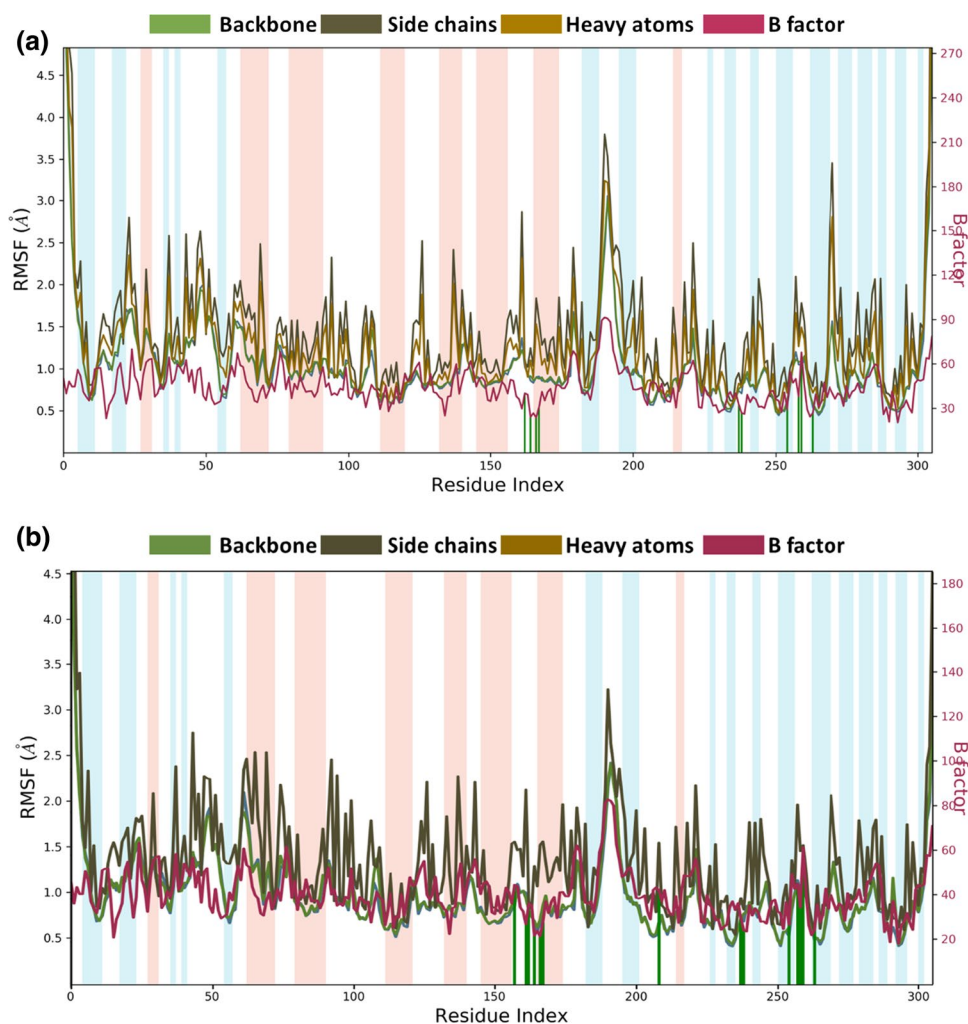


course. Commonly, the protein tails (N- and C-terminal) change the greatest than other inside locales of the protein. Auxiliary locales of proteins like alpha helices and beta strands are by and large more unbendable and inflexible than the unstructured areas and subsequently sway not actually like circle framing segments of protein. Alpha-helical and beta-strand regions are included in red and blue colours, independently. Both complexes, for this situation, SARS-CoV-2-PLpro-GRL0617 (Fig. 5a) and SARS-CoV-2-PLpro-Caesalpin-iaphenol A (Fig. 5b) portray protein cooperation with the ligands and patterns for the values of RMSF and B-factor compare similarly. Proposing that both the protein–ligand complexes have less change iterating a lesser flexing of protein backbone.

The interaction between protein and ligand were observed throughout course of MD simulation. This interaction types are classified into four kinds: Hydrogen Bonds, Hydrophobic connections, Ionic contacts and Water Bridges, which can be researched through the graphical representation of 'Simulation Interactions Diagram'. The stacked bar traces are normalized cumulative interaction profile: for example, an assessment of 0.8 suggests that 80% of the time during

simulation the corresponding interaction remains durable. Characteristics over 1.0 are possible as some amino acids may more than one types of contacts of the equivalent subtype with the ligand. In figs. 6 and 7 it is seen the interaction profile observed in the results of docking are authenticated during MD simulation for both, GRL0617 and Caesalpin-iaphenol A where the normal associations incorporate amino acids, Leu162, Asp164, Arg166, Glu167, Pro247, Pro248, Tyr264, Tyr268, Gln269 and Tyr273. Mulling over here, the crucial amino acid Tyr268; the reported inhibitor part interaction value is 0.9 (~90%) while for the lead hit, Caesalpin-iaphenol A, this value is nearing 1.0 (~100%). A depiction of the affiliations and contacts (Hydrogen bonds, Hydrophobic, Ionic, Water ranges) is shown in the Fig. 6a for SARS-CoV-2-PLpro-GRL0617 complex and Fig. 7a for SARS-CoV-2-PLpro-Caesalpin-iaphenol A complex. Henceforth, this complex can robust stable interaction all through the length of 100 ns of simulation. A timeline of the affiliations and contacts (Hydrogen bonds, Hydrophobic, Ionic, Water ranges) is showed up in the Fig. 8a for SARS-CoV2-PLpro-GRL0617 complex and Fig. 8b for SARS-CoV2-PLpro-Caesalpin-iaphenol A complex. These figures depict

Fig. 5 A 100 ns simulation profile of Protein–ligand interaction, root-mean-square fluctuation (RMSF) for **a** SARS-CoV2-PLpro-GRL0617 **b** SARS-CoV2-PLpro- Caesal-piniaphenol A



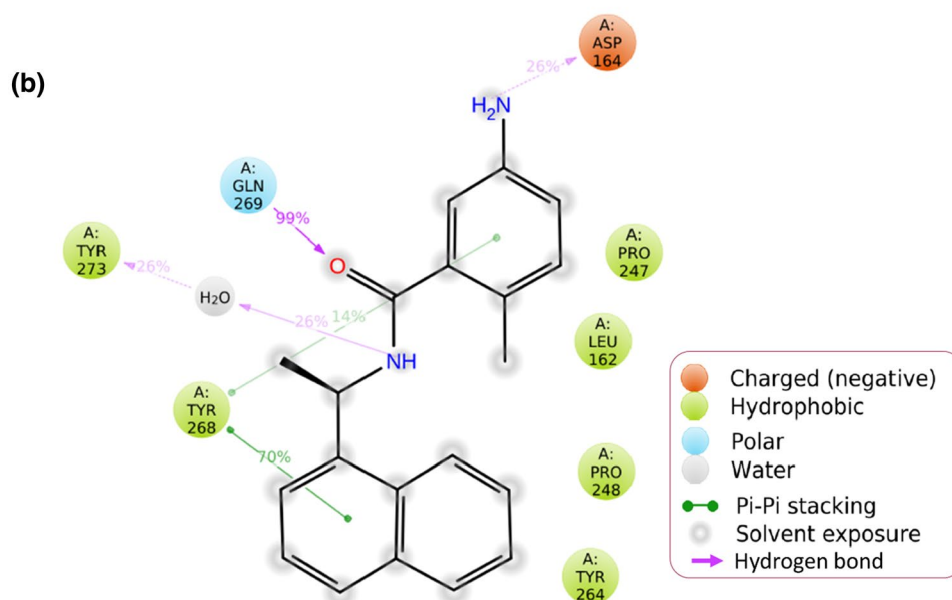
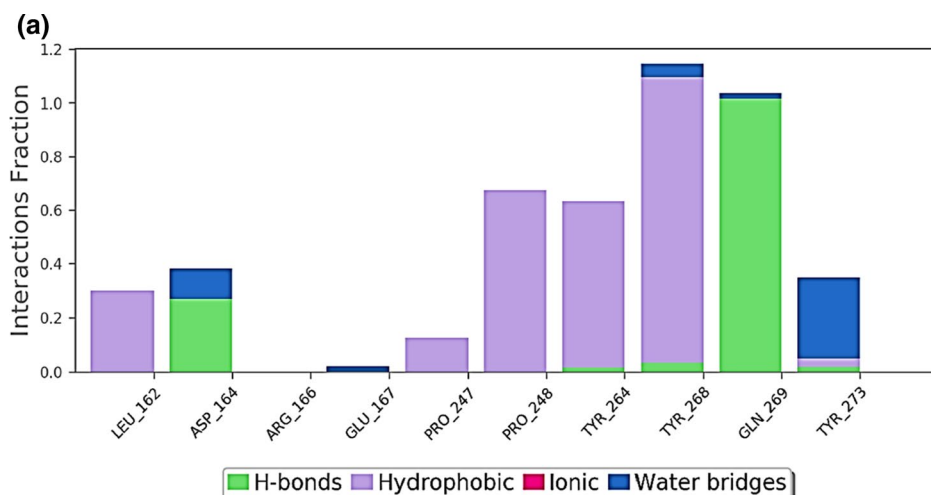
which amino acids interact with the ligand toward each path layout. Few amino acids make more than one express contact with the ligand, which is appeared by a hazier shade of orange, as demonstrated by the scale aside of the plot. The plots verify the discoveries of docking recommending the interactions proposed by the docked pose obtained by molecular docking, are being made by same amino acids during MD simulations.

Pharmacokinetics assessment using ADMET prediction

The top hits from NPASS database and the control inhibitor were assessed for drug-similarity properties by measuring their different physiochemical boundaries those are fundamental for drug disclosure. Our top docked hits namely, Caesal-piniaphenol A, Sappanone B, 3'- Deoxysappanone B, 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid and Clausine-Z possess adequate scope of physiochemical properties drug like molecules by abiding the Lipinski's rule of 5 (Table 2). Furthermore, these screened compounds were

assessed for their pharmacokinetic boundaries (Table 4). The absorption results uncover that 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid is profoundly water soluble as it has absorption value of $-2.367 \log \text{ mol/L}$ while GRL-0617 has this value of $-4.678 \log \text{ mol/L}$ and so is less dissolvable in water. Further, 3'- Deoxysappanone B, GRL-0617 and Clausine-Z showed an extraordinary % human intestinal absorption, proposing that these compounds could be absorbed by intestine upon ingestion. The last assertion has been additionally affirmed by the consequences of Caco-2 cell monolayer model which function on an in-vitro model of intestinal retention. All the top five top hits are substrate of P-glycoprotein and none of them is an inhibitor of P-glycoprotein-I and P-glycoprotein-II. In addition, all the compounds other than 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid have lesser degree of steady-state volume of distribution. While 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid has the highest unbound fraction in the human blood which means it can efficiently transverse cellular membranes or diffuse. All the bioactive compounds except GRL-0617 have relatively low blood–brain barrier

Fig. 6 a Interaction profile of crucial interacting amino acids of the SARS-CoV2-PLpro in contact with GRL0617 **b** GRL0617 Ligand interaction diagram displaying total time (in %) a particular amino acid of the protein over the course of simulation



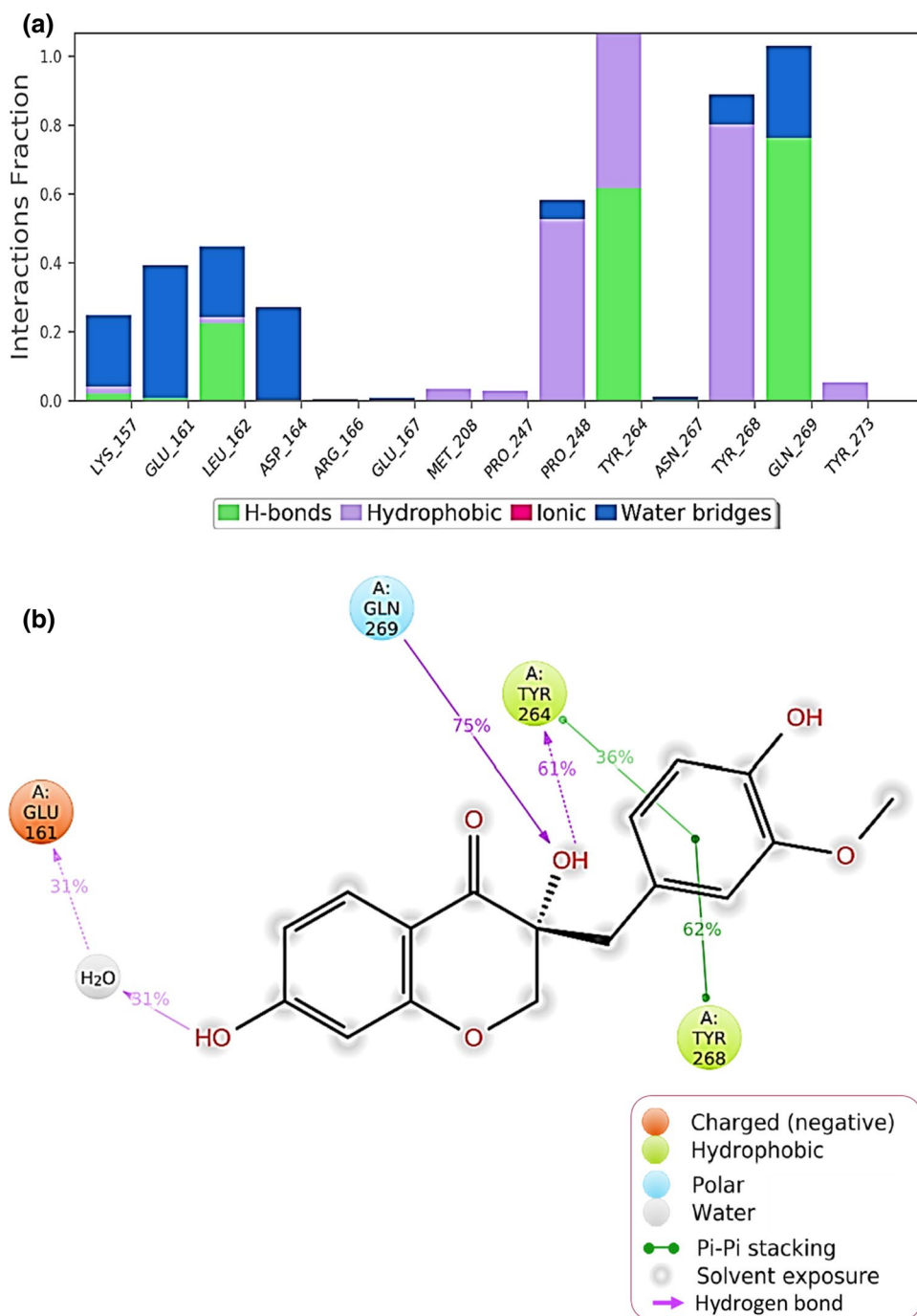
and CNS permeability values. Also, the hits were validated for human cytochrome P450 promiscuity. The predicted results revealed that only 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid was the substrate of CYP2D6 and none of them being its inhibitor. Similarly, GRL-0617 and 3'-Deoxysappanone B was reported as substrate of CYP3A4 and only GRL-0617 being its inhibitor. In addition, except Sappanone B and 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic others were inhibitor of CYP1A2. Further none of the bioactive compound is reported as substrate of renal organic cation transporter which is a renal uptake transporter that plays key role in disposition and renal clearance of drugs and endogenous compound. It is quite important to forecast the toxicity profile of all tested bioactive compounds in order to save time and resources during clinical screening process of drug discovery. 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic has no mutagenic potential against

bacteria (AMES test) but others could be toxic to bacteria. The compounds have reported no side effects on the hepatic as well as dermal cells. None of the bioactive compounds are inhibitors of human ether-a-go-go-related gene (hERG) hERG I and hERG II except GRL-0617 which is hERG II inhibitor. All the top lead compounds display relatively low recommended tolerated dose values as shown in Table 4.

Discussion

Microbes are omni present and they exist in various forms, to name few the dominating ones are fungi, bacteria and viruses [30, 31]. In the twenty-first century we are generation, which is envisaging the good, the bad and the ugly of microbes. Humans exploit microbes to produce alcoholic beverages, antibiotics, flavoring agents etc. forms the good

Fig. 7 a Interaction profile of crucial interacting amino acids of the SARS-CoV2-PLpro in contact with Caesalpiniaphenol A **b** Caesalpiniaphenol A Ligand interaction diagram displaying total time (in %) a particular amino acid of the protein over the course of simulation

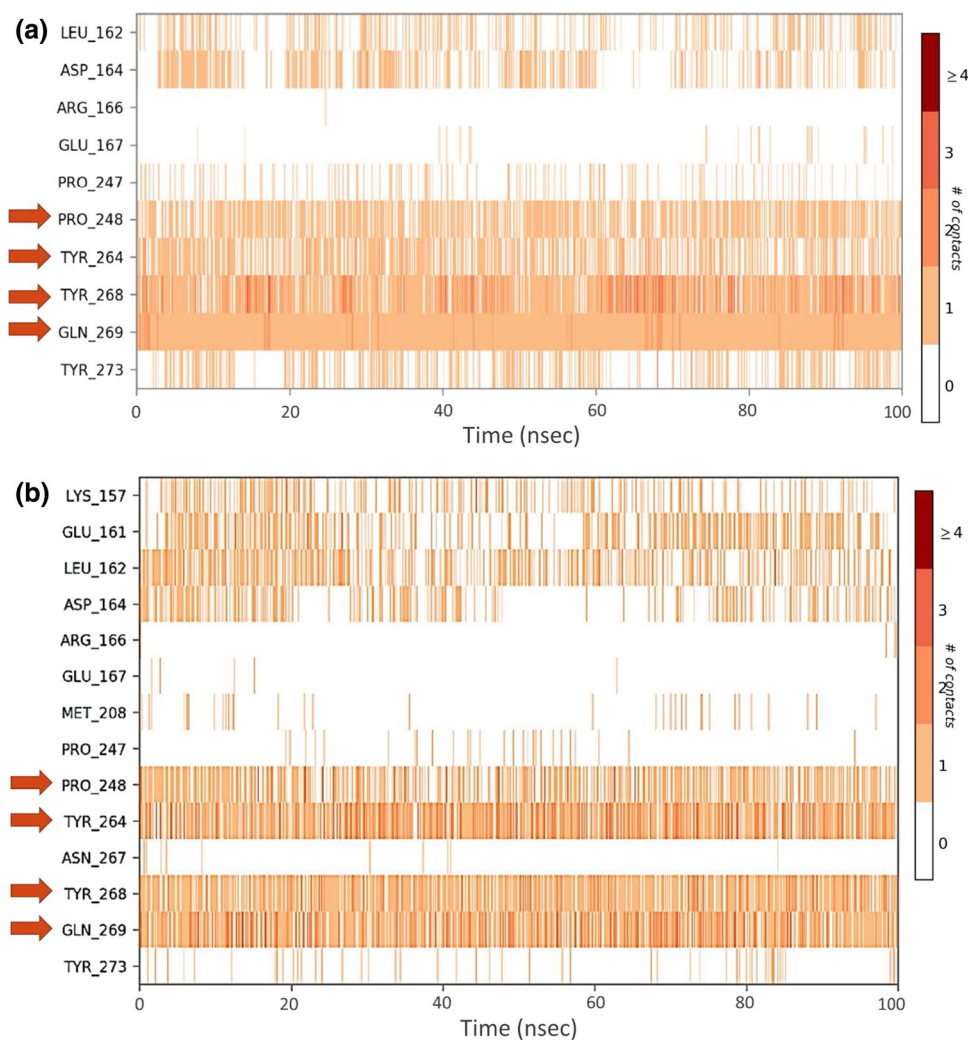


part of microbes [32, 33]. The bad is the crop spoilage by fungi and other saprophytes while the glimpse of the ugly side is the current COVID-19 pandemic that we are experiencing [6]. Before 600 years ago, the the human CoV NL63 and bat CoV ARCoV2 are sought to have evolved from a common coronavirus (CoV) ancestor and there exist at least 100 more CoVs in bats that are lethal if they gain ability to infect humans [34]. Moreover, In 2007 it was predicted by Cheng and colleagues that the ability of CoVs to undergo rapid stable genetic mutation and recombination leading the

development of new genotypes, is virtually a ticking time bomb for lethal CoVs for human to develop [35].

To address the alarming issue, there are several strategies being implemented to combat the pandemic, the most obvious one it to develop an efficient vaccine which is the most traditional approach. Where other approaches use unconventional routes where the researchers target important viral proteins and inhibit its function. One such protein is PLpro of SARS-CoV2. The importance of targeting this protein is vividly described by various publications in

Fig. 8 Timeline representation of the interactions of ligand with amino acids for the complex **a** SARS-CoV2-PLpro-GRL0617 **b** SARS-CoV2-PLpro-Caesalpiniafenol A



renowned journals such as Nature [14], Nature communications [15] and in Nature Signal Transduction and Targeted Therapy [10]. Owing to the importance of targeting PLpro few researchers have tried to work out strategies to inhibit its function. In one examination by Klemm and associates, directed 3727 extraordinary endorsed medications to fulfill the reason for drug repurposing concluded Remdesivir and Hydroxychloroquine to collaborate with PLpro to inhibit its function and showed promising outcomes under in-vitro tests [9]. In another investigation, Elekofehinti and associates focused around 50,000 natural compounds from IBS information repository (<https://www.ibscreen.com/natural.shtml>) to discover inhibitors for PLpro, their potent natural molecules were STOCK1N-69160 and STOCK1N-69160 which could at greatest make just a single Pi-Pi stacked bonds with Tyr268. As mentioned earlier, the interaction with Tyr268 is essential for inducing inhibitory effect on PLpro. Additionally, their exploration didn't utilized any certain control like GRL0617 to analyze their outcomes. Under current study we targeted natural compounds from NPASS

database where we found Caesalpiniafenol A, Sappanone B, 3'-Deoxysappanone B, 1,2,3,4-tetrahydro-beta-carboline-3-carboxylic acid and Clausine Z to all effectively interact with Try268 of PLpro, moreover MM-GBSA assessment of docked compounds showed that of all top five hits, Caesalpiniafenol A to be best molecule and further its interaction was effectively validated with MD simulations. Caesalpiniafenol A is largely found from the plant *Caesalpinia digyna*, which is also named as *Moullava digyna*. The root of this plant is known to produce this phytochemical and there are several reports suggesting Caesalpiniafenol A possessing strong antioxidant activity [36–38]. Endeavors are likewise utilizing QSAR based information mining of assorted bioactive compounds using QSAR based virtual screening to recognize the inhibitor for viral PLpro [39]. In a detailed review published by France and partners, distinguished various inhibitors for PLpro and they are as per their literature audit the molecules identified based on drug repurposing methodologies which are as follows Carbamate, Cefamandole, Chloramphenicol, Chlorphenesin, Darunavir,

Table 4 ADMET properties of native inhibitor and screened aromatic compounds

Property	Descriptor	Predicted description for top screened compounds						Unit
		GRL-0617	Caesal-piniaphenol A	Sappanone B	3'-Deoxy-sappanone B	1,2,3,4-tetrahydro-beta-carboline-3-carboxylic acid	Clausine Z	
Absorption	Water solubility	-4.678	-2.937	-3.29	-3.341	-2.367	-3.477	Numeric (log mol/L)
Absorption	Caco2 permeability	1.302	-0.031	-0.2	1.294	0.619	0.95	Numeric (log Papp in 10–6 cm/s)
Absorption	Intestinal absorption (human)	92.815	78.428	71.471	94.099	79.974	91.491	Numeric (% Absorbed)
Absorption	Skin Permeability	-2.785	-2.739	-2.735	-2.74	-2.735	-2.761	Numeric (log Kp)
Absorption	P-glycoprotein substrate	Yes	Yes	Yes	Yes	Yes	Yes	Categorical (Yes/No)
Absorption	P-glycoprotein I inhibitor	No	No	No	No	No	No	Categorical (Yes/No)
Absorption	P-glycoprotein II inhibitor	Yes	No	No	No	No	No	Categorical (Yes/No)
Distribution	VDss (human)	0.086	0.089	0.851	0.305	-1.325	-0.085	Numeric (log L/kg)
Distribution	Fraction unbound (human)	0	0.1	0.164	0.096	0.432	0.18	Numeric (Fu)
Distribution	BBB permeability	0055	-0.706	-0.922	-0.733	-0.292	-0.658	Numeric (log BB)
Distribution	CNS permeability	-1.604	-3.247	-3.273	-2.4	-2.401	-2.099	Numeric (log PS)
Metabolism	CYP2D6 substrate	No	No	No	No	Yes	No	Categorical (Yes/No)
Metabolism	CYP3A4 substrate	Yes	No	No	Yes	No	No	Categorical (Yes/No)
Metabolism	CYP1A2 inhibitor	Yes	Yes	No	Yes	No	Yes	Categorical (Yes/No)
Metabolism	CYP2C19 inhibitor	Yes	No	No	Yes	No	No	Categorical (Yes/No)
Metabolism	CYP2C9 inhibitor	Yes	No	No	Yes	No	No	Categorical (Yes/No)
Metabolism	CYP2D6 inhibitor	No	No	No	No	No	No	Categorical (Yes/No)
Metabolism	CYP3A4 inhibitor	Yes	No	No	No	No	No	Categorical (Yes/No)
Excretion	Total Clearance	0.211	0.091	0.051	0.006	0.786	0.491	Numeric (log ml/min/kg)
Excretion	Renal OCT2 substrate	No	No	No	No	No	No	Categorical (Yes/No)
Toxicity	AMES toxicity	Yes	Yes	Yes	Yes	No	Yes	Categorical (Yes/No)
Toxicity	Max. tolerated dose (human)	-0.043	0.093	0.216	-0.117	0.837	0.424	Numeric (log mg/kg/day)
Toxicity	hERG I inhibitor	No	No	No	No	No	No	Categorical (Yes/No)
Toxicity	hERG II inhibitor	Yes	No	No	No	No	No	Categorical (Yes/No)
Toxicity	Oral Rat Acute Toxicity (LD50)	0.2.472	2.154	2.327	2.208	2.31	1.823	Numeric (mol/kg)
Toxicity	Oral Rat Chronic Toxicity (LOAEL)	0.462	2.303	2.425	1.82	1.172	1.272	Numeric (log mg/kg_bw/day)
Toxicity	Hepatotoxicity	No	No	No	No	No	No	Categorical (Yes/No)
Toxicity	Skin Sensitisation	No	No	No	No	No	No	Categorical (Yes/No)
Toxicity	<i>T. Pyriformis</i> toxicity	0.529	0.374	0.335	0.407	0.285	0.737	Numeric (log ug/L)
Toxicity	Minnow toxicity	1.936	1.424	2.77	1.619	1.216	0.625	Numeric (log mM)

Levodropropizine, Luteolin, Lopinavir, Methicillin, Omeprazole, Ribavirin, Ritonavir, Thymidine, Tigecycline, Tolazamide and Vlganciclovir. The review also highlights the molecule GRL0617 as an only inhibitor recognized till date for both SARS-CoV's and SARS-CoV-2's PLpro. [40]. This

supports our notion to select GRL0617 as a positive control and reference inhibitor in this in-silico study. With confined augmentation to work with SARS-CoV2, as it requires Biosafety Level 4 (BSL4) setup, progressively more investigation through in-silico approach with docking and MD

reenactments is at the heart of computational assessment, paying little mind to such assessment ailing in-vitro examinations, massive volumes of data showing the coordinated efforts of a bioactive compound or phytochemical with an ability to inhibit function of viral protein is been brought to the experts territory, which can come exceptionally helpful for those having BSL4 setup for executing in-vitro assessments for favoring computational gauges finally saving a huge load of time.

The more extensive portrayal of bioactive natural molecules that we have identified from NPASS database to be compelling for restraining PLpro will be centered around their ADMET properties. The aftereffects of the bioavailability, pharmacodynamics, pharmacokinetics and toxicological profiles of Caesalpiniaaphenol A, Sappanone B, 3'-Deoxysappanone B, 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid, Clausine-Z and GRL-0617 are represented in Table 2 and Table 4. These screened hits were fundamentally investigated to avoid any disaster during clinical improvement of these medications because of the ADME/Tox inadequacies. The Lipinski rule of five is a dependable guideline to anticipate the medication drug likeness property or to foresee the pharmacological or natural movement of a compound that makes the compound an ideal contender for orally dynamic medication [41]. As none of the top hits violated the Lipinski's rule of 5, these phytochemicals could be kept up in the framework at appropriate concentrations. Further, to decide the bioavailability of medication competitors, the human intestinal assimilation and Caco-2 penetrability are the two most extreme significant parameters. The tested compounds (Caesalpiniaaphenol A, Sappanone B, 3'-Deoxysappanone B, 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid and Clausine-Z) have relatively low Caco-2 permeability and so predicted to be easily absorbed through the human intestine [42, 43]. Moreover, all the tested ligands were predicted to be substrates of P-glycoprotein which is an efflux transporter and a member of ABC transporter family present majorly in epithelial cells. On the contrary, none of the ligands except native ligand GRL-0617 was found to be the P-glycoprotein inhibitor. This implies that that ligands i.e., Caesalpiniaaphenol A, Sappanone B, 3'-Deoxysappanone B, 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid, Clausine-Z doesn't interfere with the normal physiological activities of P-glycoprotein that includes the active uptake and the distribution of drugs [44]. The volume of distribution calculated using steady-state volume of distribution (VD_{ss}) revealed the theoretical dose necessary for even distribution in the plasma. 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid showed lowest theoretical dose among all and Caesalpiniaaphenol A showed the highest dose require to distribute the drug in the tissue and plasma. The increasing order of the diffusion across plasma membrane is 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic

Acid < Clausine-Z < Sappanone B < Caesalpiniaaphenol A < 3'-Deoxysappanone B as the fraction that is in the unbound state. The forecast investigation on the distribution across central nervous system of the tested compound revealed that lipophilicity of the phytochemicals is directly proportional to the degree of permeability across the nervous system and blood-brain barrier. Cytochrome P450 is known to play vital function in drug activation, drug metabolism and in the drug, toxicity effects as well. Only 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid is substrate of CYP2D6 and 3'-Deoxysappanone B is substrate of CYP3A4. None of the tested phytochemical were reported as inhibitor of both CYP2D6 and CYP3A4. Additionally, none of the tested compounds were predicted as substrate of renal organic cation transporter, which suggests that there might be possibilities of drug being cleared through other routes i.e., sweat, bile secretion and many more. Total drug clearance is other important filter to validate dosing rates to achieve steady-state concentrations. Predicted results revealed that Caesalpiniaaphenol A, Sappanone B, 3'-Deoxysappanone B have relatively low clearance as compared to 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid and Clausine-Z. The toxicity profile of the tested ligands disclosed that all the compounds are safe on skin and possess neither dermal toxicity nor hepatotoxicity. Likewise, none of the bioactive compound is inhibitor of hERG I and hERG II. hERG transporter inhibition could possibly delay ventricular repolarization resulting in severe disturbance in the normal cardiac rhythm and interfere hepatic functioning [45]. To determine the safety level of ligands when administered, acute and chronic toxicity were evaluated. One of the noteworthy concerns in numerous treatment options is the side effects of drugs when given in low to moderate concentration over a long period of time. Filters such as oral rat acute toxicity (LD₅₀) and oral rat chronic toxicity (LOAEL) revealed no adverse effects of the tested bioactive compounds. The results of the bioavailability, pharmacodynamics, pharmacokinetics and toxicological profiles of Caesalpiniaaphenol A, Sappanone B, 3'-Deoxysappanone B, 1,2,3,4-Tetrahydro-Beta-Carboline-3-Carboxylic Acid, Clausine-Z and GRL-0617 are presented in Table 2 and Table 4. These screened hits were significantly scrutinized to dodge any fiasco during clinical development of these drugs due to the ADME/Tox deficiencies. The Lipinski rule of five is a rule of thumb to predict the drug likeness property or to predict the pharmacological or biological activity of a compound that makes the compound a perfect candidate for orally active drug [41]. As none of the tested compounds violated the Lipinski's rule of 5, these phytochemicals could be maintained in the system at appropriate concentrations.

Glancing through bioactive compounds that can control SARS-CoV2 is of the most raised requirement for researchers. To market such molecules as drugs for SARS-CoV2,

natural biomolecules from the plant and microbial origins are being searched, due to their low toxicity, easy to purify from natural life forms and being recognized by people. Lately, our research team recently predicted Pyranonigrin A and Flaviolin to have the prospects to interfere with Mpro of SARS-CoV2 to inhibit its function [46]. Further we have recently proposed we have postulated Fonsecin, a fungal metabolite to be the structural–functional analogue of GRL0617 and current version of manuscript is an extension of our work in search for structure–function analogue of GRL0167 [16]. A sizable portion of the approach for distinguishing the inhibitors of viral proteins uses atomic docking and MD reenactments as the primary strategies for computational assessment, and by using these methods a few lead compounds are perceived to intrude with the organic chemistry driven by viral proteins of SARS-CoV2. There are several evidences in the sphere of biosciences where the computational methodologies that make use of docking and MD generations have been of an amazing use [32, 47, 48]. It is without a doubt a troublesome time yet as disarray and vulnerability prevail, there surely is chaos but also hope. Researchers and specialists all throughout the planet are utilizing the web to distribute data quicker than any time in recent history. This has yielded some potential solutions and we with this article would help fill the knowledge gap needed for developing strategies to combat the pandemic.

Acknowledgements Authors are thankful to Department of Microbiology and Biotechnology, School of Sciences, Gujarat University, DST-FIST Sponsored Department, for providing necessary facilities to perform experiments. Authors are also thankful to UGC (Start-up Research Grant) for providing fund under the F.30-521/2020(BSR), Under Secretary FD-III Section, University Grant Commission, New Delhi 110002. We acknowledge GSBTM, DST, Government of Gujarat for providing Bioinformatics Node facility and Finishing School support. We acknowledge GUCOST, DST, Government of Gujarat for Super-computing facility provision.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All authors have seen and approved the manuscript.

Human or animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Chen Y, Guo Y, Pan Y, Zhao ZJ (2020) Structure analysis of the receptor binding of 2019-nCoV. *Biochem Biophys Res Commun*. <https://doi.org/10.1016/j.bbrc.2020.02.071>
- Song Z, Xu Y, Bao L, Zhang L, Yu P, Qu Y, Zhu H, Zhao W, Han Y, Qin C (2019) From SARS to MERS, thrusting coronaviruses into the spotlight. *Viruses*. <https://doi.org/10.3390/v11010059>
- Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YZYZL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZYZL (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579:265–269. <https://doi.org/10.1038/s41586-020-2008-3>
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Rao P, Shukla A, Parmar P, Rawal RM, Patel B, Saraf M, Goswami D (2020) Reckoning a fungal metabolite, Pyranonigrin A as a potential Main protease (Mpro) inhibitor of novel SARS-CoV-2 virus identified using docking and molecular dynamics simulation. *Biophys Chem*. <https://doi.org/10.1016/j.bpc.2020.106425>
- Shukla A, Parmar P, Kapoor G, Goswami D, Jha CK, Patel B, Saraf M (2021) Curse of La Corona: unravelling the scientific and psychological conundrums of the 21st century pandemic. *Mol Divers* 1:3. <https://doi.org/10.1007/s11030-020-10167-2>
- Kong R, Yang G, Xue R, Liu M, Wang F, Hu J, Guo X, Chang S (2020) COVID-19 docking server: an interactive server for docking small molecules, peptides and antibodies against potential targets of COVID-19. *J ArXiv Prepr*. <https://arxiv.org/abs/2003.00163>
- Joshi RS, Jagdale SS, Bansode SB, Shankar SS, Tellis MB, Pandya K, Chugh A, Giri AP, Kulkarni MJ, Joshi S (2020) Discovery of potential multi-target-directed ligands by targeting host-specific SARS-CoV-2 structurally conserved main protease. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1760137> (Taylor Fr)
- Klemm T, Ebert G, Calleja DJ, Allison CC, Richardson LW, Bernardini JP, Lu BG, Kuchel NW, Grohmann C, Shibata Y, Gan ZY, Cooney JP, Doerflinger M, Au AE, Blackmore TR, Hedenvan-Noort GJ, Geurink PP, Ovaas H, Newman J, Riboldi-Tunnicliffe A, Czabotar PE, Mitchell JP, Feltham R, Lechtenberg BC, Lowes KN, Dewson G, Pellegrini M, Lessene G, Komander D (2020) Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *EMBO J* 39:1–17
- McClain CB, Vabret N (2020) SARS-CoV-2: the many pros of targeting PLpro. *Signal Transduct Target Ther* 5:1–2. <https://doi.org/10.1038/s41392-020-00335-z>
- Boozari M, Hosseinzadeh H (2020) Natural products for COVID-19 prevention and treatment regarding to previous coronavirus infections and novel studies. *Phyther Res*. <https://doi.org/10.1002/ptr.6873> (ptr.6873)
- Elekofehinti OO, Iwaloye O, Josiah SS, Lawal AO, Akinjiyan MO, Ariyo EO (2020) Molecular docking studies, molecular dynamics and ADME/tox reveal therapeutic potentials of STOCKIN-69160 against papain-like protease of SARS-CoV-2. *Mol Divers*. <https://doi.org/10.1007/s11030-020-10151-w>
- Murugan NA, Kumar S, Jeyakanthan J, Srivastava V (2020) Searching for target-specific and multi-targeting organics for Covid-19 in the Drugbank database with a double scoring approach. *Sci Rep* 10:1–16. <https://doi.org/10.1038/s41598-020-75762-7>
- Shin D, Mukherjee R, Grewe D, Bojkova D, Baek K, Bhattacharya A, Schulz L, Widera M, Mehdi-pour AR, Tascher G, Geurink PP, Wilhelm A, van der Heden GJ, van Noort H, Ovaas S, Müller KP, Knobloch K, Rajalingam BA, Schulman J, Cinatl G, Hummer S, Ciesek ID (2020) Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* 587:657–662. <https://doi.org/10.1038/s41586-020-2601-5>
- Fu Z, Huang B, Tang J, Liu S, Liu M, Ye Y, Liu Z, Xiong Y, Zhu W, Cao D, Li J, Niu X, Zhou H, Zhao YJ, Zhang G, Huang

- H (2021) The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. *Nat Commun* 12:488. <https://doi.org/10.1038/s41467-020-20718-8>
16. Rao P, Patel R, Shukla A, Parmar P, Rawal RM, Saraf M, Goswami D (2021) Identifying structural–functional analogue of GRL0617, the only well-established inhibitor for papain-like protease (PLpro) of SARS-CoV2 from the pool of fungal metabolites using docking and molecular dynamics simulation. *Mol Divers* 1:3. <https://doi.org/10.1007/s11030-021-10220-8>
 17. Zeng X, Zhang P, He W, Qin C, Chen S, Tao L, Wang Y, Tan Y, Gao D, Wang B, Chen Z, Chen W, Jiang YY, Chen YZ (2018) NPASS: natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Res* 46:D1217–D1222. <https://doi.org/10.1093/nar/gkx1026>
 18. Jorgensen WL, Tirado-Rives J (1988) The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666. <https://doi.org/10.1021/ja00214a001>
 19. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236. <https://doi.org/10.1021/ja9621760>
 20. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W (2010) Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *J Chem Theory Comput* 6:1509–1519. <https://doi.org/10.1021/ct900587b>
 21. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J Med Chem* 49:6177–6196. <https://doi.org/10.1021/jm051256o>
 22. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69:146–148. <https://doi.org/10.1111/j.1747-0285.2007.00483.x>
 23. Wang W, Donini O, Reyes CM, Kollman PA (2001) Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein–ligand, protein–protein, and protein–nucleic acid noncovalent interactions. *Annu Rev Biophys Biomol Struct* 30:211–243. <https://doi.org/10.1146/annurev.biophys.30.1.211>
 24. Wang J, Hou T, Xu X (2006) Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Curr Comput Aided-Drug Des* 2:287–306. <https://doi.org/10.2174/157340906778226454>
 25. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 33:889–897. <https://doi.org/10.1021/ar000033j>
 26. Massova I, Kollman PA (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect Drug Discov Des* 18:113–135. <https://doi.org/10.1023/A:1008763014207>
 27. Chikhale RV, Sinha SK, Patil RB, Prasad SK, Shakya A, Gurav N, Prasad R, Dhaswadikar SR, Wanjari M, Gurav SS (2020) In-silico investigation of phytochemicals from *Asparagus racemosus* as plausible antiviral agent in COVID-19. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1784289>
 28. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: *Proceedings 2006 ACM/IEEE Conference Supercomput. SC'06*, ACM Press, New York, New York, USA, p 84. <https://doi.org/10.1145/1188455.1188544>
 29. Pires DEV, Blundell TL, Ascher DB (2015) pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem* 58:4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>
 30. Shukla A, Parmar P, Goswami D, Patel B, Saraf M (2021) Exemplifying an archetypal thorium–EPS complexation by novel thoriotolerant *Providencia thoriotolerans* AM3. *Sci Rep*. <https://doi.org/10.1038/s41598-021-82863-4>
 31. Shukla A, Parmar P, Goswami D, Patel B, Saraf M (2020) Characterization of novel thorium tolerant *Ochrobactrum intermedium* AM7 in consort with assessing its EPS–Thorium binding. *J Hazard Mater*. <https://doi.org/10.1016/j.jhazmat.2020.122047>
 32. Shukla A, Parmar P, Rao P, Goswami D, Saraf M (2020) Twin peaks: presenting the antagonistic molecular interplay of curcumin with LasR and LuxR quorum sensing pathways. *Curr Microbiol*. <https://doi.org/10.1007/s00284-020-01997-2>
 33. Patel D, Patel A, Patel M, Goswami D (2021) *Talaromyces pinophilus* strain M13: a portrayal of novel groundbreaking fungal strain for phytointensification. *Environ Sci Pollut Res* 28:8758–8769. <https://doi.org/10.1007/s11356-020-11152-w>
 34. Perlman S, Netland J (2009) Coronaviruses post-SARS: update on replication and pathogenesis. *Nat Rev Microbiol* 7:439–450. <https://doi.org/10.1038/nrmicro2147>
 35. Cheng VCC, Lau SKP, Woo PCY, Kwok YY (2007) Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clin Microbiol Rev* 20:660–694. <https://doi.org/10.1128/CMR.00023-07>
 36. Cuong D, Hung TM, Kim JC, Kim EH, Woo MH, Choi JS, Jeong H, Lee B, S. Min (2012) Phenolic compounds from *Caesalpinia sappan* heartwood and their anti-inflammatory activity. *ACS Publ* 75:2069–2075. <https://doi.org/10.1021/mp3003673>
 37. Shukla S, Mehta A, John J, Singh S, Mehta P, Vyas SP (2009) Antioxidant activity and total phenolic content of ethanolic extract of *Caesalpinia bonducella* seeds. *Food Chem Toxicol* 47:1848–1851. <https://doi.org/10.1016/j.fct.2009.04.040>
 38. Srinivasan R, Chandrasekar MJN, Nanjan MJ, Suresh B (2007) Antioxidant activity of *Caesalpinia digyna* root. *J Ethnopharmacol* 113:284–291. <https://doi.org/10.1016/j.jep.2007.06.006>
 39. Amin SA, Ghosh K, Gayen S, Jha T (2020) Chemical-informatics approach to COVID-19 drug discovery: Monte Carlo based QSAR, virtual screening and molecular docking study of some in-house molecules as papain-like protease (PLpro) inhibitors. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1780946>
 40. France A, Hognon C, Micolot T, García-Iriepa C, Iriepa I, Terenzi A, Grandemange S, Barone G, Marazzi M, Monari A (2020) Molecular basis of SARS-CoV-2 infection and rational design of potential antiviral agents: modeling and simulation approaches. *Cite This J Proteome Res* 19:4291–4315. <https://doi.org/10.1021/acs.jproteome.0c00779>
 41. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 64:4–17. <https://doi.org/10.1016/j.addr.2012.09.019>
 42. Broccatelli F, Salphati L, Plise E, Cheong J, Gobbi A, Lee ML, Aliagas I (2016) Predicting passive permeability of drug-like molecules from chemical structure: where are we? *Mol Pharm* 13:4199–4208. <https://doi.org/10.1021/acs.molpharmaceut.6b00836>
 43. Press B, Di Grandi D (2008) Permeability for intestinal absorption: Caco-2 assay and related issues. *Curr Drug Metab* 9:893–900. <https://doi.org/10.2174/138920008786485119>

44. Raghava KM, Lakshmi PK (2012) Overview of P-glycoprotein inhibitors: a rational outlook. *Braz J Pharm Sci* 48:353–367. <https://doi.org/10.1590/S1984-82502012000300002>
45. Oso BJ, Oyewo EB, Oladiji AT (2019) Influence of ethanolic extracts of dried fruit of *Xylopiya aethiopica* (Dunal) a rich on haematological and biochemical parameters in healthy Wistar rats. *Clin Phytoscience*. <https://doi.org/10.1186/s40816-019-0104-4>
46. Rao P, Shukla A, Parmar P, Rawal R, Patel B, Saraf M, Goswami D (2020) Proposing a fungal metabolite-Flaviolin as a potential inhibitor of 3CLpro of novel coronavirus SARS-CoV-2 identified using docking and molecular dynamics. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2020.1813202>
47. Parmar P, Shukla A, Rao P, Saraf M, Patel B, Goswami D (2020) The rise of gingerol as anti-QS molecule: darkest episode in the LuxR-mediated bioluminescence saga. *Bioorg Chem* 99:103823. <https://doi.org/10.1016/j.bioorg.2020.103823>
48. Kumar SP, Patel CN, Rawal RM, Pandya HA (2020) Energetic contributions of amino acid residues and its cross-talk to delineate ligand-binding mechanism. *Proteins Struct Funct Bioinforma* 88:1207–1225. <https://doi.org/10.1002/prot.25894>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.