ORIGINAL ARTICLE

Structure-based virtual screening, in silico docking, ADME properties prediction and molecular dynamics studies for the identification of potential inhibitors against SARS-CoV-2 M^pro

Anbuselvam Mohan[1] · Nicole Rendine[2] · Mohammed Kassim Sudheer Mohammed[3] · Anbuselvam Jeeva[4] · Hai-Feng Ji[2] · Venkateswara Rao Talluri[5]

Abstract

COVID-19 is a viral pandemic caused by SARS-CoV-2. Due to its highly contagious nature, millions of people are getting affected worldwide knocking down the delicate global socio-economic equilibrium. According to the World Health Organization, COVID-19 has affected over 186 million people with a mortality of around 4 million as of July 09, 2021. Currently, there are few therapeutic options available for COVID-19 control. The rapid mutations in SARS-CoV-2 genome and development of new virulent strains with increased infection and mortality among COVID-19 patients, there is a great need to discover more potential drugs for SARS-CoV-2 on a priority basis. One of the key viral enzymes responsible for the replication and maturation of SARS-CoV-2 is M^pro protein. In the current study, structure-based virtual screening was used to identify four potential ligands against SARS-CoV-2 M^pro from a set of 8,722 ASINEX library compounds. These four compounds were evaluated using ADME filter to check their ADME profile and druggability, and all the four compounds were found to be within the current pharmacological acceptable range. They were individually docked to SARS-CoV-2 M^pro protein to assess their molecular interactions. Further, molecular dynamics (MD) simulations was carried out on protein–ligand complex using Desmond at 100 ns to explore their binding conformational stability. Based on RMSD, RMSF and hydrogen bond interactions, it was found that the stability of protein–ligand complex was maintained throughout the entire 100 ns simulations for all the four compounds. Some of the key ligand amino acid residues participated in stabilizing the protein–ligand interactions includes GLN 189, SER 10, GLU 166, ASN 142 with PHE 66 and TRP 132 of SARS-CoV-2 M^pro. Further optimization of these compounds could lead to promising drug candidates for SARS-CoV-2 M^pro target.

Keywords Covid-19 · SARS-CoV-2 M^pro protein · CADD · Virtual screening · ADME · Molecular dynamics simulation · Docking · RMSD · RMSF · Ligand · Computer aided drug design · Hydrogen bond · Optimization · Asinex library · Schrodinger software

✉ Venkateswara Rao Talluri
vrtalluri@gmail.com

1 Department of Biotechnology, Selvamm Arts and Science College (Autonomous), Namakkal, Tamil Nadu 637 003, India

2 Department of Chemistry, Drexel University, Philadelphia, PA 19104, USA

3 Department of Botany, Government Arts and Science College (Autonomous), Coimbatore, Tamil Nadu 641 018, India

4 Department of Animal Science, Bharathidasan University, Tiruchirappalli, Tamil Nadu 620 024, India

5 Prof.TNA Innovation Centre, Varsha Bioscience and Technology India Private Limited, Jiblakpally, Yadadri District, Hyderabad, Telangana 508 284, India

Introduction

COVID-19 is currently the most prevalent zoonotic infectious disease caused by SARS-CoV-2 virus. According to the World Health Organization (WHO), COVID-19 has affected over 186 million people and resulted in the death of 4 million people as of July 09, 2021[1, 2]. There are seven different strains of corona viruses that cause severe life-threatening infections in human's viz. HCoV-OC43, HCoV-HKU1, HCoV-229E, HCoV-NL63, MERS-CoV, SARS-CoV and SARS-CoV-2. These are responsible for the most life-threatening infectious diseases including those that cause complications with the respiratory and nervous systems, hepatic function and gastrointestinal system in humans [3].

SARS-CoV-2 virus is the most prominent viral infection that triggered the current global COVID-19 pandemic [4]. The COVID-19 infection is transmitted between humans primarily through respiratory droplets, fecal matter, and close contact with an infected person. The most common symptoms of COVID-19 are the onset of a fever, fatigue, dry cough, and diarrhea. Based on the symptoms, COVID-19 infections are classified into four groups: mild, moderate, severe, and critical. In a mild infection of COVID-19, the symptoms include headache, muscle pain, dry cough, mild fever, nasal congestion, sore throat and malaise. In the moderate COVID-19 infection, the symptoms include breathing problems and tachypnoea. In severe & critical COVID-19 infection, acute respiratory distress syndrome (ARDS), sepsis and the dysfunction of vital organs [5] occur.

SARS-CoV-2 binds to ACE2 receptors present on the surface of the human host cell. ACE2 receptor in humans is primarily found in the nasopharynx and interstitial epithelia, especially type II alveolar cells in the lung. The genomic RNA of SARS-CoV-2 directly attaches to the host's ribosomes and initiates viral genome transcription. This occurs through the proteolytic process and is carried out by the papain-like protease (PLpro) and corona virus main protease (M^{pro}). SARS-CoV-2's entire viral genome is 30 kb that encompass 29, 903 nucleotides and encodes 14 open reading frames (ORFs). Non-structural proteins are generated by M^{pro} and thus participate in viral replication. The 3′ end of the viral genome encodes the following structural proteins: spike(S)-glycoprotein, envelope protein (E), membrane protein (M), nucleoprotein (N), and nine putative accessory factors. These proteins are responsible for viral replication and viral functioning inside the host. Therefore, M^{pro} is a promising anti-viral drug target for COVID-19[6, 7, 8, 9, 10, 11 and 12].The M^{pro} of SARS-CoV-2 is a homodimer that contains 306 amino acids and three significant functional domains.

Domain I and II consists of an anti-parallel β-sheet structure, similar to chymotrypsin. Domain III consists of alpha-helices, with the third region connected to the second region through a long loop region containing 185–200 residues. The active site contains a Cys-His catalytic dyad located in a cleft between Domains I and II. Domains II and III are connected by a long loop of 185–200 amino acid residues.

The currently in practice vaccines or anti-viral therapeutics including darunavir, atazanavir, nelfinavir, indinavir, amprenavir, ritonavir, lopinavir, hydroxychloroquine, chloroquine, remdesivir and azithromycin are partially effective in controlling the severity of the disease and mortality. Therefore, there is an urgent need for the development of effective drugs against COVID-19 [13]. From the prehistoric times, medicinal plants played a significant role in drug development due to presence of rich biochemical's such as alkaloids, flavanoids, phenols, chalcones, coumarines, lignans, polyketides, alkanes, alkenes, alkynes, simple aromatics, peptides, terpenes and steroids with high therapeutic properties. According to World Health Organization (WHO), 80% of traditional medicines are prepared from medicinal plants [14].Computer-aided drug design plays a significant role in drug discovery and development. In the current study, structure-based virtual screening, ADME properties filtration and molecular dynamic simulations were carried out against M^{pro} of SARS-CoV-2 to identify the best possible drug/s that can stop viral multiplication at an early stage and reduce the mortality.

Materials and methods

Structure-based virtual screening

Computer-aided drug design (CADD) plays a crucial role in the preliminary phases of drug discovery and development, especially structure-based virtual screening (SBVS).This is a very popular in silico technique due to its speed, accuracy and low cost. SBVS is one of the most important techniques to contribute to the determination of lead compounds from a vast chemical library based on the 3-Dimensional structure of the drug target [15]. Molecular docking is one of the most prominent structure based drug design (SBDD) approaches. It accurately predicts the binding modes of a ligand within the binding pocket of a specific therapeutic drug target based on ligand interactions with the molecular entities found in the binding pocket.

In the current study, the 3-D coordinates of therapeutic target SARS-CoV-2M^{pro} (PDB ID: 6LU7) with co-crystallized inhibitor N3 was obtained from the protein data

bank [16]. Prior to docking, crystal structure coordinates of SARS-CoV-2 M^{pro} target proteins and compounds were prepared using protein preparation panel of Schrodinger software to bring them to their biological active state. During this process, the co-crystallized ligand and unwanted water molecules were removed, missing heavy atoms and formal charges were added, bond orders were adjusted, hydrogen atoms were added, proper charges were assigned to the target structure, and optimization and energy minimization were completed with the help of an OPLS3 force filed [17]. The prepared protein structure was subsequently taken to gird generation. In the present study, Schrodinger Glide version 5.5 was used for docking. The binding pocket residues of co-crystallized SARS-CoV-2M^{pro}(PDB ID: 6LU7) ASP 187, HIS 41, GLU 166, MET 49, TYR 54, CYS 145, GLN 189, SER 10, HIS 164 and ASN 142 shows extensive interaction with the co-crystal ligand N3[16]. Therefore, a grid box was generated by selecting the center of the co-crystal ligand (i.e., N3) as the centroid of the grid box (20 Å×20 Å×20 Å) by utilizing the receptor grid generation panel implanted within the Schrodinger suite.

For the ligand preparation, initially, 8,722 anti-viral compounds were downloaded from the ASINEX database in SDF format and reference co-crystal ligand (i.e., N3) 3-D structure was obtained in the PDB format [18, 19].LigPrep module of Schrodinger software suite was utilized for ligand preparation. During ligand preparation, the following criteria including (i) OPLS3 force field, (ii) generation of all possible ionization states at pH 7.0 ± 2.0, (iii) *desalt* option, (iv) generation of tautomers for all conformers, (v) generation of one low energy conformer per ligand was implemented. The prepared ligands were subsequently taken through a structure-based virtual screening process [20, 21] using Glide module of Schrodinger suite. The HTVS mode was employed for filtering the chemical entities suggested from the library of molecules, SP mode was applied for further screening, and XP mode was utilized to obtain more accurate docking calculations [22].

Assessment of ADME properties

Assessment of the pharmacological profile is one of the most important steps in the early stages of a drug discovery and development process. Schrodinger QikProp module was used to calculate ADME properties and screen the compounds. Some of the very important properties utilized for ligand screening includes partition coefficient (QP log P octanol/water), water solubility (QPlogS), percentage of human oral absorption, apparent MDCK permeability

(QPPMDCK), prediction of blood–brain barrier permeability (QPlogBB), molecular weight and hydrogen bond donors and acceptors [23, 24].

Molecular dynamics simulation

Understanding the molecular interactions between simple and macromolecular structures at the atomic level are a challenging task for a molecular biologist. Desmond software [25] was utilized for the analysis of molecular interactions at different time scales. From the virtual screening, four best protein–ligand complexes were selected based on their Glide score. SPC water model was used for salvation of the protein–ligand complexes with the orthorhombic water boundary box set to a box size of X = 10 Å, Y = 10 Å, and Z = 10 Å. Counter ions were added to the complex to neutralize the solvated system. The complex energy was minimized using an OPLS3e force field [26, 27]. In the current study, isothermal-isobaric (NPT) ensemble was employed. The molecular dynamics simulation run time was set at 100 ns, with a recording interval of 100 ps and energy of 1.2 ps. After the molecular dynamics run, RMSD and RMSF of the complexes were analyzed using the simulation interaction diagram panel embedded within Maestro.

Results and discussion

Selection of potential inhibitors for therapeutic targets of SARS-CoV-2M^{pro}

Theoretical and experimental information of therapeutic targets is crucial for structure-based virtual screening. In order to identify potential inhibitors for SARS-CoV-2M^{pro}, anti-viral compounds from the ASINEX database were docked using Glide software in the virtual screening study. Finally, potential inhibitors were selected based on their Glide score.

Molecular docking with SARS-CoV-2M^{pro}

M^{pro} is a crucial enzyme that participates in the transcription of the SARS-CoV-2 viral genome and, therefore, is an attractive drug target. To evaluate the effectiveness of the Glide docking protocols, a crystal structure of SARS-CoV-2 M^{pro} (PDB: 6LU7) from PDB was initially obtained and a co-crystal inhibitor N3 was extracted from the binding pocket of the SARS-CoV-2M^{pro} protein target. The N3 inhibitor was then redocked into the same position in SARS-CoV-2

M^pro using three different Glide docking protocols: HTVS, SP, and XP. The scores for each protocol were noted for the reference ligand N3 (-5.52 kcal/mol, -6.24 kcal/mol, and -6.63 kcal/mol, respectively). In the XP docking mode, the reference ligand N3 maintained hydrogen bond with following residues: ASP 197, HIS 41, and GLU 166 which has two hydrogen bond interactions. The residues had following bond lengths: 2.27 Å, 2.26 Å, 2.22 Å, and 1.90 Å, respectively. The amino acid residues MET 49, CYC 44, TYR 254, LEU 141, PHE 140, VAL 186, MET 165, PHE 181, and CYS 145 play a crucial role in hydrophobic interactions. The native crystal structure of SARS-CoV-2 and the docked conformers of SARS-CoV-2 were superimposed and the RMSD of the superimposed structures was calculated. Interestingly, the RMSD was found to be 0.166 Å. This result clearly indicates the dynamic nature of Glide docking protocol.

Initially, a total of 8,722 anti-viral compounds from ASINEX library were docked into the active site of SARS-CoV-2 to obtain Glide scores, which represent the binding affinities. The Glide score was used to rank the compounds. A high throughput virtual screening confirmed stronger binding of these compounds with M^pro in comparison to reference ligand N3, with a Glide score of -5.25 kcal/mol. The compounds were then allowed to dock with the binding pocket of SARS-CoV-2 in the standard precision mode yielding 700 compounds with a Glide score lower than -6.24 kcal/mol. These 700 compounds were allowed to dock using the extra precision mode in Glide leading to identification of four potential ligands exhibiting better binding than the reference compound N3 with binding free energy of -6.63 kcal/mol. The overall workflow was reported in Fig. 1.

The four potential, drug-like compounds exceed the pharmacokinetic factors, or requirements that are defined for drugs for human consumption and, therefore, qualify as potential drug-like molecules. The interacting residues for these lead compounds are shown in Table 1 and their docking pose are depicted in Fig. 2.

Glide score (Kcal/mol); Glide energy (Kcal/mol); No of hydrogen bond interaction; Interacting residues; Distance between the protein and ligand (Å).

ADME properties prediction

The ADME parameters of the four newly identified hit compounds were evaluated using QikProp software module. The drug-like properties of the four newly identified hit compounds were assessed based on Lipinski's rule of five and all the four hit compounds expressed drug-like behavior including molecular mass less than 500 Da, hydrogen bond donors less than 5, 10 hydrogen bond acceptors and less than 5 octanol-water partition coefficient(logP). The predicted cell permeability (QPPcaco2) ranged from 113 to 520 nm s^{-1}. The predicted value clearly indicates all the four compounds have excellent intestinal absorption capabilities. The predicted blood–brain barrier passage was within a safe range of -1.579 to -0.506 nm s^{-1}. This value indicates the compounds do not cause any adverse side effects to the brain. The human oral absorption percentage ranged from 65 to 93%. The predicted aqueous solubility (QPlogS) ranged from -5.254 to -3.320. The predicted binding to human serum albumin (QPksha) was fairly low, -0.132 to -0.405. The predicted MDCK value was also below the average range, 46 to 512. The number of rotatable bonds ranged from 4000 to 7000. All the pharmacokinetic parameters fit well within the satisfactory range. The previously stated values clearly indicate the hit compounds possess good solubility, membrane permeability and efficiently reach the target site to produce the desired biological activity. The results of the ADME properties are reported in Table 2 (Fig. 3).

Molecular weight (<500 Da), Hydrogen bond donors (<5). Hydrogen bond acceptors (<10).

Predicted octanol/water partition co-efficient log p (acceptable range: 22.0 to 6.5). Predicted aqueous solubility; S in mol/L (acceptable range: 26.5 to 0.5). Predicted Caco-2 cell permeability in nm/s (acceptable range: 25 is poor and 0.500 is great). Prediction of binding to human serum albumin (acceptable range: 21.5 to 1.5). QP log BB for brain/blood (-3.0 to 1.2). Predicted percentage of human oral absorption is poor (25%).

Fig. 1 The overall workflow of virtual screening
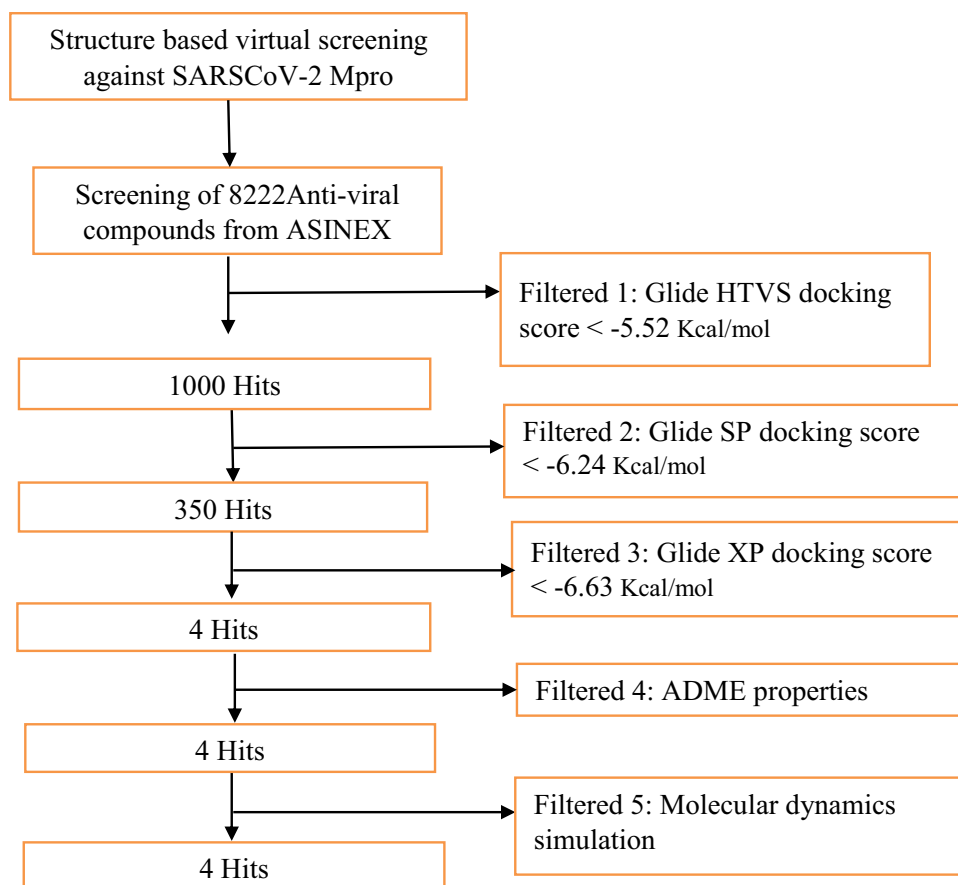
Table 1 Glide extra-precision (XP) results for the four potential compounds

| S. No	Compound Name	Glide score	Glide Energy	No. of hydrogen bonds	Interaction residues	Distance (Å)
1	[3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamicacid tert-butyl ester	−8.21	−64.35	5	GLN 189 (2) HIS 164 LEU 141 ASN 142	2.38, 1.84 2.04 2.04 1.94
2	1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylicacid (4-fluoro-phenyl)-amide	−7.89	−68.94	4	GLN 189 PHE 140 GLU166 HIS 164	2.16 2.10 2.11 2.26
3	1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide	−7.87	−76.70	4	GLN 189 PHE 140 GLU166 HIS 164	2.13 2.26 1.93 2.30
4	1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide	−7.45	−62.67	3	GLY 143 LEU 141 GLN 189	2.07 1.93 1.56
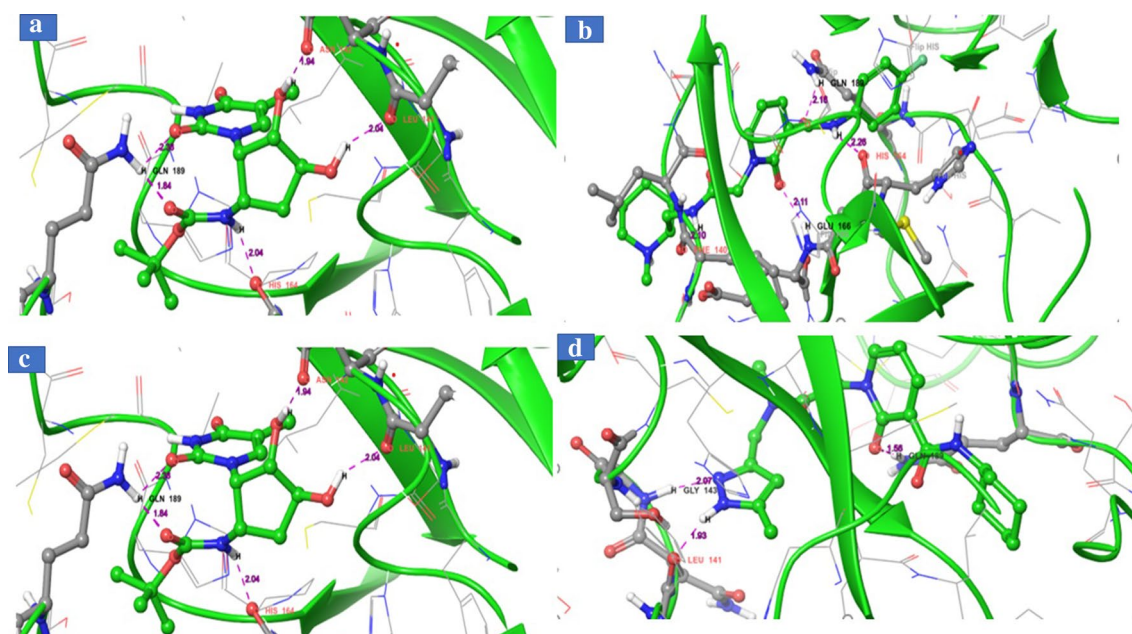
Fig. 2: 3-D interaction representation of the top four molecules in the active site of the Mpro. **a** [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)cyclopentyl]-carbamicacid tert-butyl ester: **b**. 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2dihydro-pyridine-3-carboxylicacid (4-fluoro-phenyl)-amide. **c**. 1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide. **d**. 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide

Docking study

Binding analysis of [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester

The docking pose of [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester inside the binding cavity of the SARS-CoV-2Mpro was analyzed. The Glide score and Glide energy were found to be −8.21 kcal/mol and -64.35 kcal/mol, respectively. The docking pose revealed that the [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester formed five hydrogen bonds with the SARS-CoV-2 Mpro. The side-chain hydrogen atom of GLN 189 of SARS-CoV-2 Mpro was involved in the formation of a hydrogen bond with the oxygen of the [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)cyclopentyl] carbamic acid

tert-butyl ester. This hydrogen bond has a bond length of 1.84 Å, 2.38 Å. The hydrogen of [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl) cyclopentyl]-carbamic acid tert-butyl ester interacts with the oxygen of LEU 141 of SARS-CoV-2 Mpro target. This hydrogen bond has a bond length of 2.04 Å. The hydrogen of [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester interacts with the oxygen of residue HIS 164 of SARS-CoV-2 Mpro to form a hydrogen bond with a bond length of 2.04 Å. The hydrogen of the [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester participates in a hydrogen bond with the side chain oxygen. A series of key amino acids located in the active site of SARS-CoV-2 Mpro, such as LEU 27, PRO 52, TYR 54, PHE 181, LEU 141, MET 165, MET 49 and CYC 145 accounts for the hydrophobic interactions in the protein–ligand complex.

Table 2 Assessment of drug-like properties of the four hit compounds

Compound Name	MW	HBD	HBA	QPLogPo/w	Caco-2	Khasa	QPLogB/B	MDCK	Logp S	Rb	% Oral absorption
[3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamicacid tert-butyl ester	341.363	4.000	9.400	0.283	113	−0.405	−1.579	46	−3.320	4.000	65
1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylicacid (4-fluoro-phenyl)-amide	400.452	1.000	9.000	1.942	171	−0.339	−0.506	173	−3.600	6.000	78
1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide	423.446	1.000	8.500	3.050	482	−0.132	−1.138	512	−5.254	7.000	93
1-[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl]-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide	385.465	1.000	8.500	2.297	520	−0.234	−1.070	333	−4.465	6.000	89

Binding mode of 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide

The docking pose of the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide inside the binding cavity of the SARS-CoV-2M^{pro} was examined. The docking pose revealed four hydrogen bond interactions between the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide and the binding pocket region of the SARS-CoV-2 M^{pro}. The hydrogen atom of the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide formed a hydrogen bond with an oxygen of the backbone residue PHE 140. This hydrogen bond had a bond length of 2.10 Å. The side-chain hydrogen atom of the polar residue GLN 189 formed a hydrogen bond with the oxygen atom of the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide. This hydrogen bond had a bond length of 2.16 Å. The backbone hydrogen atom of the negatively charged residue GLU 166 formed a hydrogen bond with the oxygen atom of the1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide. This hydrogen bond had a bond length of 2.11 Å. The hydrogen atom of the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide interacted with the oxygen atom of the backbone residue HIS 164 to form a hydrogen bond with a bond length of 2.26 Å. Similarly, a series of key amino acids located in the active site, such as LEU 141, MET 165, MET 49, CYC 145, PRO 52, TYR 54 and CYS 44, accounts for the hydrophobic interactions. The Glide score and Glide energy were found to be −7.89 kcal/mol and −68.94 kcal/mol, respectively.

Binding mode of—({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide

The docking pose of -({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide within the binding cavity of the SARS-CoV-2M^{pro} was examined. The Glide score and Glide energy were obtained −7.87 kcal/mol and -76.70 kcal/mol, respectively. The docking pose displayed only four hydrogen bond interactions formed between the-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide and the

binding cavity of the SARS-CoV-2 M^{pro}. The following residues participated in hydrogen bond interactions: PHE 140, GLU 166, HIS 164, and GLN 189. The bond lengths of the preceding hydrogen bonds were as follows: 2.26 Å, 1.93 Å, 2.30 Å, and 2.13 Å, respectively. Similarly, a series of key amino acids located in the active site, such as PRO 52, LEU 141, MET 165, MET 49, CYC 145, TYR 54 and CYS 44, accounts for hydrophobic interactions.

Binding mode of 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide

The docking pose of the 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide within the binding cavity region of the SARS-CoV-2 M^{pro} was examined. The Glide score and Glide energy were obtained to be -7.45 kcal/mol and −62.67 kcal/mol, respectively. The binding pose exhibited three hydrogen bond interactions between the 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide and the SARS-CoV-2 M^{pro}. The side-chain hydrogen atom of the polar residue GLN 189 interacts with an oxygen atom of the

1-{[Methyl-(5-me+thyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide to form a hydrogen bond with a bond length of 1.56 Å. A hydrogen atom of the 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylicacid cyclohexylamide formed a hydrogen bond with an oxygen atom of the backbone residue LEU 141. This hydrogen bond with a bond length of 1.93 Å. A hydrogen atom of the backbone residue GLY 143 interacts with a nitrogen atom of 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine 3-carboxylic acid cyclohexylamide to produce a bond length of 2.07 Å. Furthermore, amino acid residues VAL 186, MET 49, MET 165, LEU 167 and PHE 140 contribute to the hydrophobic interactions between the ligand and the main protease of SARS-CoV-2.

Molecular dynamics simulation analysis

MD analysis of SARS-CoV-2 M^{pro}− 3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester

A MD simulation for binding stability of the M^{pro}-3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,
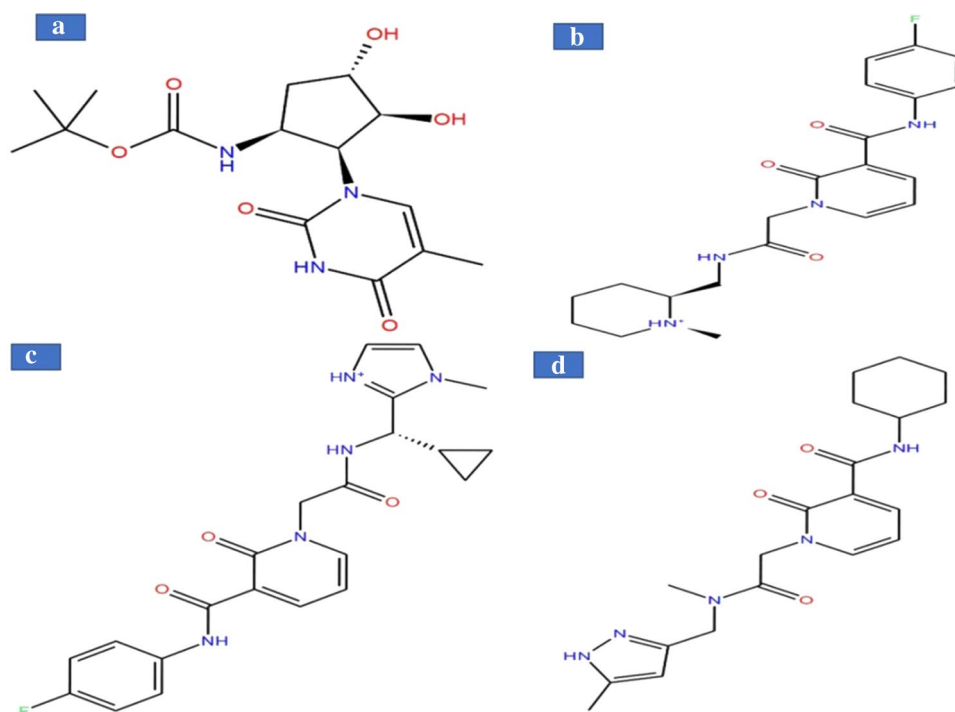

Fig. 3 Chemical structures of the four hit molecules. **a** [3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl) cyclopentyl]-carbamicacid tert-butyl ester: **b** 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2dihydro-pyridine-3 carboxylic acid (4-fluoro-phenyl)-amide. **c** 1-({[Cyclopropyl- (1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide. **d** 1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide
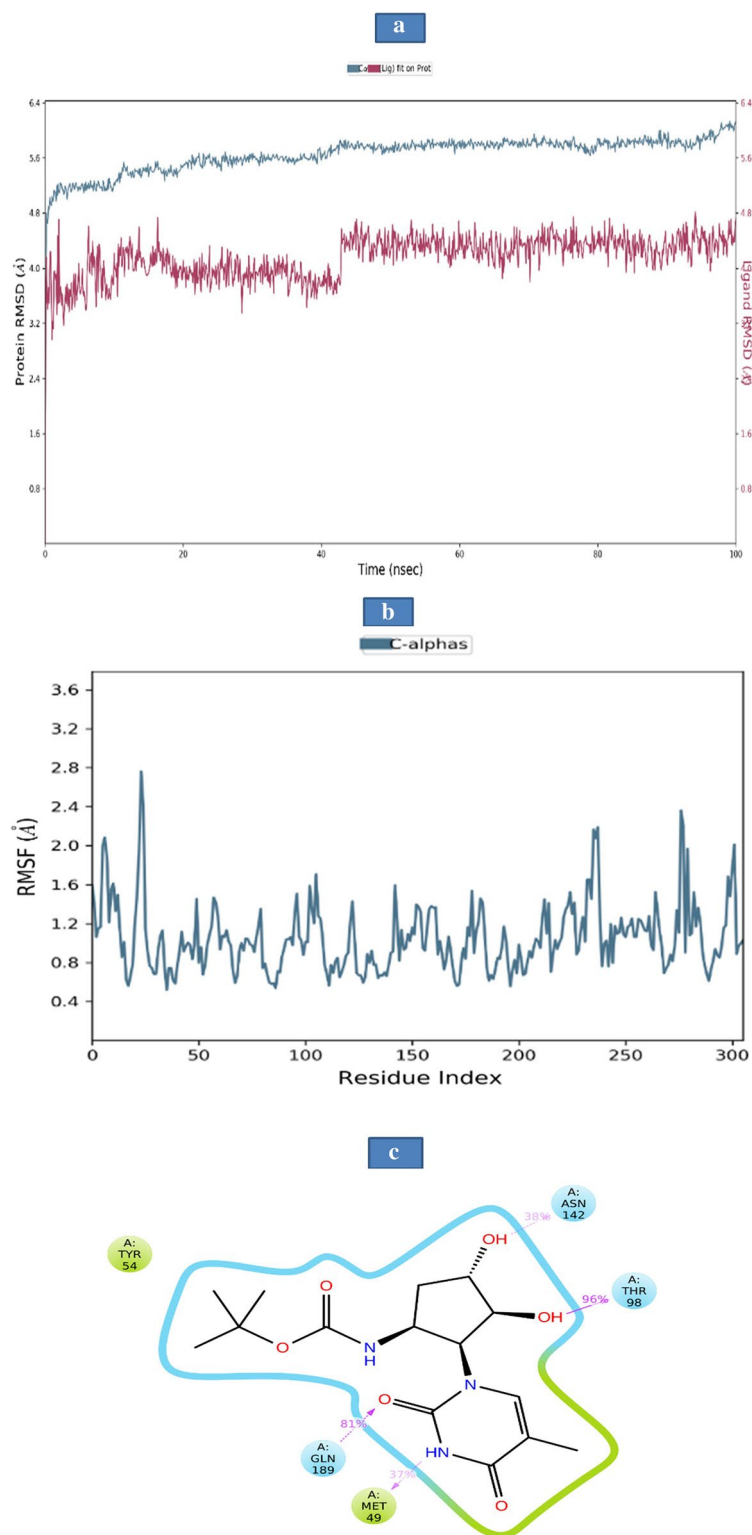
Fig. 4 a The RMSD plot of M^pro-3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester complex during 100-ns simulations. **b** The RMSF plot of M^pro during the 100 ns simulations. **c** The 2d diagram of M^pro-3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester contacts during 100 ns simulations
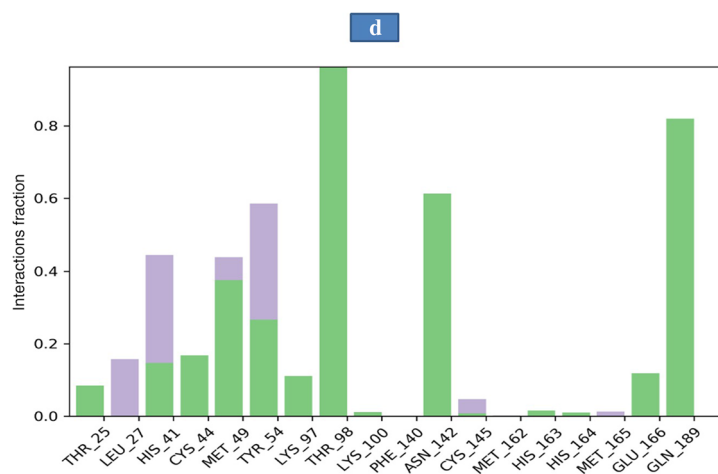
1-yl)-cyclopentyl]-carbamic acid tert-butyl ester complex at end of the 100 ns simulations. **d** The bar diagrams of M^pro-3,4-Dihydroxy-2-(5-ethyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester contacts during 100 ns simulations

Fig. 4 (continued)

4-dihydro-2H-pyrimidin-1-yl)-cyclopentyl]-carbamic acid tert-butyl ester was performed and the results were analyzed. The ligand RMSDs fluctuated with a maximum value of 4.7 Å during the first 3 ns followed by more stable interactions for the rest of the simulation. The RMSD of the protein backbone graph for complex 1 was obtained from the MD experiment at a certain time scale. It showed that the complex RMSD initially had slight fluctuations, ranging from 4.7 Å to 5.8 Å, during the first 45 ns followed by sustained stability and a value of 6.1 Å for the remainder of the simulation.

The stability of the protein–ligand complex during dynamics analysis was listed as the RMSD values of the protein backbone as shown in Fig. 4a. The root-mean-square fluctuation (RMSF) values for each residue in the protein backbone are illustrated graphically in Fig. 4b. In these plots, the peaks indicate residues that fluctuates the most over the entire simulation. The fluctuations started at 1.5 Å and continued up to 2.7 Å and most of the fluctuations were observed between residues 5 to 25. The amino acid residues ASN 142, THR 98, GLN 189 and MET 49 formed hydrogen bond interactions at the end of the dynamics simulation resulting in the increased stability of the protein–ligand complex as depicted in Fig. 4c.

Simulation results indicate that the protein–ligand complex exhibited good hydrogen bonding between the residues (green). The residue THR 98 exhibited the most hydrogen bonding throughout the simulation period followed by GLN 189(0.9) and ASN 142 (0.6). The residue THR 54 exhibited hydrogen bonding up to 30% of the simulation time followed by hydrophobic interactions is reported in Fig. 4d. The simulation indicates presence of hydrogen bonds and hydrophobic interactions throughout the simulation period. This indicates stability of the complex.

MD analysis of 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide

The binding stability of the 1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide on M^{pro} was examined for 100 ns under different time scales. Initially, the RMSD of the ligand reached a maximum of 6.3 Å for 80 ns. This was followed by steady interactions at 5.6 Å until the end of the simulation. It was noted that the protein backbone was initially 6.6 Å to 7 Å for 5 ns, followed by stable fluctuations up to 7.3 Å for the remaining simulation period. A slight deviation observed on the protein backbone indicates some changes in the binding stability of compound 2 within the binding region of the M^{pro}. This is reported in Fig. 5a. The RMSF value for each residue in the target protein backbone was analyzed and is depicted in Fig. 5b. The protein backbone fluctuations ranged from 0.9 Å to 1.9 Å. Most of the fluctuations were detected between the residues of 140 to 290. It was observed that the following residues were involved in the formation of hydrogen bonds with the binding site region of the M^{pro} during the docking simulation: GLN 189, PHE 140, GLU166 and HIS 164. Even though, the key residues, CYS 145, THR 26, GLU 166 and PRO 168 participated in the formation of hydrogen bonds with the target protein at the end of the simulation, they also enhanced the stability of the protein–ligand complex as reported in Fig. 5c. The strongest interactions between the protein-hit compound 2 complexes were due to the presence of steady hydrogen bonding and hydrophobic interactions over the duration
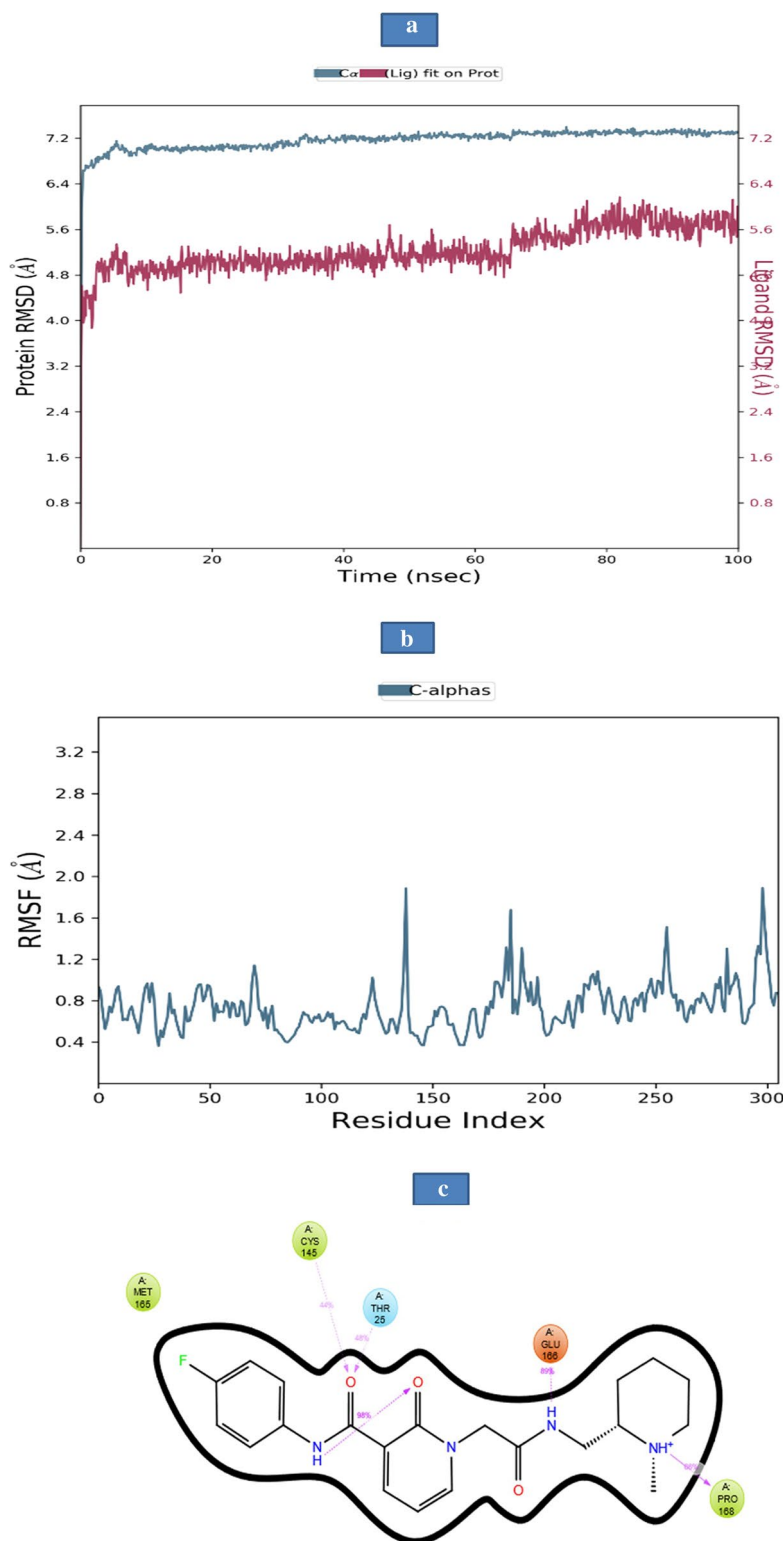
Fig. 5 **a** The RMSD plot of M^pro^-1-{[(1-Methyl-piperidin-2-ylmethyl) -carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide complex during 100-ns simulations. **b** The RMSF plot of M^pro^ during the 100 ns simulations. **c** The 2d diagram of M^pro^-1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide complex at end of the 100 ns simulations. **d** The bar diagrams of M^pro^-1-{[(1-Methyl-piperidin-2-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide contacts during 100 ns simulations
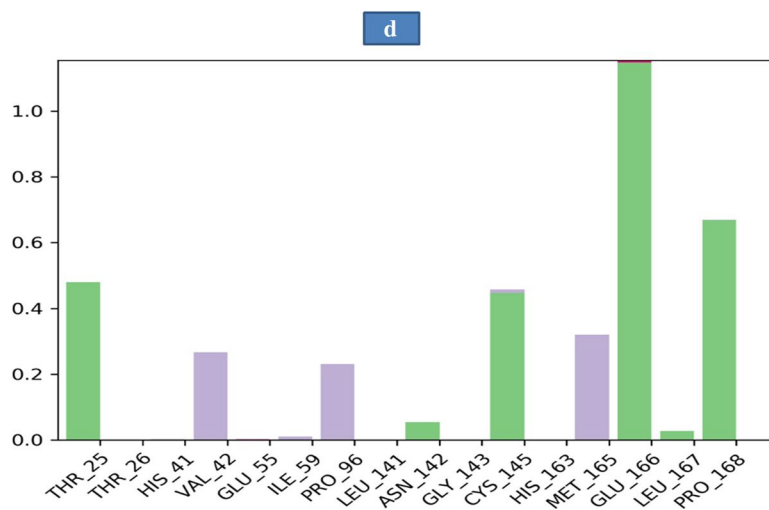
Fig. 5 (continued)

of the simulation. This is reported in Fig. 5d. The residue GLN 166 exhibited a hydrogen bonding network with the M^{pro} throughout the simulation followed by PRO 168 (0.7) and THR 25 (0.47). The residue CYS 145 showed hydrogen bonding up to 42% of the simulation time followed by hydrophobic interactions. The simulation indicates presence of hydrogen bonds and hydrophobic interactions throughout the simulation period. This also indicates the increase in stability of the protein–ligand complex.

MD analysis of M^{pro}- 1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1, 2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide

The binding stability of the 1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide within the active site of the M^{pro} was analyzed. The ligand RMSD fluctuated up to 8.1 Å for the first 3 ns but gradually became more stable prior to 42 ns. From 42 to 100 ns, the stability of the ligand RMSD increased based on the lowest RMSD reaching 8.8 Å. Fluctuations in the RMSD protein backbone were recorded. These ranged from 7.3 Å to 10.8 Å for up to 100 ns and are reported in Fig. 6a. The RMSF value for each residue in the protein backbone is illustrated graphically in Fig. 6b. In these plots, the peaks indicate areas of the

residues that fluctuates the most over the entire simulation. The fluctuations started around 0.6 Å to 1.6 Å and most of the fluctuations were observed between residues 5 to 230. It was observed that GLN 189, PHE 140, GLU166, and HIS164 residues participated in hydrogen bonding with the active site region of the M^{pro} during the docking simulation. At the end of the molecular dynamics simulation, it was observed that PHE 8, SER 10, ASN 142 and GLU 166 residues also formed hydrogen bonds to improve the stability of the protein–ligand complex. This is reported in Fig. 6c. The protein–ligand interaction histogram of the hit compound 3 is reported in Fig. 6d. The protein–ligand complex throughout the simulation shows good hydrogen bonding between the residues (green). The residue GLN 166 exhibited the most hydrogen bonding throughout the simulation period followed by SER 10 (0.40) and ASN 142 (0.35). The residue PHE 8 participated in hydrophobic interactions. The simulation indicates the presence of hydrogen bonds and hydrophobic interactions throughout the simulation period. This demonstrates the stability of the protein–ligand complex.

MD analysis of M^{pro}-1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid cyclohexylamide

The binding stability of the MD analysis of M^{pro}-1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-
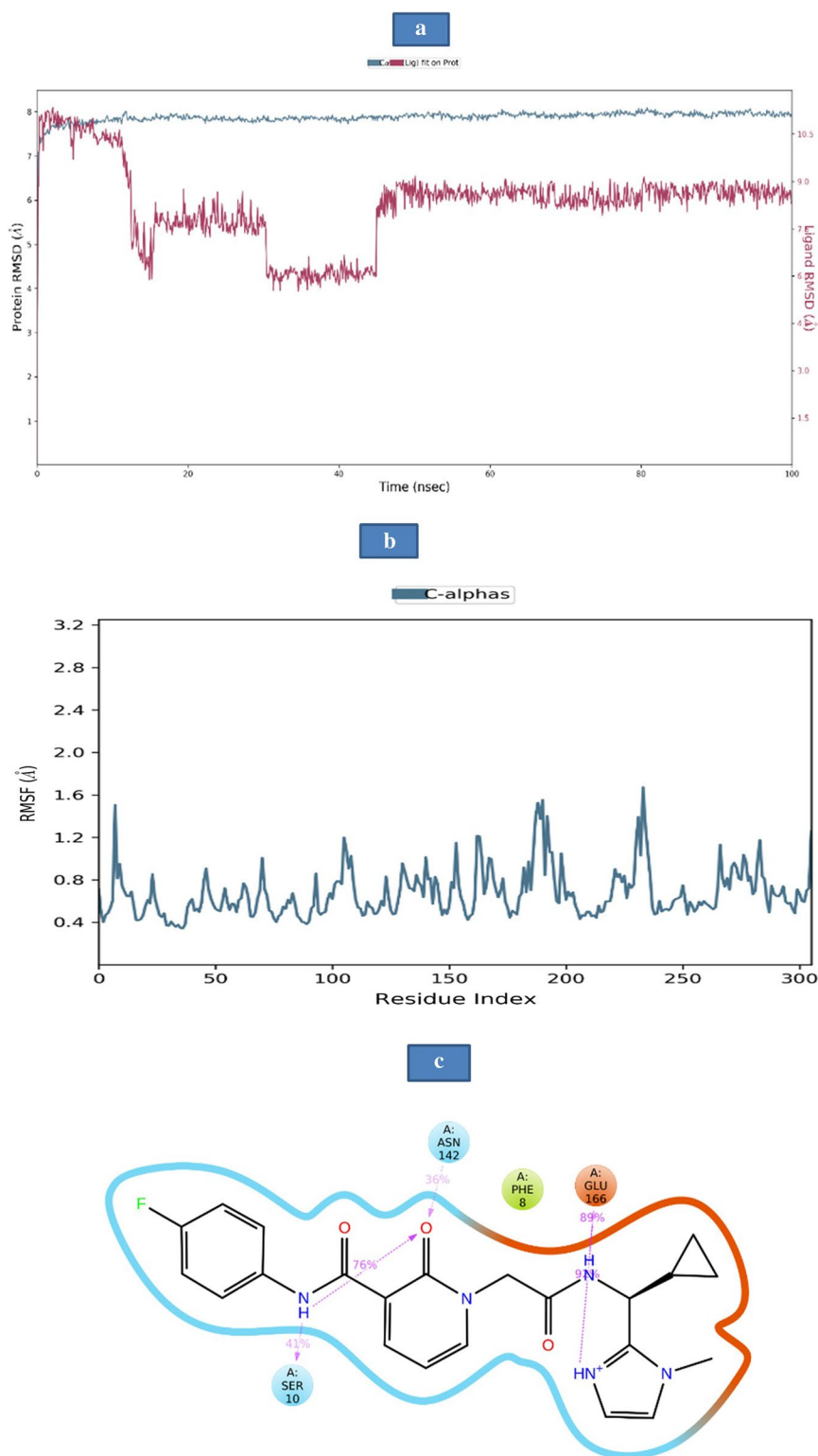
Fig. 6 a The RMSD plot of M^{pro}-1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide complex during 100-ns simulations. b The RMSF plot of M^{pro} during the 100nssimulations.cThe2ddiagramofM^{pro}-1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyri-dine-3-carboxylic acid (4-fluoro-phenyl)-amide complex at end of the 100nssimulations.d).ThebardiagramsofM^{pro}-1-({[Cyclopropyl-(1-methyl-1H-imidazol-2-yl)-methyl]-carbamoyl}-methyl)-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid (4-fluoro-phenyl)-amide contacts during 100 ns simulations
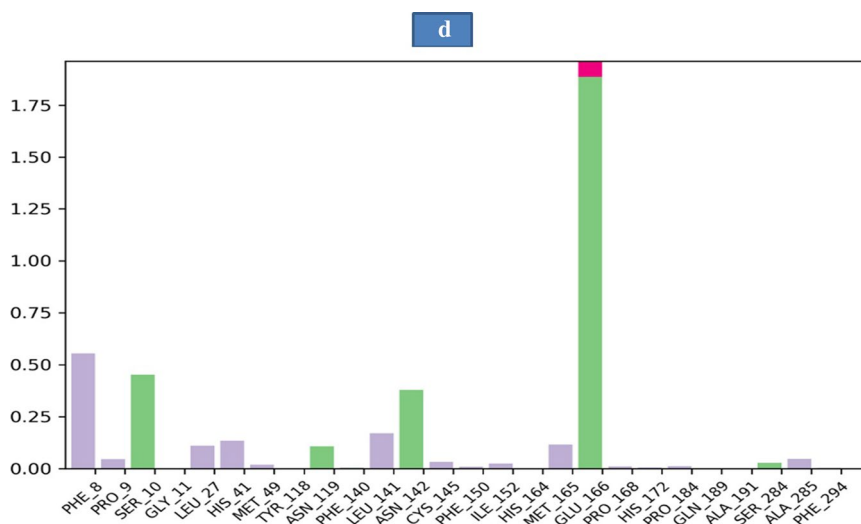
Fig. 6 (continued)

methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylicacidcy-clohexylamide were analyzed. The RMSD of the ligand initially recorded fluctuates up to a maximum of 5.9 Å until 42 ns. This was followed by stable interactions at 4 Å until the end of the simulation. The protein backbone initially had recorded fluctuations ranging from 5.8 Å to 9.6 Å up to 25 ns. This was followed by stable fluctuations up to 10.6 Å through the end of the simulation. Large RMSD fluctuations represent a large binding conformation variation and movement of the compound inside the binding site. This is shown for this experiment in Fig. 7a. The RMSF value for each residue in the protein backbone is illustrated graphically in Fig. 7b. The fluctuations ranged from 2 Å to 4.8 Å and most of the fluctuations occurred between residues 170 to 280. It was observed that GLU 166 and GLN 189 residues participate in hydrogen bonding with the active site region of the M^pro during the docking simulation. At the end of the molecular dynamics simulation, it was observed that GLN 189, SER 10 and ASN 142 also formed hydrogen bonds which can be attributed to the increase in stability of the protein–ligand complex as reported in Fig. 7c. The protein–ligand complex throughout the simulation showed good hydrogen bonding between the residues (green). The residue PRO 9 exhibited the strongest hydrogen bonding throughout the simulation period followed by GLN 189(0.65) and ASN 142 (0.6). The residue CYS 145 initially exhibited hydrogen bonding

up to 50% of the simulation time followed by hydrophobic interactions. The simulation indicates presence of hydrogen bonds and hydrophobic interactions throughout the simulation period. This demonstrates the stability of the complex as reported in Fig. 7d.

Conclusion

COVID-19 is one of the most threatening pandemic diseases. The control of this viral multiplication in human beings is a challenging task for Pharma investigators. SARS-CoV-2 M^pro protein is one of the very important drug targets available to inhibit the viral multiplication. In the current study, over eight thousand anti-viral compounds from ASINEX database were screened using structure-based virtual screening. Compounds passing through ADME filter were individually docked to the M^pro protein target. Molecular dynamics studies were performed for individually for all the four Protein–ligand complexes at 100 ns. The binding stability was evaluated and the structural conformational changes of the protein–ligand complex were evaluated to understand the efficiency of probable drug-receptor complexes in a biological system. From these in silico studies, it was concluded that all the four hit compounds could be further studied to understand their suitability for therapeutic applications.
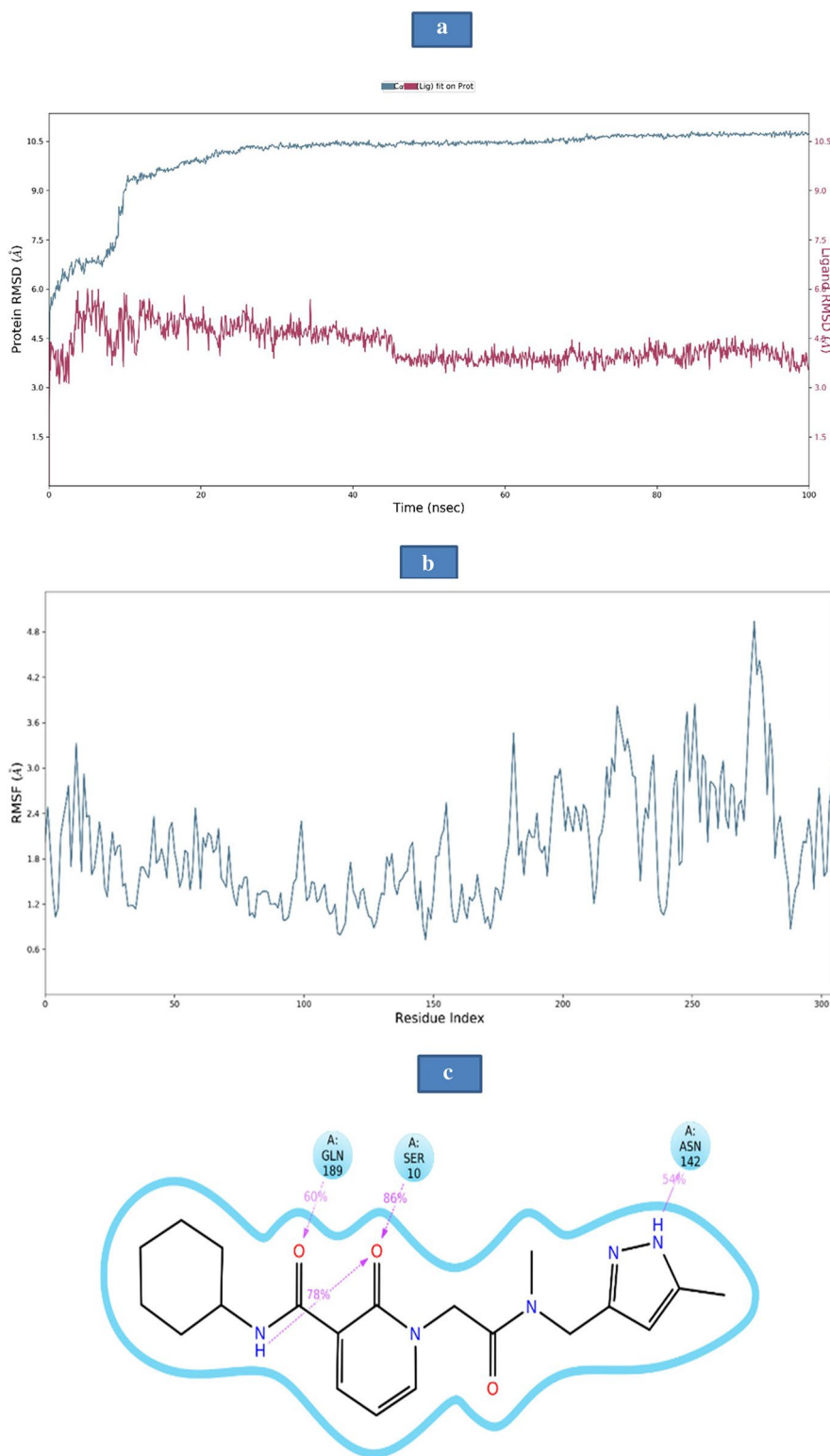
Fig. 7 a The RMSD plot of M^pro-1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid complex during 100-ns simulations. **b** The RMSF plot of M^pro during the 100 ns simulations. **c** The 2d diagram of M^pro-1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acidcomplex at end of the 100 ns simulations. **d** The bar diagrams of M^pro-1-{[Methyl-(5-methyl-1H-pyrazol-3-ylmethyl)-carbamoyl]-methyl}-2-oxo-1,2-dihydro-pyridine-3-carboxylic acid contacts during 100 ns simulations
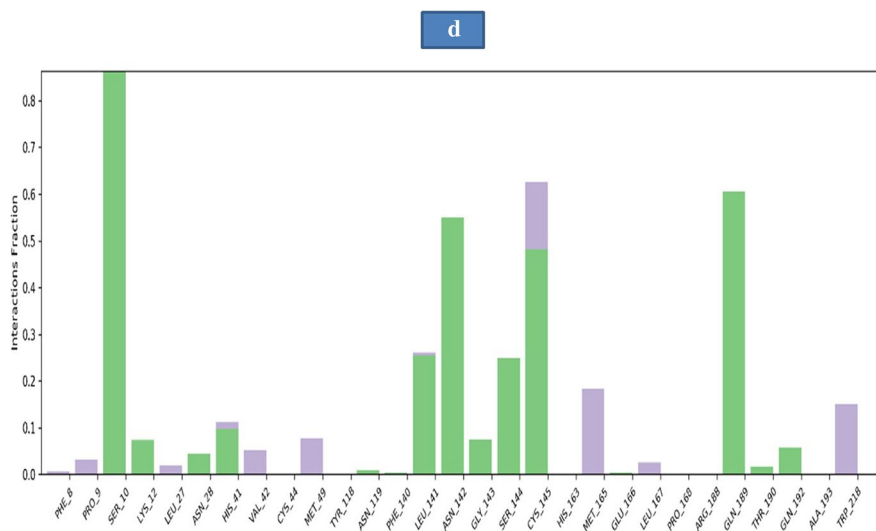
Fig. 7 (continued)

References

1. https://www.who.int/emergencies/diseases/novel-coronavirus-2019?

2. Gentile D, Patamia V, Scala A, Sciortino MT, Piperno A, Rescifina A (2020) Putative inhibitors of SARS-CoV-2 main protease from a library of marine natural products: a virtual screening and molecular modeling study. Mar Drugs 18(4):225

3. Zhou M, Zhang X, Jieming Qu (2020) Corona virus disease 2019 (COVID-19): a clinical update. Front Med 4(2):126–135. https://doi.org/10.1007/s11684-020-0767-8

4. Anthony RF, Stanley Perlman C (2015) Corona viruses: an overview of their replication and pathogenesis. Methods Mol Biol 1282:1–23

5. Hassan SA, Sheikh FN, Jamal S, Ezeh JK, Akhtar A (2020) Corona virus (COVID-19): a review of clinical features, diagnosis, and treatment. Cureus. https://doi.org/10.7759/cureus.7355

6. Lianga J, Pitsilloua E, Karagiannis C, Darmawan KK, Ng K, Hung A, Karagiannis T (2020) Interaction of the prototypical α-ketoamide inhibitor with the SARS-CoV-2 main protease active site in silico: Molecular dynamic simulations highlight the stability of the ligand-protein complex. ComputBiol Chem 87:107292. https://doi.org/10.1016/j.compbiolchem.2020.107292

7. Nisha Muralidharana R, Sakthivela DV, Michael Gromiha M (2020) Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-CoV-2 protease against COVID-19. J BiomolStruct Dyn. https://doi.org/10.1080/07391102.2020.1752802

8. Dai W, Zhang B, Jiang X-M, HaixiaSu JL, Zhao Y, XiongXie ZhenmingJin, JingjingPeng FL, Li C, Li Y, Bai F, Wang H, Cheng Xi, Cen X, Shulei Hu, Yang X, Wang J, Liu X, Xiao G, Jiang H, ZiheRao L-K, YechunXu HY, Liu H (2020) Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science 368(6497):1331–1335. https://doi.org/10.1126/science.abb4489

9. Yadav R, Imran M, Dhamija P, Suchal K, Handu S (2020) Virtual screening and dynamics of potential inhibitors targeting RNA binding domain of nucleocapsid phosphoprotein from SARS-CoV-2. J Biomol Struct Dyn 2020(22):1–16. https://doi.org/10.1080/07391102.2020.1778536

10. Oso BJ, Adeoye AO, Olaoye IF (2020) Pharmacoinformatics and hypothetical studies on allicin, curcumin, and gingerol as potential candidates against COVID-19-associated proteases. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1813630

11. Kumar A, Choudhir G, Shukla SK, Sharma M, Tyagi P, Bhushan A, Rathore M (2020) Identification of phytochemical inhibitors against main protease of COVID-19 using molecular modelling approaches. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1772112

12. Chikhale RV, Gurav SS, Patil RB, Sinha SK, Prasad SK, Shakya A, Shrivastava SK, Gurav NS, Prasad RS (2020) Sars-cov-2 host entry and replication inhibitors from Indian ginseng: an *in-silico* approach. JBiomolStruct Dyn. https://doi.org/10.1080/07391102.2020.1778539

13. Maia EHB, Assis LC, Alves T, de Oliveira A, da Silva M, Taranto AG (2020) Structure-based virtual screening: from classical to artificial intelligence. Front Chem 28(8):343. https://doi.org/10.3389/fchem.2020.00343

14. Kumar Anuj, Kumar Rohit, Sharma Mansi, Upendra Kumar MNV, Gajula Prasad, Singh Krishna Pal (2018) Uttarakhand Medicinal Plants Database (UMPDB): a platform for exploring genomic, chemical, and traditional knowledge. Data 3(1):7. https://doi.org/10.3390/data3010007

15. Ferreira LG, dos Santos RN, Oliva G, Andricopulo AD (2015) Molecular docking and structure-based drug design strategies. Molecules 20(7):13384–13421. https://doi.org/10.3390/molecules200713384

16. Jin Z, Xiaoyu Du, Yechun Xu, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, YinkaiDuan JY, Wang L, Yang K, Liu F, Jiang R, Yang X, You T, Liu X, Yang X, Bai F, Liu H, Liu X, Guddat LW, Wenqing Xu, Xiao G, Qin C, Shi Z, Jiang H, Rao Z,

Yang H (2020) Structure of M^pro from SARS-CoV-2 and discovery of its inhibitors. Nature 582:289–293

17. Protein preparation Wizard Maestro (2018) New York: Schrodinger LLC

18. http://www.asinex.com

19. Dhanachandra Singh KH, Karthikeyan M, Kirubakaran P, Nagamani S (2011) Pharmacophore filtering and 3D-QSAR in the discovery of new JAK2 inhibitors. J Mol Graph Model. https://doi.org/10.1016/j.jmgm.2011.07.004

20. Ligprep, Version 2.3, Schrodinger, LLC, New York, NY

21. Glide, Version 5.5 (2009) Schrodinger, LLC, New York, NY

22. Ma D-L, Chan D-H, Leung C-H (2011) Molecular docking for virtual screening of natural product databases. Chem Sci 2:1656–1665

23. Dong J, Wang N-N, Yao Z-J, Zhang L, Cheng Y, Ouyang D, Ai-Ping Lu, Cao D-S (2018) ADMET lab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. J Cheminform 10(1):29. https://doi.org/10.1186/s13321-018-0283-x

24. Teli MK, Rajanikant GK (2012) Identifications of novel potential HIF-prolyl hydroxylase inhibitors by in silico screening. Mol Divers 16(1):193–202. https://doi.org/10.1007/s11030-011-9338-x

25. Veeramachaneni GK, Kranthi Raj K, Chalasani LM, Bondili JS, Talluri VR (2015) High-throughput virtual screening with e-pharmacophore and molecular simulations study in the designing of pancreatic lipase inhibitors. Drug Des Devel Ther 9:4397–4412. https://doi.org/10.2147/DDDT.S84052

26. Salmas RE, Unlu A, Bektaş M, Yurtsever M, Mestanoglu M, Durdagi S (2017) Virtual screening of small molecules databases for discovery of novel PARP-1 inhibitors: combination of in silico and in vitro studies. J Biomol Struct Dyn 35(9):1899–1915. https://doi.org/10.1080/07391102.2016.1199328

27. Vijayakumar S, Manogar P, Prabhu S, Singh RASK (2018) Novel ligand-based docking; molecular dynamic simulations; and absorption, distribution, metabolism, and excretion approach to analyzing potential Acetylcholinesterase inhibitors for Alzheimer's disease. J Pharm Anal. 8(6):413–420