ORIGINAL ARTICLE

Molecular docking studies, molecular dynamics and ADME/tox reveal therapeutic potentials of STOCK1N-69160 against papain-like protease of SARS-CoV-2

Olusola Olalekan Elekofehinti[1] · Opeyemi Iwaloye[1] · Sunday Solomon Josiah[2] · Akeem Olalekan Lawal[1] · Moses Orimoloye Akinjiyan[1] · Esther Opeyemi Ariyo[1]
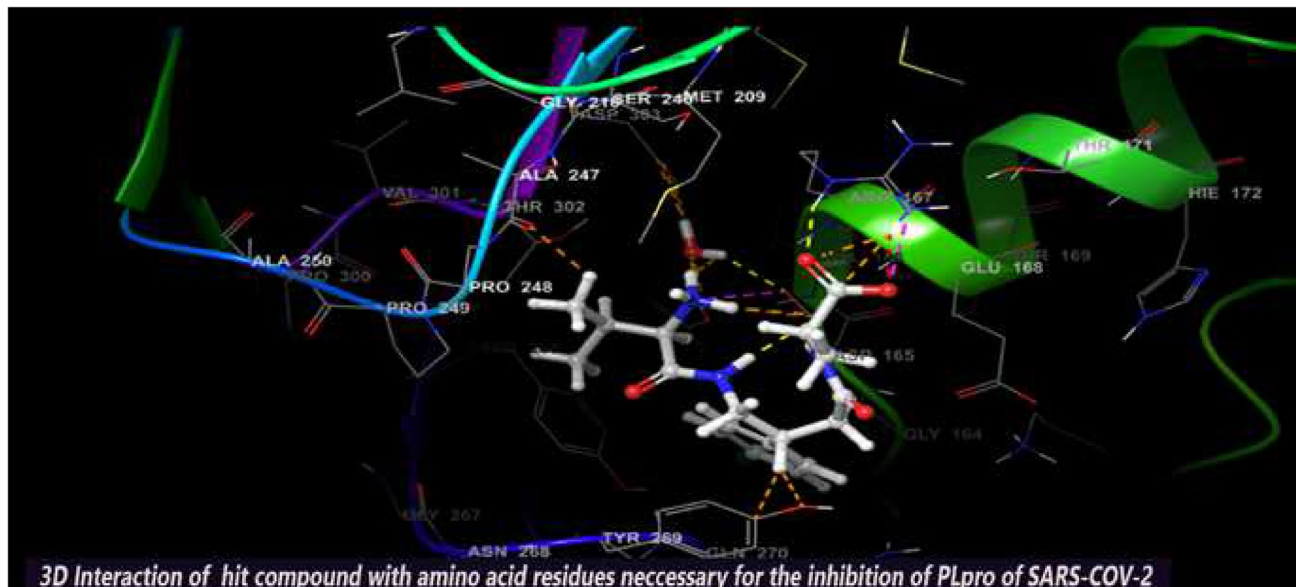
Abstract

SARS-CoV-2 is a new strain of Coronavirus that caused the pneumonia outbreak in Wuhan, China and has spread to over 200 countries of the world. It has received worldwide attention due to its virulence and high rate of infection. So far, several drugs have experimented against SARS-CoV-2, but the failure of these drugs to specifically interact with the viral protease necessitates urgent measure to boost up researches for the development of effective therapeutics against SARS-CoV-2. Papain-like protease (PLpro) of the viral polyproteins is essential for maturation and infectivity of the virus, making it one of the prime targets explored for SARS-CoV-2 drug design. This study was conducted to evaluate the efficacy of ~ 50,000 natural compounds retrieved from IBS database against COVID-19 PLpro using computer-aided drug design. Based on molecular dock scores, molecular interaction with active catalytic residues and molecular dynamics (MD) simulations stud- ies, STOCK1N-69160 [(S)-2-((R)-4-((R)-2-amino-3-methylbutanamido)-3-(4-chlorophenyl) butanamido) propanoic acid hydrochloride] has been proposed as a novel inhibitor against COVID-19 PLpro. It demonstrated favourable docking score, the free energy of binding, interacted with key amino acid residues necessary for PLpro inhibition and also showed significant moderation for parameters investigated for ADME/tox (Adsorption, distribution, metabolism, excretion and toxicological) properties. The edge of the compound was further established by its stability in MD simulation conducted for 30 ns employ- ing GROMACS software. We propose that STOCK1N-69160 is worth further investigation for preventing SARS-CoV-2.

✉ Olusola Olalekan Elekofehinti
ooelekofehinti@futa.edu.ng

1   Bioinformatics and Molecular Biology Unit, Department of Biochemistry, Federal University of Technology Akure, Ondo, Nigeria

2   Phytomedicine Biochemical Pharmacology and Toxicology Unit, Department of Biochemistry, Federal University of Technology Akure, Ondo, Nigeria

Graphic abstract

3D Interaction of hit compound with amino acid residues neccessary for the inhibition of PLpro of SARS-COV-2

Introduction

In April 2018, World Health Organization (WHO) made a list of pathologies which entails middle east respiratory syndrome (MERS), severe acute respiratory syndrome (SARS) and Disease X as diseases with pandemic potential caused by unknown pathogens [1, 2]. In December 2019, there was an outbreak of pneumonia with an unknown aetiology in Wuhan, China and was considered as the first Disease X following the announcement by WHO [3, 4]. The first confirmed case of this new disease was traced back to 17 November 2019 [5].

Not long after, a novel coronavirus, 2019-nCoV, as coded by WHO [6], was identified as the pathogen causing the Coronavirus disease, COVID-19 [7, 8]. A group of scientist showed that 2019-nCoV has 79.5% sequence identity to SARS-CoV and 96% sequence correlation to bat Coronavirus (SL-CoV-RaTG13) [9]. This viral disease was later renamed as SARS-CoV-2 by the coronaviridae study group (CSG) of the International Committee on Taxonomy of Viruses (ICTV) [10]. As of 25 June 2020, not less than 9.5 million cases of SARS-CoV-2 viral infection has been reported and documented worldwide; spreading across 188 countries of the 196 officially recognized nations by United Nation (UN) resulting in ~ 483, 000 deaths. Impressively about 4.5 million patients have recuperated [11]. Sadly, the infection is still multiplying and neither vaccines nor specific antiviral treatment has been approved to be effective.

Coronaviruses (CoVs) have been regarded as the largest RNA viruses identified, which belongs to the family *coronaviridae*. They are divided into four genera, α-, β-, δ- and γ-coronaviruses, while the β-coronaviruses are further divided into A, B, C and D lineages. The seven CoVs that can infect humans (HCoVs) include HCoV-229E and HCoV-NL63 in the α-coronaviruses, HCoV-OC43 and HCoV-HKU1 in the β- coronaviruses lineage A, SARS-CoV and SARS-CoV-2 in the β- coronaviruses lineage B (β-B coronaviruses) and MERS-CoV in the β-coronaviruses lineage C [12, 13]. SARS-CoV-2 mostly affects the lungs (it also affects other organs such as kidney, intestine and blood vessels) due to its access to the host cells via a critical enzyme responsible for regulating blood pressure known as angiotensin-converting enzyme 2 (ACE2), attached to the cell membrane of various organs [14, 15]. The virus enters the host cell by the use of spike protein, a unique surface glycoprotein that connects the virus to human ACE2 [16]. Once the virus binds to ACE2 (receptor), it enters host cells through endocytosis or membrane fusion. The viral RNA enters the nucleus of human cells to replicate and to make viral proteins and new viral particles which are then released to infect new cells [17]. People infected with SARS-CoV-2 showed fever as the predominant symptom [18]. Other possible symptoms include cough, loss of appetite, shortness of breath (in a severe case of people with an immunocompromised system), loss of smell and taste, vomiting, organ failure [18–20].

The nucleocapsid, spike, membrane and envelope are the structural proteins of SARS-CoV-2, and these proteins together formed the envelope of SARS-CoV-2 [17]. They are responsible for the virus assembly and budding, host cellular response to viral infection and the shape of the virus envelop [21]. While the polymers spike proteins are most notable for giving SARS-COV-2 a crown-like shape (Coronavirus), nucleocapsid protein is involved in the processing of the viral genetic make-up [21].

Two therapeutic targets for SARS-CoV-2 encoding cysteine proteases have been defined; they are papain-like protease (PLpro) and the 3-chymotrypsin-like protease (3Clpro, also known as the main protease-Mpro). The processing of the viral polyproteins is essential for maturation and infectivity of the virus [22]. Due to the crucial roles played by these proteases in the life cycle of the virus, they serve as prime targets for drug design. Recently published articles on COVID-19 employed the use of SARS-CoV-2 PLpro to screen drug-like compounds approved by the Federal Drug Agency (FDA) as novel inhibitors [23–25]. Hence, the present study exploited the same target (SARS-CoV-2 PLpro) to reveal potent anti-COVID-19 agents from a library of natural products. The computational methods employed in this study include molecular docking studies, ADME/tox (Adsorption, distribution, metabolism, excretion and toxicological) and molecular dynamics simulation which aided in the discovery of lead compounds.

Materials and methods

Preparation of protein crystal structure via *protein preparation wizard*

The X-ray crystallographic structures of SARS-CoV PLpro and SARS-CoV-2 PLpro were retrieved from the protein databank (https://www.rcsb.org/), with accession ID 3e9s and 7cmd. The protein preparation wizard of the Glide software (Schrödinger Suite 2018–1) was used to rectify specific errors such as missing hydrogen atoms in the protein during X-ray crystallography. The protocol used in this study has been previously described [26, 27]. The bond orders were assigned, and possible missing hydrogen atoms in the PBD structure were added. Also, disulphide bonds were created between two sulphur atoms in proximity and other options were left as default. Full energetic optimization was performed in the final refinement step using OPLS3 force field and the RMSD of heavy atoms was set at 0.3 Å [28].

Preparation of library of natural compounds

About fifty thousand (50,000) compounds from natural origin were downloaded from IBS database (https://www.ibscr

een.com/natural.shtml) and uploaded on the workspace of Maestro molecular Interface (version 11v5) of Schrödinger suite. Low energy 3D conformers with satisfactory bond lengths and angles for each two-dimensional structure were generated [29]. The possible ionization states for each ligand structure were generated at a physiological pH of 7.2 ± 0.2. All other options were kept as default and the ligands were minimized using an optimized potential liquid simulation (OPLS3) force field [28]. The Ramachandran plot of PLpro from SARS-CoV-2 is illustrated in Fig. 1.

Molecular docking studies

The glide grid file was generated via receptor grid generation panel by picking the co-crystallized ligand located at the active site of the protein to automatically reveal x, y, z coordinates (-31.02, 21.89 and 30.09), respectively. Glide uses the position and size of the ligand to calculate a default centre and a default size for the region for which grids will be calculated.

The virtual screening of the ~50,000 compounds was carried out using three hierarchical GLIDE docking filtering, namely: high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP). HTVS docking is intended for the rapid screening of vast numbers of ligands. HTVS has much more restricted conformational sampling than SP docking, standard-precision (SP) docking is appropriate for screening ligands of unknown quality in large numbers. Extra-precision (XP) docking and scoring is a more robust and discriminating procedure, which takes longer to run than SP. XP is designed to be used on ligand poses that have been determined to be high-scoring using SP docking. The database was run through SP docking. First, the top 30% of the final poses was further docked using XP,
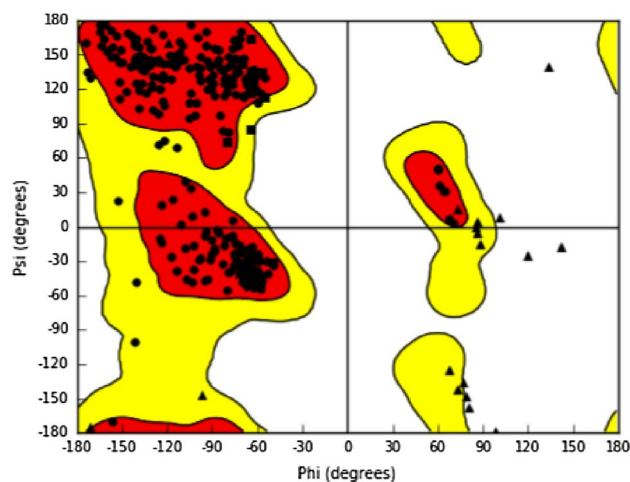

Fig. 1 Ramachandran plot of the retrieved crystal protein structure (3E9S) from PDB repository

to score the compounds binding affinity via more expensive docking simulation on worthwhile poses.

To accurately predict the binding affinity of the novel inhibitors with the protein crystal, induced-fit docking (IFD) was implemented. IFD protocol aims to improve the docking of ligands in which it is believed that the receptor adjusts significantly to the presence of the ligand. The protocol performs a constrained minimization of the receptor followed by initial Glide docking of the ligands using a softened potential. A selected set of the docked poses is passed on to prime for a refinement step [29]. After a prime side-chain prediction and minimization, the best receptor structures for each ligand are passed back to Glide for re-docking of the ligand [30].

Prime/MM-GBSA calculations

The docking complexes were carried out using MM-GBSA calculation [31]. The docked complexes were minimized by using local optimization feature in Prime. The OPLS-AA 2005 force field was employed to determine the binding energy for a set of receptor and ligand. The following equation was used to calculate the binding free energy:

$$\Delta GBind = \Delta EMM + \Delta GSolv + \Delta GSA.$$

Where ΔEMM is the variance between the minimized energy of the protein–ligand complexes, while $\Delta GSolv$ is the variation between the GBSA solvation energy of the protein–ligand complexes and the sum of the solvation energies for the protein and ligand. In ΔGSA contains some of the surface area energies in the protein and ligand and the difference in the surface area energies for the complexes. The minimization of the docked complexes was done using a local optimization feature of prime.

ADME/tox

Qikprop of Maestro v11.5 was used to determine the physicochemical and pharmacokinetic parameters necessary for a drug to be considered as drug candidates [32]. The toxicological properties of the compounds were predicted by an online server ADMETsar [33].

Molecular Dynamics (MD) simulation

The MD simulation was performed using GROMACS 4.6.5 MD package running on Ubuntu 18.04 (Hp-635 CPU@ 2.5 GHz Intel Core™ i5 RAM 6 GB) with GROMOS force field [34]. SARS-CoV-2 PLpro was analyzed for MD simulation. The process started with initial cubic solvation with single point charge water model, followed by ionization and neutralization in Na and Cl ions (0.15 M). The solvated system was subjected to further energy minimization to remove stearic conflicts between protein and water molecules using the steepest descent integrator. The minimized model was subjected to position-restrained MD under NPT conditions by keeping the number of particles (N), the system pressure (P) and the temperature (T) constant. This was executed for 100,000 steps for a total of 200 ps. The temperature of 300 k and a pressure of 1 atm were maintained. After equilibration of the system with constant temperature and pressure, the production MD run of 10,000,000 steps was performed for 30 ns to carry out the structural analysis on PLpro and its complex with the lead compound.

Results and discussion

The COVID-19 pandemic has taken the world by storm and the rate of infection continues to spread like wide-fire. Despite the low mortality rate (about 5.1% of the confirmed case) [35], it has continued to pose significant problems globally. Experts are worried about the spread of COVID-19 in Africa because of the inadequate healthcare systems [36]. As of 21 July 2020, the number of reported cases of death from COVID-19 stands at about 600,000 with about 16 million confirmed cases. The economy of numerous countries, however, could fall into recession due to this pandemic if the spread of this virus continues in the next couple of months [35]. Hence urgent attention is needed from scientists across the globe to find a possible cure for this viral infection [35]. In many viruses, proteases play essential roles in viral replication; therefore, proteases are often used as protein targets during the development of antiviral therapeutics [37]. PLpro antagonistic activities have been shown to block the production of essential cytokines involved in the activation of the host's innate immune response against viral infection, including the Type I interferon b (IFNb) and chemokines [38].

Molecular docking studies, Post docking analysis and binding free energy

Molecular docking is of great importance in the planning and design of new drugs. It correctly predicts the experimental binding mode and affinity of a native molecule within the binding site of the drug target [39]. To study the molecular basis of interaction and conformation of hit compounds within the binding pose of our protein target, glide XP docking and induced fit docking (IFD) were carried out. Table 1 revealed the 2D structure of lead compounds with their IUPAC name. The results of the molecular docking are enlisted in Table 2. STOCK1N-69160 had the most favourable docking score with PLpro (SARS-CoV) with docking score of − 8.414 kcal/mol. Next in the ranking are

Table 1 IUPAC name and 2D structure of lead compounds

s/n	Compound name	IUPAC name	2D structure
1	STOCK1N-69160	(S)-2-((R)-4-((R)-2-amino-3-methylbutanamido)-3-(4-chlorophenyl)butanamido)propanoic acid hydrochloride	
2	STOCK1N-68604	N-(2,4-dimethylphenyl)-2-((2S,11aS)-2-hydroxy-5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)acetamide	
3	STOCK1N-66718	(S)-2-((S)-2-(3-oxo-1,2,3,4-tetrahydroquinoxaline-1-carboxamido)-3-phenylpropanamido)pentanedioic acid	
4	STOCK1N-69717	2-((2S,11aS)-2-hydroxy-5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)-N-phenylacetamide	
5	STOCK1N-84011	3-(5-(2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-4,8-dioxo-4,8,9,10-tetrahydropyrano[2,3-f]chromen-10-yl)furan-2-yl)benzoic acid	
6	STOCK1N-59841	(S)-2-((S)-2-(9-oxo-9H-fluorene-4-carboxamido)-3-phenylpropanamido)propanoic acid	
7	STOCK1N-67398	methyl 2-(8-(((3,4-dimethoxybenzyl)amino)methyl)-7-hydroxy-4-methyl-2-oxo-2H-chromen-3-yl)acetate hydrochloride	
8	STOCK1N-65477	(S)-2-(2-(3-oxo-1,2,3,4-tetrahydroquinoxaline-1-carboxamido)-3-phenylpropanamido)acetic acid	

Table 2 Docking, post docking analysis and binding free energy (Prime/MM-GBSA) of the compounds with papain-like protease of SARS-CoV

s/n	Compound name	Docking score	IFD score	Interacting residues	No of H-bond	Binding free energy
1	STOCK1N-69160	− 8.414	− 673.83	**Tyr265*, Tyr269***, Arg167, Asp165, Gln270, Gly164	6	− 32.47
2	STOCK1N-68604	− 8.011	− 672.67	Gln270, Tyr269, Gly267, **Tyr265***	4	− 65.76
3	STOCK1N-66718	− 7.940	− 676.26	Arg167, Tyr274, Tyr269, **Tyr269*,** Asp165, Leu163, Gln270	7	− 13.64
4	STOCK1N-69717	− 7.819	− 678.27	Tyr274, Arg167, Asp165	3	− 57.22
5	STOCK1N-84011	− 7.803	− 670.66	**Tyr265, Tyr269,** Tyr269, Glu168	3	− 59.56
6	STOCK1N-59841	− 7.701	− 672.75	Arg167, **Arg167***, Asp165, Tyr265, Gln270	6	− 44.12
7	STOCK1N-67398	− 7.268	− 675.98	**Tyr269***, Gln270, Asp165, Tyr265	6	− 76.80
8	STOCK1N-65477	− 7.211	− 674.25	Arg167, Asp165, Gln270, Glu168	6	− 52.58

"*" are the residues that formed pi-pi interactions with the phenyl ring of the respective compounds

STOCK1N-68604 and STOCK1N-66718 (− 8.011 kcal and − 7.940 kcal/mol). A total of 12 natural compounds emerged as promising in term of docking score.

Induced fit docking algorithm which makes use of Glide and refinement module for calculating the binding energy of ligand with receptor-based on concomitant structural changes of the ligand within the binding pocket of the protein target was further employed for accurate prediction [29, 30]. In Glide docking algorithm (HTVS, XP and SP), the receptor is held rigid when the ligand is docked into its binding site. It is popularly believed that the rigidity of the receptor can give invalid docking score because proteins undergo specific movements (side-chain or backbone) upon binding with small molecules. These structural changes, however, allow the receptor to modify its binding site so that it more closely conforms to the shape and binding mode of the ligand. This IFD method enables the receptor to change its binding pocket so that it conforms more closely to the ligand shape and binding mode. Unlike the results of docking score which reported STOCK1N-69160 as hit with highest binding affinity, the output of IFD revealed that STOCK1N-69717, STOCK1N-66718 and STOCK1N-67398 displayed more favourable interactions with PLpro (SARS-CoV) achieving IFD score of -678.27 kcal/mol, -676.26 kcal/mol and -675.98 kcal/mol, respectively.

In contrast, STOCK1N-69160 showed IFD score of -673.83 kcal/mol. Several studies have found a variety of natural products as inhibitors of viral PLpro activity [40–42]. Docking score resonates the inhibitory prowess of ligands in the protein–ligand complex [43]. Therefore, this result corroborates that the top docked natural compounds may have inhibitory activities against PLpro of SARS-COV.

Interaction of the small molecules with amino acid residues at the active site of PLpro is necessary for the inhibition of the protease [44, 45]. The residues in the binding pocket of PLpro are Tyr265, Tyr269, Arg167, Asp165, Gln270, Gly164, Glu168, Tyr274, Tyr269, Gly267, Thr302,

Ala247, Leu163, Pro248, Met209, Pro249 and Val166. The results of docking analysis (Table 2, Fig. 2a) showed that STOCK1N-69160 formed H-bonds with PLpro residues Arg167, Asp165, Gln270 and Gly164. STOCK1N-68604 formed H-bonds with Gln270, Tyr269 and Gly267. STOCK1N-66718 formed the highest number of H-bond with the residues Arg167, Tyr274, Tyr269, Asp165, Leu163 and Gln270. Overall the lead compounds showed similar H-bond interaction with three residues Asp165, Arg167 and Gln270. Docking analysis results, including the H-bonds and pip-pi stacking that interacted with 3e9s amino acids, are presented in Table 2. All of the H-bonds interacted with amino acids in the SARS-COV-2 PLpro active site. The binding energy results are related to the number of H-bonds formed with the active site of SARS-COV-2 PLpro.

The docking analysis of the lead compounds was validated by calculating the binding free energy of protein–ligand complexes. Two different studies have demonstrated the reliability of MM-GBSA (Molecular mechanism surface area continuum salvation) post docking method for rating affinity of the protein–ligand complex [46, 47]. The energy released (ΔG_{bind}) due to bond formation, or rather interaction of the ligand with protein is in the form of binding energy and it determines the stability of any given protein–ligand complex. The free energy of a favourable reaction is negative. The results of the energetic analysis of the complexes are provided in Table 2. Prime MM-GBSA (ΔG_{bind}) range was from -76.80 kcal/mol (STOCK1N-67398) to -13.64 kcal/mol (STOCK1N-66718). Among the lead compounds, STOCK1N-67398 and STOCK1N-68604 had the highest binding score compared to others. The results showed that significant contributions to the ligand-binding were nonpolar salvation terms (ΔG_{lipo}), van der Waals (ΔG_{vdW}) and covalent energy ($\Delta G_{covalent}$).

In this study, we utilized the x-ray crystal structure of PLpro from the SARS virus to perform virtual screening, due to unavailability of PLpro X-ray structure from

Fig. 2 **a** 2D interaction of STOCK1N-69160 [(S)-2-((R)-4-((R)-2-amino-3-methylbutanamido)-3-(4-chlorophenyl) butanamido) propanoic acid hydrochloride] with amino acid residue within the active site of papin-like protease of SARS-CoV, **b** 2D interaction of STOCK1N-69160 [(S)-2-((R)-4-((R)-2-amino-3-methylbutanamido)-3-(4-chloro-phenyl) butanamido)propanoic acid hydrochloride] with amino acid residue within the active site of papin-like protease of SARS-CoV-2
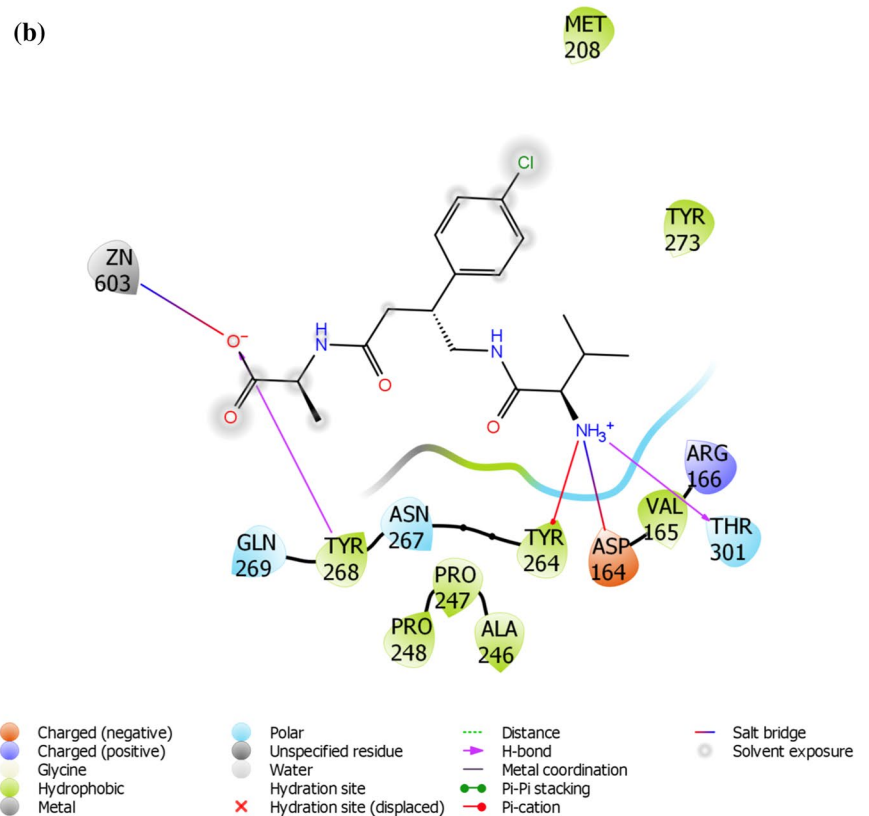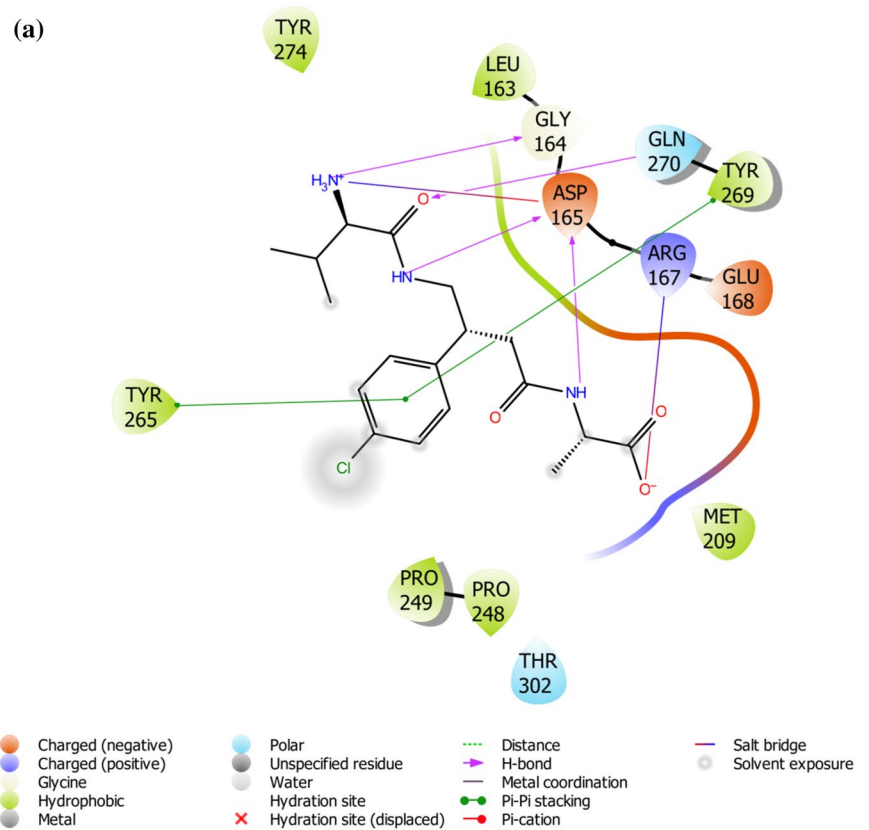
SARS-COV-2 when the study began. However, many efforts have been made to characterized SARS-COV-2 crystal structure using X-ray diffraction. Hence, we now have PLpro X-ray structure from SARS-COV-2 in the protein databank repository. Therefore, we further validated the results obtained from this study by docking STOCK1N-69160 with SARS-COV-2 (PDB ID: 7cmd) and the docking and post docking results are shown in Table 3. STOCK1N-69160 had docking and IFD score of—9.084 kcal/mol and -655.84 kcal/mol, respectively, while interacting with the residues Tyr268 and Thr301 located at the active site of SARS-COV-2. 2D interaction is shown in Fig. 2a.

Determination of ADME/tox for lead compounds

Several rules have been employed to determine the drug-likeness of compounds, with the most popular being the Lipinski Rule of Five (RO5) [48]. The rule states that for a compound to be considered as the drug-like molecule it must not violate more than one of the following criteria: molecular weight < 500 Da, octanol–water partition coefficient < 5, hydrogen bond donor ≤ 5, hydrogen bond acceptor ≤ 10. Results from this study (Table 2) revealed that all the compounds except for STOCK1N-69717 and STOCK1N-84011 obeyed Lipinski rules. As more compounds break some of these rules and enter the market [49], attempts have been made to modify these parameters [50–53]. The physicochemical properties calculated by Qikprop are shown in Table 4. The pharmacokinetic properties of the compounds were analyzed by predicting the binding to human serum albumin (QPlogKhsa), IC_{50} value for the blockage of HERG K + channels (QPlogHERG), MDCK cell permeability in nm/sec (QPPMDCK), brain/blood partition coefficient (QPlogBB) and Caco-2 cell permeability in nm/sec (QPPCaco). Except for STOCK1N-84011 with a polar surface area of 212.758, all other compounds fall within the recommended range for molecular weight, predicted octanol/water partition coefficient, number of hydrogen bond donors and number of hydrogen bond acceptors. The lead compounds were within the values recommended for QPlogKhsa, STOCK1N-68604 and STOCK1N-67398 were in the range for the blockage of HERG K + channels, Caco-2 cell permeability and MDCK cell permeability, the compounds STOCK1N-69717 and STOCK1N-84011 did not comply with the predicted brain/blood partition coefficient (Range from –3.0 to 1.2). Organ and genomics

toxicological assessment were also determined using an online server. The parameters taken into considerations were carcinogenicity, eye corrosion, eye irritation, Ames mutagenesis, micronuclear, hepatotoxicity androgen receptor binding and PPAR gamma. The results depict that only STOCK1N-69160 showed significant moderations for these parameters (Table 5).

Molecular dynamics (MD) simulation

After careful examination of the docking, post-docking analysis and ADME/tox (Adsorption, distribution, metabolism, excretion and toxicological) properties of the natural compounds, STOCK1N-69160 was chosen as an ideal drug candidate against SARS-CoV-2 (PLpro); therefore, to investigate the validity of the docking data, the docked complex of STOCK1N-69160 with PLpro was subjected to 30 ns MD simulation. Root mean square deviation (RMSD) provides information about a protein concerning its backbone structure [54]. The RMSD for STOCK1N-69160-complex was found to be approximate 0.44 nm and small initial fluctuation was observed, which was stabilized after 0.5 ns due to the reduction in amplitude of fluctuation with time. The protein in Apo-state, however, achieved stabilization between 0.7 ns and 2.5 ns (Fig. 3a).To observe the dynamic behaviour of residues of different complexes of PLpro, atomic positional fluctuations of backbone residues of each PLpro was computed. The root means square fluctuation (RMSF) gives information about the dynamic behaviour of residues. The complexes have less fluctuation depicting a lower degree of flexibility in backbone residues. Fluctuations only occur between 1900–2000 and between 2300–2400 (Fig. 3b). The Rg (radius of gyration) was also calculated for the STOCK1N-69160-complex and PLpro in apo-state to assess the compactness of the complex and protein structure. The Rg range of the STOCK1N-69160 bound PLpro complex is between 2.37 and 2.52 nm. This complex appeared to be stable between 0 and 20 ns. In Apo-state, the significant fluctuation was observed towards the end of the simulation, which corresponds to a decrease in Rg values (2.4–2.25nm^2) (Fig. 3c). The compartment of the hydrophobic core of the PLpro crystal structure was examined by investigating the change in solvent-accessible surface (SASA). The SASA range of STOCK1N-69160 bound complex lies between 156 and 180 nm^2. This complex showed stability all through the duration of the simulation (Fig. 3d). Additionally, unbound

Table 3 Docking, post docking analysis and binding free energy (Prime/MM-GBSA) of STOCK1N-69160 with papain-like protease of SARS-CoV-2

s/n	Compound name	Docking score	IFD score	Interacting residues	No of H-bond	Binding free energy
1	STOCK1N-69160	− 9.084	− 655.84	Tyr268, Thr301,	2	− 12.56

Table 4 Physicochemical and pharmacokinetics properties of the lead compounds

s/n	Compound name	[a]mol_MW	[b]donorHB	[c]accptHB	[d]QPlogPo/w	[e]PSA	[f]RuleOfFive
1	STOCK1N-69160	383.874	4.25	7.25	− 0.666	140.754	0
2	STOCK1N-68604	393.441	2	10.2	2.083	104.865	0
3	STOCK1N-66718	468.465	3.5	9.5	1.473	199.74	1
4	STOCK1N-69717	638.734	5	14.75	1.576	199.588	2
5	STOCK1N-84011	542.455	4	9.5	2.018	212.758	2
6	STOCK1N-59841	442.47	1.5	7.5	3.617	144.614	0
7	STOCK1N-67398	427.453	2	8.25	2.889	122.859	0

s/n	Compound name	[a]QPlogKhsa	[b]QPlogHERG	[c]QPPMDCK	[d]QPlogBB	[e]QPPCaco
1	STOCK1N-69160	− 0.822	− 1.349	8.342	− 1.548	2.431
2	STOCK1N-68604	− 0.173	− 5.65	324.327	− 0.915	676.635
3	STOCK1N-66718	− 0.949	1.177	0.476	− 2.638	0.433
4	STOCK1N-69717	− 0.513	− 4.049	10.348	− 3.652	12.142
5	STOCK1N-84011	0.185	− 3.957	0.184	− 3.525	0.536
6	STOCK1N-59841	0.06	− 3.826	12.305	− 2.108	15.768
7	STOCK1N-67398	0.268	− 6.786	44.501	− 1.228	98.106
8	STOCK1N-65477	− 0.641	− 1.73	5.536	− 2.3	5.842

[a]Molecular weight of the molecule (Range:130.0–725.0)

[b]Predicted octanol/water partition coefficient. (Range: − 2.0–6.5)

[c]Number of hydrogen bond donors (Range: 0.0–6.0)

[d]Number of hydrogen bond acceptors (Range: 2.0–20.0)

[e]PSA:Van der Waals surface area of polar nitrogen and oxygen atoms. Range from 7.0 to 200.0

[f]Number of violations of Lipinski's rule of five (Range: maximum is 4)

[a]QPlogKhsa:Prediction of binding to human serum albumin. Range from − 1.5 to + 1.5

[b]QPlogHERG:Predicted IC50 value for blockage of HERG K + channels. Concern below –5

[c]QPPMDCK:Predicted apparent MDCK cell permeability in nm/sec. MDCK cells are considered to be a good mimic for the blood–brain barrier. QikProp predictions are for non-active transport. < 25 poor, > 500 great

[d]QPlogBB:Predicted brain/blood partition coefficient. Range from − 3.0 to 1.2

[e]QPPCaco: Predicted apparent Caco-2 cell permeability in nm/sec. Caco-2 cells are a model for the gutblood barrier. QikProp predictions are for non-active transport. < 25 poor, > 500 great

PLpro structure showed large fluctuation between 20 and 30 ns. The total number of hydrogen bonds produced during the simulation period of 30 ns is illustrated in Fig. 3e.

Conclusion

Molecular docking studies, ADME/tox and MD simulation, were performed with ~ 50,000 selected natural compounds from IBS database to evaluate prospective antiviral agents against SARS-COV-2. The analysis revealed 8 top docked compounds which were further filtered by post docking analysis. Significantly, STOCK1N-69160 satisfied all the parameters investigated such as docking score, glide energy, ADME/tox and trajectories analysis. We, therefore, propose that the inhibitory potentials of STOCK1N-69160 against PLpro of SARS-COV-2 should further be explored in the subsequent wet experiment. This study facilitates the initiation of the SARS-COV-2 discovery process for antiviral activity.

Table 5 Assessment of organ and genomics toxicity by AMETsar

s/n	Compound name	A	B	C	D	E	F	G	H
1	STOCK1N-69160	−	−	−	−	−		+ −	−
2	STOCK1N-68604	−	−	−	−	+		+ +	+
3	STOCK1N-66718	−	−	−	+	+		+ +	+
4	STOCK1N-69717	−	−	−	−	−		+ +	+
5	STOCK1N-84011	−	−	−	+	+		+ +	+
6	STOCK1N-59841	−	−	−	−	+		+ +	+
7	STOCK1N-67398	−	−	−	+	+		+ +	+
8	STOCK1N-65477	−	−	−	−	−		+ +	+

Active = (+); Inactive = (−)

A = Carcinogenicity

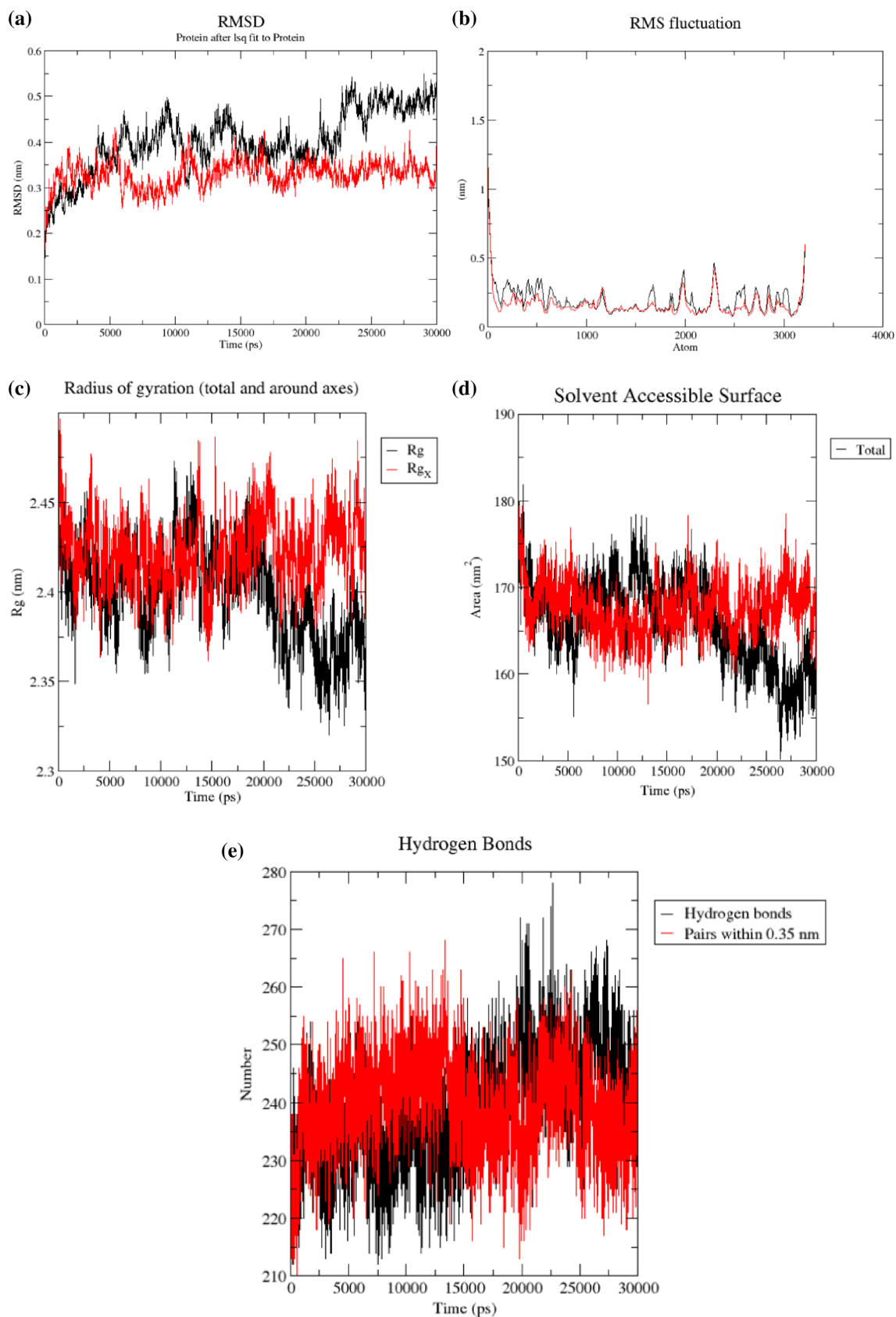B = Eye corrosion

C = Eye irritation

D = Ames mutagenesis

E = Micronuclear

F = Hepatotoxicity

G = Androgen receptor binding

H = PPAR gamma

Fig. 3 a The RMSD of the backbone atom of 3e9s (SARS-CoV ▶ PLpro) in Apo (red) and bound state (black) over a time period of 30 ns denoting a value of 4.4 Å, b RMSF plot of backbone atoms of 3e9s (SARS-CoV PLpro) in Apo and bound state over a time period of 30 ns, c Radius of gyration (Rg) plot of Apo SARS-CoV PLpro (Black) compared to complexed state (Red) for 30 ns Simulation, d solvent accessible surface (SASA) of STOCK1N-69160 bound PLpro complex with respect to the unbound (apo) reference structure, e The total number of the hydrogen bonds produced during the simulation period of 30 ns

(a) RMSD
Protein after lsq fit to Protein

(b) RMS fluctuation

(c) Radius of gyration (total and around axes)

(d) Solvent Accessible Surface

(e) Hydrogen Bonds

Compliance with ethical standards

Conflict of interest Authors declare no conflict of interests.

References

1. World Health Organization (2015) Blueprint for R&D preparedness and response to public health emergencies due to highly infectious pathogens. https://www.who.int/medicines/ebolatreatment/WHO-list-of-top-emerging-diseases/en/

2. World Health Organization (2019) Prioritizing diseases for research and development in emergency contexts https://www.who.int/activities/prioritizing-diseases-for-research-anddevelopment-in-emergency-contexts

3. Hui DS, Azhar E, Madani TA, Ntoumi F, Kock R, Dar O et al (2020) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 201 9 novel coronavirus outbreak in Wuhan, China. International Journal of Infectious Diseases 91:264–266. https://doi.org/10.1016/j.ijid.2020.01.009

4. World Health Organization (2020) Director-General's opening remarks at the media briefing on COVID-19. https://www.scmp.com/news/china/society/article/3074991/coronavirus-china-first-confirmed-covid-19-case-traced-back

5. World Health Organization (2020) Coronavirus disease. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

6. Zhu N, Zhang D, Wang W, Li X et al (2020) A novel Coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 382:727–733. https://doi.org/10.1056/NEJMoa2001017

7. Wu F, Zhao S, Yu B, Chen Y, Wang C, Song Z et al (2020) A new coronavirus associated with human respiratory disease in China. Nature 579:265–269. https://doi.org/10.1038/s41586-020-2008-3

8. Harapan H, Naoya I, Amanda Y, Wira W, Synat K et al (2020) Coronavirus Disease (COVID-10): Aliterature review. Nature 579:270–273. https://doi.org/10.1016/j.jiph.2020.03.019

9. Gorbalenya A, Baker SC, Groot JR, Baric R, Drosten C, Gulyaeva AA et al (2020) The species Severe acute respiratory syndromerelatedcoronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. https://doi.org/10.1038/s41564-020-0695-z

10. Johns Hopkins University (2020) COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE). https://www.coronavirus.jhu.edu

11. Zhou P, Yang X, Wang X, Xian W, Ben H et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579:270–273. https://doi.org/10.1038/s41586-020-2012-7

12. Shuai X, Meiqin L, Chao W et al (2020) Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targetingits spike protein that harbors a high capacity to mediate membrane fusion. Cell Res 30:343–355. https://doi.org/10.1038/s41422-020-0305-x

13. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H (2004) Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus A first step in understanding SARS pathogenesis. J Path. 203(2):631–637. https://doi.org/10.1002/path.1570

14. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N (2000) A Novel AngiotensinConverting Enzyme-Related Carboxypeptidase (ACE2) Converts Angiotensin I to Angiotensin-9. Circ Res 87(5):1–9. https://doi.org/10.1161/01.RES.87.5.e1

15. Letko M, Marzi A, Munster V (2020) Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nature Microbiology 5(4):562–569. https://doi.org/10.1038/s41564-020-0688-y

16. Chan JF, Kok K, Zhu Z, Chu H, To KK, Yuan S, Yuen K (2020) Genomic characterization of the 201 9 novel humanpathogenic coronavirus isolated from a patient with a typical pneumonia after visiting Wuhan. Emerging Microbes & Infections 9(1):221–236. https://doi.org/10.1080/22221751.2020.1719902

17. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL et al (2020) The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 1 48 studies from 9 countries. PLoS ONE 15(6):e0234765. https://doi.org/10.1371/journal.pone.0234765

18. Wei XS, Wang X, Niu YR, Ye LL, Peng WB, Wang ZH et al (2020) Clinical Characteristics of SARS-CoV-2 Infected Pneumonia with Diarrhea. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3546120

19. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR (2020) Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. Int J Antimicrob Agents 55(3):105924. https://doi.org/10.1016/j.ijantimicag.2020.105924

20. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ (2003) The coronavirus spike protein is a classI virus fusion protein: structural and functional characterization of the fusion core complex. Journal 77:8801–8811

21. Chen Y, Qianyun L, Deyin G (2020) Coronaviruses: genome structure, replication and pathogenesis. J Med Virology. https://doi.org/10.1002/jmv.25681

22. Xu Z, Cheng P, Yulong S, Zhengdan Z, Kaijie M, Xiaoyu W, Weiliang Z (2020) Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease byan integrative approach combining homology modelling, molecular docking and binding free energy calculation. BioRxiv. https://doi.org/10.1101/2020.01.27.921627

23. Liu X, Xiu-Jie W (2020) Potential inhibitors for 2019-nCoV coronavirus M protease from clinically approved medicines. bioRxiv. https://doi.org/10.1101/2020.01.29.924100

24. Li Y, Jinyong Z, Ning W, Haibo L, Yun S, Gang G, Kaiyun L, Hao Z, Quanming Z (2020) Therapeutic Drugs Targeting 2019-nCoV Main Protease by High Throughput Screening. bioRxiv. https://doi.org/10.1101/2020.01.28.922922

25. Elekofehinti OO, Iwaloye O, Famusiwa CD, Rocha JBT (2020) Identification of Main Protease of coronavirus SARS-COV-2 (Mpro) Inhibitors from Melissa officinalis. Curr Drug Discov Technol. https://doi.org/10.2174/1570163817999200918103705

26. Iwaloye O, Elekofehinti OO, Oluwarotimi EA, Babatomiwa K, Momoh I (2020) In silico molecular studies of natural compounds as possible anti-Alzheimer's agents: ligand-based design. Network Modeling Analysis in Health Informatics and Bioinformatics 9:54

27. Iwaloye O, Elekofehinti OO, Oluwarotimi EA et al (2020) Insight into glycogen synthase kinase-3β inhibitory activity of phyto-constituents from Melissa officinalis: in silico studies. SilicoPharmacol 8:2

28. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY et al (2015) OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. J ChemTheoryComput. https://doi.org/10.1021/acs.jctc.5b00864

29. Schrödinger Release 201 8-4: LigPrep (2018) Schrödinger. LLC, New York, NY. 2018

30. Sherman W, Beard HS, Farid R (2006) Use of an Induced Fit Receptor Structure in Virtual Screening. ChemBiol Drug Des 67:83–84

31. Schrödinger Release 2020-2: Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY (2016) Prime, Schrödinger, LLC, New York, NY, 2020.

32. Schrödinger Release 2020–2: Prime, Schrödinger, LLC, New York, NY, (2020).

33. Schrödinger Release 201 8-4:QikProp (2018) Schrödinger, LLC, New York, NY, 2018.

34. Hongbin Y, Lixia S, Zhuang W, Weihua L, Guixia L, Yun T (2018) ADMETopt: A Web Server for ADMET Optimization in Drug Design via Scaffold Hopping. J ChemInf Model 58(10):2051–2056

35. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) GROMACS: fast, flexible and free. J ComputChem 26:1701–1718

36. Lumber JL (2020) African Countries Respond Quickly To Spread Of COVID-19. https://www.npr.org/sections/goatsandsoda/2020/03/21/818894991/africancountriesrespond-quickly-to-spread-of-covid-19

37. Lambel F (2020) Somali medics report rapid rise in deaths as Covid-19 fears grow. https://www.theguardian.com/world/2020/may/02/somali-medics-report-rapid-rise-in-deathsas-covid-19-fears-grow

38. Chang KO, Kim Y, Lovell S, Rathnayake AD, Groutas WC (2019) Antiviral drug discovery: Norovirus proteases and development of inhibitors. Viruses 11(2):1–14. https://doi.org/10.3390/v11020197

39. Mielech AM, Chen Y, Mesecar AD, Baker SC (2014) Nidovirus papain-like proteases: Multifunctional enzymes with protease, deubiquitinating and deISGylating activities. Virus Res. 194:184–190

40. Isabella AG, Camila SM, Laurent ED (2014) Receptor-Ligand molecular docking Biophys Rev 6(1):75–87

41. Park JY, Jeong HJ, Kim JH, Kim YM, Park SJ, Kim D, Park KH, Lee WS, Ryu YB (2012) Diarylheptanoids from Alnus japonica inhibit papain-like protease of severe acute respiratory syndrome coronavirus. Biol Pharm Bull 35:2036–2042

42. ParkY KJH, Kim YM, Jeong HJ, Kim DW, Park KH, Kwon HJ, Park SJ, Lee WS, Ryu YB (2014) Tanshinones as selective and slow-binding inhibitors for SARS-CoV cysteine proteases. Bioorg Med Chem 20:5928–5935

43. Cho JK, Curtis-Long MJ, Lee KH, Kim DW, Ryu HW, Yuk HJ, Park K (2014) Geranylated flavonoids displaying SARS-CoV papain-like protease inhibition from the fruits of Paulownia tomentosa. Bioorg Med Chem 21:3051–3057

44. Rampogu S, Baek A, Zeb A, Lee KW (2018) Exploration for novel inhibitors showing back-to-front approach against VEGFR-2 kinase domain (4AG8) employing molecular docking mechanism and molecular dynamics simulations. BMC cancer 18(1):264

45. Baez-Santos YM, Barraza SJ, Wilson MW, Agius MP, Mielech AM et al (2014) X-ray structural and biological evaluation of a series of potent and highly selective inhibitors of human coronavirus papain-like proteases. J Med Chem 57:2393–2412

46. Baez-Santos YM, Mielech AM, Deng X, Baker S, Mesecar AD (2014) Catalytic function and substrate specificity of the PLpro domain of nsp3 from the Middle East Respiratory Syndrome Coronavirus (MERS-CoV). J Virol. 88(21):12511–12527

47. Maffucci I, Hu X, Fumagalli V, Contini A (2018) An efficient implementation of the Nwat MMGBSA method to rescore docking results in medium-throughput virtual screenings. Front Chem 6:43

48. Sun H, Li Y, Shen M, Tian S, Xu L, Pan P et al (2014) Assessing the performance of the MM/PBSA and MM/GBSA methods. 5, improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. PhysChemChemPhys 16:22035–22045

49. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46(1–3):3–26

50. Zhang MQ, Wilkinson B (2007) Drug discovery beyond the "rule-offive." CurrOpinBiotechnol 18(6):478–488

51. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45(12):2615–2623

52. Congreve M, Carr R, Murray C, Jhoti H (2003) A rule of three for fragment-based lead discovery? Drug Discov Today 8(19):876–877

53. Köster H, Craan T, Brass S, Herhaus C, Zentgraf M, Neumann L, Heine A, Klebe GA, (2011) Small Nonrule of 3 Compatible Fragment Library Provides High Hit Rate of Endothiapepsin Crystal Structures with Various Fragment Chemotypes. J Med Chem 54(22):7784–7796

54. Baell JB, Congreve M, Leeson P, Abad-Zapatero C (2013) Ask the experts: past, present and future of the rule of five. Fut Med Chem 5(7):745–752