Identification of potential SARS-CoV-2 M^pro inhibitors integrating molecular docking and water thermodynamics

M. Elizabeth Sobhia, Ketan Ghosh, Srikanth Sivangula, Siva Kumar and Harmanpreet Singh

National Institute of Pharmaceutical Education and Research (NIPER), SAS Nagar, Punjab, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

The COVID-19 pandemic is an ongoing global health emergency caused by a newly discovered coronavirus SARS-CoV-2. The entire scientific community across the globe is working diligently to tackle this unprecedented challenge. *In silico* studies have played a crucial role in the current situation by expediting the process of identification of novel potential chemotypes targeting the viral receptors. In this study, we have made efforts to identify molecules that can potentially inhibit the SARS-CoV-2 main protease (M^pro) using the high-resolution crystal structure of SARS-CoV-2 M^pro. The SARS-CoV-2 M^pro has a large flexible binding pocket that can accommodate various chemically diverse ligands but a complete occupation of the binding cavity is necessary for efficient inhibition and stability. We augmented glide three-tier molecular docking protocol with water thermodynamics to screen molecules obtained from three different compound libraries. The diverse hits obtained through docking studies were scored against generated WaterMap to enrich the quality of results. Five molecules were selected from each compound library on the basis of scores and protein-ligand complementarity. Further MD simulations on the proposed molecules affirm the stability of these molecules in the complex. MM-GBSA results and intermolecular hydrogen bond analysis also confirm the thermodynamic stability of proposed molecules. This study also presumably steers the structure determination of many ligand-main protease complexes using x-ray diffraction methods.

Introduction

Newly emerging pathogens are becoming worldwide threats to public health in recent times. In December 2019, a cluster of patients with 'pneumonia of unknown origin' was reported in Wuhan city of Hubei province, China. The etiology of this disease was later associated with a virus of the coronavirus family. The virus was later termed as Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) by experts of the International Committee on Taxonomy of Viruses (ICTV) due to its apparent similarity to SARS-CoV from 2002. The World Health Organization (WHO) labelled this infectious disease as Coronavirus Disease-2019 (COVID-19) and declared it as a global health emergency. Though the earlier outbreaks of severe acute respiratory syndrome coronavirus (SARS-CoV) in China and middle east respiratory syndrome coronavirus (MERS-CoV) in the Middle east, posed a significant threat, with the COVID-19 numbers increasing exponentially, the global public health has been under serious stress (Paules et al., 2020; Wu et al., 2020). As of 6 September 2020, WHO reports more than 27 million confirmed cases and 900,000 deaths worldwide (WHO, 2020).

An enormous amount of research is underway to develop a vaccine or identify drugs against this virus (Abd El-Aziz & Stockand, 2020; Al-Kassmy et al., 2020; Elfiky, 2020; Huggins, 2020). Considerable work on the development of drugs for other coronaviruses such as SARS-CoV and MERS-CoV has enabled rapid identification of drug targets for SARS-CoV-2 (Anand et al., 2003; Ghosh et al., 2005; 2007; Kumar et al., 2016; 2017;). The RNA genome of SARS-CoV-2 encodes for various structural, non-structural and accessory proteins. The non-structural proteins like 3-chymotrypsin-like protease (3CL^pro), papain-like protease, helicase, and RNA-dependent RNA polymerase as well as structural proteins such as spike glycoprotein are attractive targets for developing antiviral drugs (Li & De Clercq, 2020). The main reading frame of RNA, ORF 1ab encodes for two overlapping polyproteins pp1a and pp1ab which undergo proteolysis through two viral proteases, the papain-like protease along with 3CL^pro also referred to as main protease (M^pro) to produce 16 non-structural proteins. The remaining portion of the RNA codes for accessory proteins and structural proteins viz., spike (S) protein, membrane (M) protein, envelop (E) protein, nucleocapsid (N) protein, hemagglutinin esterase (HE) glycoprotein (Ullrich & Nitsche, 2020; Ziebuhr et al., 2000). The cleavage of viral polypeptides pp1a and pp1ab is an important step for viral replication. Therefore, inhibiting viral proteases can be an attractive option to inhibit viral replication. Such an approach has been successfully applied earlier to develop

several antiviral drugs that are used to treat Hepatitis C (HCV) and HIV infections (Agbowuro et al., 2018). Since the advent of COVID-19 antivirals including ritonavir and lopinavir were suggested for the treatment of the disease however they were found ineffective or of little benefit in the treatment of disease (Li et al., 2020; Smolders et al., 2020). Several other investigations are underway testing the efficacy of broad-spectrum antivirals like interferon-α/γ (Felgenhauer et al., 2020) and other HCV or HIV protease inhibitors like danoprevir (Chen et al., 2020) and darunavir (Chen et al., 2020; Nicolini et al., 2020) to treat the infection. The structure of M[pro] is rather conserved among coronaviruses. Moreover, till now there is no known human homolog for viral M[pro] which makes it an ideal drug target (Ullrich & Nitsche, 2020; Zhang et al., 2020).

Intensive efforts are being made to identify options to treat SARS-CoV-2 infection focusing on enzymes involved in various viral pathways such as TMPRSS2 (Huggins, 2020), RNA polymerase (Ivashchenko et al. 2020; Yin et al., 2020), Spike protein (Lei et al., 2020; Liu et al., 2020), M[pro] (Fu et al., 2020; Ma et al., 2020; Zhang et al., 2020) and helicase (Shu et al., 2020; Yuen et al., 2020). Previous studies based on SARS-CoV and MERS-CoV recognised α-ketoamides as broad-spectrum inhibitors of coronaviruses (Zhang et al., 2020). This had provided leads to the scientific community to identify new inhibitors for SARS-CoV-2. Several covalent inhibitors have already been identified for SARS-CoV-2 M[pro] and their co-crystallized structures successfully solved (Fu et al., 2020; Jin et al., 2020; 2020; Ma et al., 2020; Zhang et al., 2020). Two covalent protease inhibitors of ketoamide category boceprevir and GC376 exhibited strong *in vitro* activity against SARS-CoV-2 M[pro] (Fu et al., 2020). An organoselenium compound ebselen has recently emerged as a promising drug lead candidate in cell-based assays which targets this crucial enzyme (Jin et al., 2020; Menendez et al., 2020). In this study, we focused on viral M[pro] to identify molecules with potential to bind the active site of the SARS-CoV-2 M[pro] structure (PDB ID - 5R82). Despite similar studies have published earlier, we followed a three-tiered docking protocol using commercially available ligand libraries and WM/MM-GBSA calculations to meet our objective. Our study focuses on small molecule candidates of natural or synthetic origin which can accelerate drug discovery efforts to combat the COVID-19 pandemic.

Structural information on M[pro] of SARS-CoV-2

Table S1 contains the details of all the experimentally solved crystal structures of M[pro]. The resolution of these structures having 306 residues ranges from 1.31 Å to 2.20 Å. Amino acid sequence alignment SARS-CoV-2 with SARS-CoV M[pro] reveals a high degree of sequence identity (∼96 percent). Superimposition of their crystal structures also revealed high structural similarity with conserved active site region (Ullrich & Nitsche, 2020).

This protein structure forms a homodimer, crystallized with a small fragment molecule bound to the substrate-binding site. Each monomer forms three subdomains with substrate binding site harbouring in a cleft between domain II and domain III. The 3D-structural analysis of 5R82 reveals that the binding cavity is larger than the space occupied by the co-crystallized ligand. The substrate-binding pocket is further divided into five sub-pockets S1, S2, S3, S4 and S1' which accommodate different amino acid residues of the substrate polypeptide. The two monomers of M[pro] are connected through non-covalent linkages where SER1 of each monomer interacts with GLU166 of the other monomer which stabilizes the S1 sub-pocket of the binding site in M[pro] protein. The proteolytic activity of M[pro] is believed to follow a multi-step mechanism using CYS145 and HIS41 residues (catalytic dyad) with help of a water molecule, present in S1' sub-pocket of the binding cavity (Jin et al., 2020; Ullrich & Nitsche, 2020). A ligand that can form interactions with these catalytic residues along with hydrophilic and hydrophobic interactions with residues of other sub-pockets can act as a potent SARS-CoV-2 M[pro] inhibitor. Figure 1 shows the biological assembly of SARS-CoV-2 homodimer with substrate-binding pocket highlighting the various sub-pockets and catalytic residues.

The residues surrounding the binding cavity are shown in Table 1. Among these residues THR26, HIE41, SER46, PHE140, GLY143, CYS145, GLU166 and GLN 189 are important for main protease inhibition.

Materials and methods

3D-Databases preparation

Three databases namely Enamine (Shivanyuk et al., 2007), Specs (www.specs.net), Natural product (Sterling & Irwin, 2015) were refined by eliminating duplicates and adding missing hydrogen atoms. All the ligand libraries were prepared using Ligprep module of Schrödinger suite (Release 2019-1: LigPrep, Schrödinger, LLC, New York, NY, 2020). The 3D structures were generated using the new improved Optimized Potential for Liquid Simulations-3e (OPLS-3e) force field (Harder et al., 2016). Ligands were desalted and possible tautomers and stereoisomers were generated while preserving the specified chiralities. Epik (Shelley et al., 2007) utility was used to generate possible ionisation states at pH 7.0 ± 2.0.

Protein preparation

We selected PDB 5R82 from the list of available structures for our study as it has the best resolution (1.31 Å) amongst all the available structures. The protein structure file was prepared using the Protein Preparation Wizard of Schrödinger's Maestro (Release 2019-1: Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2020). This involved addition and optimization of hydrogen bonds, disulphide bonds creation, atomic clashes removal, addition of formal charges to the hetero groups and then optimizing at neutral pH. Finally, the structure was minimized using the OPLS-3e force field. The crystal water molecules from the crystal structure near the catalytic site and substrate binding site were retained for the molecular docking studies.
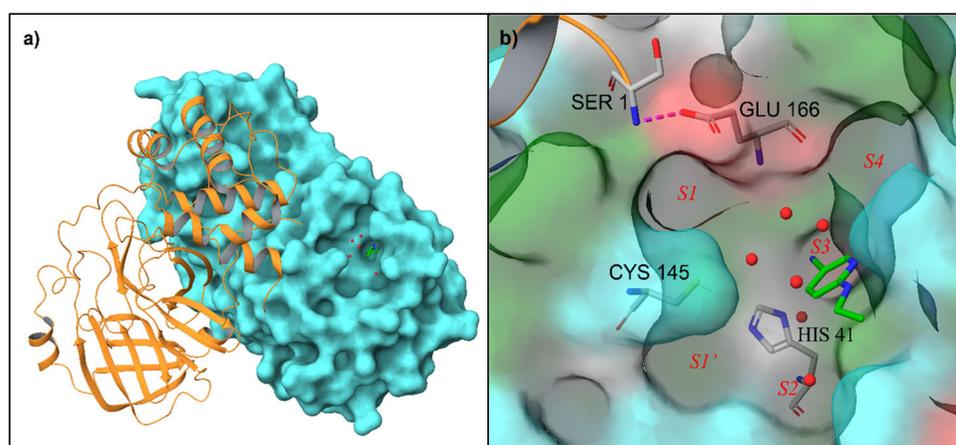
Figure 1. (a) Biological assembly of Mpro dimer showing cartoon representation of one monomer and surface of the other monomer highlighting the binding pocket with green ligand and crystal waters shown as red spheres (b) Enlarged view of binding pocket highlighting the key residues and crystal waters in form of red spheres. The sub-pockets S4-S1 in the binding cavity are indicated(clock-wise).

Table 1. Residues making up the cavity for binding.

Residue Type	Residue no. and 3 letter names
Polar residues	Gln19, Thr21, Thr24, Thr25, Thr26, Asn28, His41, Thr45, Ser46, Asn119, Asn142, Gly143, Ser144, His163, His164, Gln189, Thr190, Gln192
Hydrophobic residues	Val20, Leu27, Pro39, Val42, Ile43, Cys44, Met49, Leu50, Pro52, Tyr54, Tyr118, Phe140, Leu141, Cys145
Negatively charged residues	Glu47, Asp48, Glu166, Asp187
Positively charged residues	Arg40, Arg188

Grid-Generation

The receptor grid was generated at the centroid of the bound ligand in the crystal structure of SARS-CoV-2 M^{pro}. The grid box was extended up to 15 Å for the inner box and 26 Å for the outer box covering the entire binding site cavity completely.

Molecular docking

We used a series of hierarchical filters to search for probable locations of the ligand in the binding site of a receptor. Docking studies were performed using the Glide module of Schrödinger suite (Release 2019-1: Glide, Schrödinger, LLC, New York, NY, 2020). In a three-tiered docking strategy, the prepared ligands from the 3D databases were docked in three stages, following the docking protocols, which are High Throughput Virtual Screening (HTVS), Standard Precision (SP) and Extra Precision (XP). The top ~5 percent of the screened ligands from the first stage of HTVS docking were passed on to the second stage of SP docking. 500 screened molecules from SP docking were then docked using a more accurate and computationally intensive XP mode. The HTVS and SP docking use the same scoring function, however, HTVS reduces the number of intermediate conformers that reduce the thoroughness of the final torsional refinement. On the other hand, XP mode uses a more extensive sampling than SP and employs an improved scoring function that rewards or penalises certain interactions (Friesner et al., 2006). No functional group constraints or torsional constraints were applied during the docking process. Post docking minimization was performed for the ligands and all other parameters were set to default when following the three-tiered docking strategy.

WaterMap calculations

WaterMap calculations were used to assess the thermodynamic parameters of waters in the binding site. These parameters can be used for ligand design by examining how ligands displace/replace these hydration sites. It is based on a small molecular dynamics simulation of explicit solvent followed by clustering and generation thermodynamic properties of hydration sites. It classifies water hydration sites into two categories, namely stable water and unstable waters based on the thermodynamic properties like entropy (-TΔS), enthalpy (-ΔH), free energy (ΔG) (Abel et al., 2011; Riniker et al., 2012).

WaterMap calculations for the prepared protein structure were performed using default simulation parameters (TIP4P solvent model at 300 K, 1 atmospheric pressure and 2 ns of simulation time) and treating existing waters as a part of explicit solvent (Release 2019-1: WaterMap, Schrödinger, LLC, New York, NY, 2020). The binding site was defined using the coordinates of the co-crystallized ligand. The truncate protein option was left unchecked for this study. The generated hydration sites were examined according to their enthalpy, entropy and free energy values.

The WM/MM scoring approach was employed to score the ligands which uses a combination of MMGBSA and WaterMap data to assess the best binding pose. The scoring was performed using VSGB solvation model and defining residues within the vicinity of the ligand as non-flexible.

Molecular dynamics simulations

Two complexes of protein with ligands ZINC31167921 and ZINC67912395 were considered for Molecular dynamics (MD)

Flowchart of the study
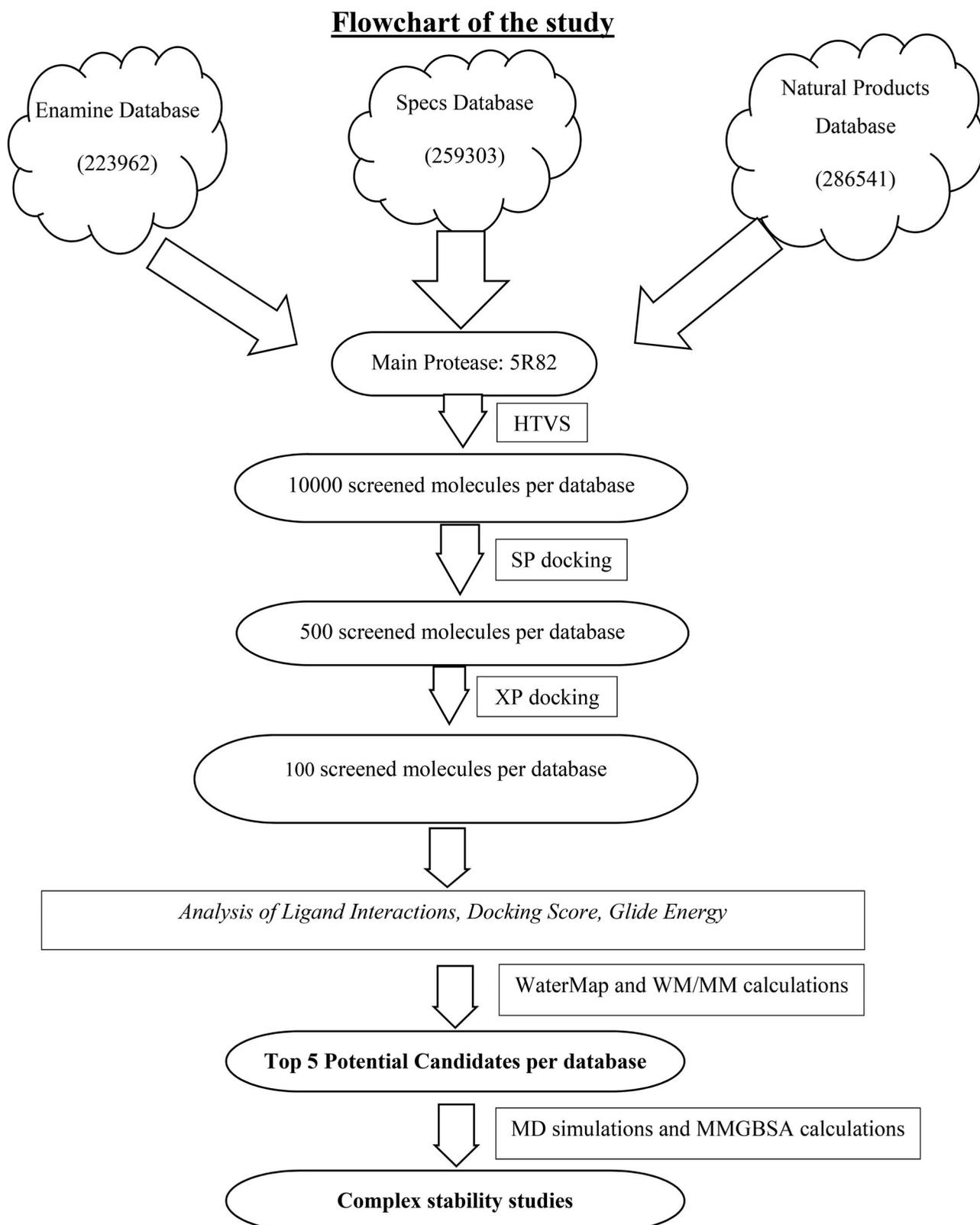
Figure 2. Flowchart representation of the study performed.

simulations to understand the dynamic changes in the complex and binding interactions. These two complexes were submitted for MD simulations of ∼100ns using AMBER18 (Case et al., 2005). The protein topologies were prepared using ff14SB force field (Maier et al., 2015) parameters with the help of AMBER LEaP module while antechamber module and GAFF force field with am1bcc charges (Wang et al., 2004) were utilized to generate the parameters for ligands. The system was solvated in TIP3P water molecules and neutralized by adding the necessary amount of Cl⁻ ions in electrostatically preferred positions as the protein is positively charged at neutral pH (Harrach & Drossel, 2014).

The electrostatic interactions were estimated using Particle Mesh Ewald (PME) method with non-bonded cutoff kept at 10 Å. The solvated complexes were minimized three times for 1000, 500 and 250000 steps through a short steepest descent minimization for 50% steps followed by conjugate gradient minimization for the remaining steps. Heating was performed using NVT ensemble for 50 ps where the protein-ligand complex was restrained with force constant of 2 kcal/mol/Å^2 and gradually heated from 0 to 300 K. Three density equilibration steps were carried out; two steps of 50 ps each with the restrained weight of 2 kcal/mol/Å^2 and 1 kcal/mol/Å^2 respectively followed by one step of 100 ps unrestrained density equilibration. The system was equilibrated to free simulation for 1 ns at 300 K temperature and 1 atm pressure. Finally, a long-range simulation for 100 ns production run was performed at 300 K temperature and 1 atm pressure. Coordinate trajectories were recorded after every 10 ns for complete MD run. MD was performed to calculate a single trajectory using pmemd program of AMBER18 running on NVIDIA Tesla K20Xm GPU workstation (Salomon-Ferrer et al., 2013).

The outline of the strategy followed for this study is mentioned in a flowchart (Figure 2).
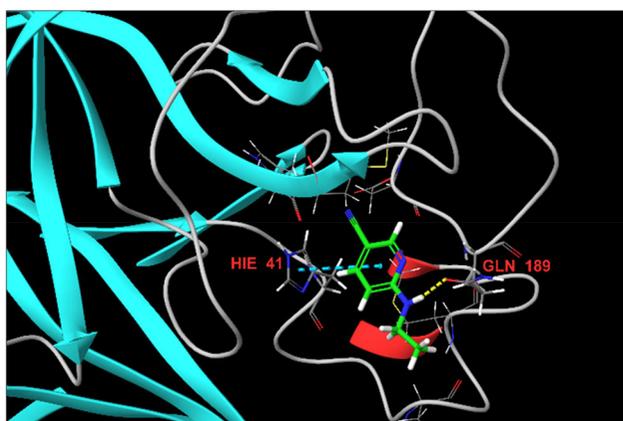
Figure 3. Interactions of co-crystalized ligand in 5R82 with labelled residues in the vicinity.

Results and discussion

Main protease crystal structure (PDB:5R82) offers the opportunity to employ a structure-based drug design strategy in identifying novel COVID-19 inhibitors. In the experimental structure, ligand RVZ forms a hydrogen bond with GLN189 and a weak pi-pi stacking interaction with HIS41 as shown in Figure 3. It is clear that the binding site is surrounded by various loop regions that owe to the flexible nature of the binding cavity. Cysteine and serine proteases usually possess three catalytic residues while M^{pro} has only two catalytic residues in its active site (Ullrich & Nitsche, 2020). It can be presumed that a water near the catalytic HIS41 residue plays the role of the third residue by making a strong hydrogen bond with it. A water molecule near the HIS41-CYS145 catalytic dyad was preserved for docking studies. Another crystal water was preserved near the catalytic site stabilized by three hydrogen bonds with HIS41, HIS164 and ASP187 residues. The molecules screened using XP docking were subsequently rescored using WM/MM-GBSA scoring to find the relative binding affinities towards SARS-CoV-2 M^{pro}. The combined WaterMap MM-GBSA scoring protocol generally provides a better correlation with experimental activity than either of them when used alone (Abel et al., 2011).

Molecular docking and database screening

Approximately 7.7 lakh molecules from three databases were screened by HTVS followed by SP and XP docking. The top docked molecules from each database were chosen for further analysis. The molecules were filtered according to the docking scores and analysed for hydrogen bond interaction with any of the residues in the binding site region especially HIS41, CYS145, GLU166 and GLN189. The data of five selected molecules from each database is presented in Table 2. These results include parameters like docking scores, XP GScore, MMGBSA ΔG and WM/MM ΔG values. Average ligand binding free energies for selected molecules range from around −45 kcal/mol to −60 kcal/mol.

The 3D interaction diagram of the top 5 molecules from each database are shown in Supplementary Figure S1. The

Table 2. Binding energy parameters of top 5 screened database molecules complexed with enzyme.

Sr. no.	Title	Docking Score*	XP Gscore*	MMGBSA ΔG bind*	WaterMap ΔG bind*	WM/MM ΔG bind*
			Natural products Database			
1	ZINC38139713	−11.458	−11.458	−59.351	−25.324	−23.082
2	ZINC31165691	−9.936	−9.936	−53.785	−23.596	−23.072
3	ZINC31167921	−11.598	−11.598	−49.451	−28.468	−23.013
4	ZINC31165101	−6.836	−6.836	−60.325	−21.950	−22.253
5	ZINC38143727	−10.197	−10.197	−49.690	−24.955	−21.075
			Enamine Database			
6	Z3040398264	−6.296	−6.477	−56.295	−27.642	−22.702
7	Z2706637055	−6.167	−6.204	−51.977	−32.733	−22.167
8	Z2229246315	−6.164	−6.171	−58.873	−24.021	−22.029
9	Z3061991917	−6.306	−6.413	−49.288	−26.015	−21.125
10	Z2063273911	−7.979	−7.986	−45.759	−18.565	−18.420
			Specs Database			
11	ZINC49781425	−10.380	−10.380	−56.958	−25.528	−26.646
12	ZINC67912395	−10.135	−10.135	−61.343	−29.317	−26.414
13	ZINC67911229	−11.057	−11.057	−58.837	−17.338	−23.609
14	ZINC15675289	−7.357	−7.357	−59.439	−20.246	−22.150
15	ZINC44305556	−10.684	−10.684	−44.280	−19.291	−21.914

*Sign indicates kcal/mol.

Table 3. Predicted ADME/Tox properties of the selected molecules.

Title	PlogPo/w	PlogHERG	PPCaco	PlogBB	PlogKhsa	Rule of Five	Rule of Three	DILI	AMES test
			Natural products Database						
ZINC38139713	−2.121	−5.416	1.309	−4.338	−1.188	3	2	No	No
ZINC31165691	−1.867	−5.528	1.429	−4.253	−1.006	2	2	No	Yes
ZINC31167921	0.804	−5.031	65.528	−2.689	−0.76	3	1	No	No
ZINC31165101	−0.877	−3.278	0.158	−0.993	−0.745	2	1	Yes	No
ZINC38143727	−2.267	−5.061	1.702	−4.211	−1.256	3	2	No	No
			Enamine Database						
Z3040398264	1.826	−2.535	494.741	−0.749	−0.656	0	0	Yes	Yes
Z2706637055	0.34	−1.803	79.251	−1.525	−0.874	0	0	Yes	No
Z2229246315	1.586	−6.605	51.077	−1.382	−0.119	0	1	Yes	Yes
Z3061991917	3.218	−3.996	1477.416	−0.497	0.067	0	0	No	No
Z2063273911	−0.313	−2.725	6.233	−1.443	−0.833	0	1	Yes	Yes
			Specs Database						
ZINC49781425	−3.276	−5.758	0.455	−5.324	−1.593	3	2	No	No
ZINC67912395	−1.311	−3.556	0.042	−5.703	−1.313	3	2	No	No
ZINC67911229	−1.857	−5.011	2.708	−3.864	−1.148	3	2	No	No
ZINC15675289	1.338	−3.899	215.845	−1.186	−0.605	0	0	Yes	No
ZINC44305556	−2.274	−5.505	1.551	−4.266	−1.258	3	2	No	No

PlogPo/w: Predicted octanol/water partition coefficient. (-2.0 – 6.5).
PlogHERG: Predicted IC_{50} value for blockage of HERG K^+ channels. (concern below -5).
PPCaco: Predicted apparent Caco-2 cell permeability in nm/sec. (25 poor, >500 great).
PlogBB: Predicted brain/blood partition coefficient. (−3.0–1.2).
PlogKhsa: Prediction of binding to human serum albumin. (−1.5–1.5).
Rule of five: Number of violations of Lipinski's rule of five. (Maximum is 4).
Rule of three: Number of violations of Jorgensen's rule of three. (Maximum is 3).
DILI: Drug Induced Liver Injury prediction.
AMES: AMES toxicity prediction for carcinogenicity.

selection of the molecules was based on the fact that these molecules possess required bonded and non-bonded interactions with residues involved in catalytic mechanism and engulfing the S1, S1', S3 and S4 sub-pockets of 5R82. Most of the selected molecules occupy the catalytic S1 and S1' subsites showing interactions with atleast one of the residues of catalytic dyad and hydrogen bonds with the other residues of the binding site. The docking analysis revealed that a common pi-stacking interaction usually forms by the alignment of aromatic rings of the ligands and HIE41 side chain. Ligands Z2229246315 and ZINC15675289 formed both hydrogen bond and pi-stack interaction with HIE41. The selected ligands ZINC38139713, ZINC31165691, Z3061991917, ZINC49781425 and ZINC44305556 form multiple hydrogen bonds/salt bridge with GLU166 (residue involved in enzyme dimerization along with SER1 of the other monomer) that can interfere with the biological assembly of the enzyme and inhibit its activity. Ligands ZINC38139713, ZINC49781425 and ZINC67912395 from the Specs and natural products database showed a significantly stronger affinity towards the M^pro indicated by multiple hydrogen bonds and hydrophobic interactions that may help ligand to stabilize into the binding pocket. Also, a good MMGBSA ΔG and WM/MM ΔG scores for these three ligands affirmed their likelihood to act as M^pro inhibitors. Molecules like Z3040398264, Z2063273911, ZINC67911229, ZINC5675289 and ZINC44305556 contain an electrophilic warhead group (carbon atom of the carbonyl group) in close proximity to the catalytic thiol-group of CYS145 which may behave as peptidomimetic substrate to form strong covalent interaction with the enzyme. A general comparison of these ligands to active in vitro molecules boceprevir and GC376 displayed similar interaction patterns. The two crystal water molecules conserved for docking simulation didn't show any major water bridged interaction except in ligand Z2063273911 which forms a water bridged

interaction with HIE164. Although water molecules are involved in the catalytic mechanism of the protein, they may not have a significant role in ligand binding interactions. A list of the type of ligand interactions with the corresponding residues involved is provided in Table 4.

WaterMap calculations and WM/MM scoring

The WaterMap calculations were used to compute the free energy (ΔG) of the explicit waters around the binding site of protein. The hydration sites generated for the crystal structure 5R82 are displayed in Figure 4. Hydration sites are coloured according to their free energy values ranging from green to red. The most stable ones are represented using green while the most unstable or energetically unfavourable ones are shaded in red. There are particularly three hydration sites with high ΔG values of 5.47, 4.96 and 3.67 kcal/mol located in hydrophobic site S3 and site S4 of the binding cavity. These waters have low entropy due to positional restrictions and a high enthalpy due to lack of polar interactions. High ΔG values along with high enthalpy (ΔH ≫ 0kcal/mol) indicate that it is favourable to displace them. Any ligand with nonpolar atoms entering these sites will tend to displace these waters and get rewarded with a higher binding affinity. No favourable replacement sites were present in the binding cavity. Except for a few sites, most of the other hydration sites in the binding cavity are highly conserved and can only be utilized to form bridging interactions. One of the conserved water molecules used for our study partially overlaps this kind of hydration site.

Simple docking calculations ignore the free energy contributions from desolvation of binding pocket. Data from WaterMap calculations can be used to optimize the ligand pose in the binding site. WM/MM scoring was performed for

Table 4. Residues of M^{pro} involved in interactions with the selected ligands.

Sr.	Title	Hydrogen bond	Salt bridge	π-π stacking	Hydrophobic interactions
Natural products Database					
a)	ZINC38139713	THR25, THR26, SER46, GLY143, GLU166, HIE164	–	HIE41	THR25, THR26, LEU27, HIE41, SER46, MET49, LEU141, ASN142, GLY143, CYS145, HIS163, HIE164, MET165, GLU166, ASP187, ARG188, GLN189
b)	ZINC31165691	HIE41, ASN142, GLU166, GLN189	–	–	THR25, LEU27, HIE41, THR45, SER46, MET49, LEU141, ASN142, CYS145, HIS163, MET165, GLU166, LEU167, ARG188, GLN189
c)	ZINC31167921	ASN142, GLU166, GLN189, THR190	–	HIE41	THR25, THR26, LEU27, HIE41, MET49, LEU141, ASN142, GLY143, CYS145, HIE164, MET165, GLU166, LEU167, PRO168, PHE181, ASP187, ARG188, GLN189, THR190, GLN192
d)	ZINC31165101	GLU166, GLN189	–	HIE41	THR25, THR26, LEU27, HIE41, SER46, MET49, LEU141, ASN142, GLY143, CYS145, HIS163, HIE164, MET165, GLU166, ASP187, ARG188, GLN189
e)	ZINC38143727	THR26, SER46, GLY143, GLU166	–	HIE41	THR25, THR26, LEU27, HIE41, SER46, MET49, LEU50, ASN142, GLY143, CYS145, HIE164, MET165, GLU166, ASP187, ARG188, GLN189
Enamine Database					
f)	Z3040398264	GLY143	–	HIE41	THR25, THR26, LEU27, HIE41, CYS44, THR45, SER46, MET49, ASN142, GLY143, CYS145, HIE164, MET165, GLU166, LEU167, PRO168, ASP187, ARG188, GLN189, THR190
g)	Z2706637055	HIE41, CYS44, ARG188	–	–	THR25, LEU27, HIE41, CYS44, THR45, SER46, MET49, ASN142, GLY143, CYS145, HIE164, MET165, GLU166, LEU167, PRO168, ASP187, ARG188, GLN189, THR190, GLN192
h)	Z2229246315	THR25, HIE41, GLN189	–	HIE41	THR24, THR25, THR26, LEU27, HIE41, THR45, SER46, MET49, GLY143, CYS145, HIE164, MET165, GLU166, ARG188, GLN189
i)	Z3061991917	GLY143, GLU166, GLN189	GLU166	HIE41	THR25, THR26, LEU27, HIE41, CYS44, THR45, SER46, MET49, PHE140, LEU141, ASN142, GLY143, CYS145, HIS163, MET165, GLU166, ARG188, GLN189
j)	Z2063273911	LEU141, ASN142, HIE164, GLU166, GLN189	–	HIE41	THR25, LEU27, HIE41, SER46, MET49, LEU141, ASN142, CYS145, HIS163, HIE164, MET165, GLU166, ASP187, ARG188, GLN189
Specs Database					
k)	ZINC49781425	THR26, PHE140, LEU141, GLU166, ARG188	–	HIE41	THR25, THR26, LEU27, HIE41, THR45, SER46, PHE140, LEU141, ASN142, GLY143, SER144, CYS145, HIS163, MET165, GLU166, LEU167, PRO168, ARG188, GLN189, THR190, GLN192
l)	ZINC67912395	GLU166, LEU167, GLN189	–	HIE41	THR24, THR25, THR26, LEU27, HIE41, CYS44, THR45, SER46, MET49, LEU141, ASN142, GLY143, CYS145, HIS163, HIE164, MET165, GLU166, LEU167, PRO168, ASP187, ARG188, GLN189
m)	ZINC67911229	LEU141, GLY143, GLU166, GLN189	–	–	THR24, THR25, THR26, HIE41, THR45, SER46, MET49, LEU141, ASN142, GLY143, SER144, CYS145, HIS163, MET165, GLU166, LEU167, PRO168, GLN189
n)	ZINC15675289	HIE41, GLY143	–	HIE41	THR25, THR26, LEU27, HIE41, MET49, LEU141, ASN142, GLY143, CYS145, HIE164, MET165, GLU166, ASP187, ARG188, GLN189
o)	ZINC44305556	THR26, SER46, PHE140, ASN142, GLU166, GLN189, ASN142	–	HIE41	THR25, THR26, LEU27, HIE41, THR45, SER46, MET49, PHE140, LEU141, ASN142, GLY143, SER144, CYS145, HIS163, MET165, GLU166, LEU167, ARG188, GLN189

the top ligands obtained after the hierarchical docking step. Pure MM-GBSA calculations ignore information generated by WaterMap calculations and WaterMap calculations alone don't correlate well with the experimental activities. The combined WM/MM scoring incorporates energetic terms from Prime MM-GBSA calculation to the WaterMap scores. This new scoring protocol correlates well to experimental binding energies. The WM/MM scoring was performed for the top hits identified using a hierarchical docking protocol to obtain binding free energies based on cumulative thermodynamic terms from MM-GBSA and WaterMap ΔG calculations. Combining the WM/MM scoring with docking protocol changes the ligand rankings significantly which may be related to experimental binding affinities in a much better way (Abel et al., 2011). The selected molecules had WM/MM scores in a range of around −18 to −26 kcal/mol as depicted in Table 2. Molecules ZINC38139713, ZINC31165691, ZINC31167921, ZINC49781425, ZINC67912395 and ZINC67911229 had good WM/MM ΔG bind values along with high docking scores. From the visual analysis, it was observed that many of the screened molecules had significant overlap with the three most unstable hydration sites in the binding cavity and a few of them showed a more significant binding affinity in terms of the calculated parameters. Figure 4 indicates the interaction of one such ligand ZINC31167921 with the WaterMap. The ligand ZINC31167921 effectively overlapped the unfavourable hydration sites identified using WaterMap calculations. The aromatic
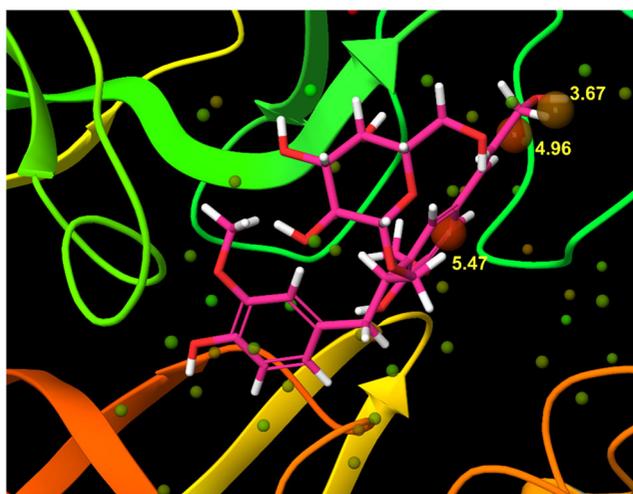

Figure 4. 5R82 WaterMap Ligand ZINC31167921 significantly overlaps unstable hydration sites contributing to binding affinity.

ring of the ligand fits snuggly into binding pocket S3 with an attached fragment entering the S4 pocket contributing to a high WaterMap ΔG bind value. In a similar fashion, ligands Z2706637055 and ZINC67912395 a satisfactory overlap with those hydration sites. Overall, WM/MM scoring helped in identification of ligands containing a balance of polar and hydrophobic interactions necessary for viral enzyme inhibition.

ADMET analysis

ADME properties of the selected compounds were predicted using Qikprop module of Schrödinger. Various basic physiochemical properties such as PlogPo/w (Predicted octanol/water partition coefficient), PlogHERG (Predicted IC50 value for the blockage of HERG K + channels), PPCaco (Predicted Caco-2 cell permeability for the gut blood barrier), PlogBB (predicted brain/blood partition coefficient) and PlogKhsa (Predicted binding to human serum albumin) of these compounds were calculated. The calculated parameter values of the selected compounds are given in Table 3. Most of the selected molecules fall in a decent range of values for the given parameters.

In silico toxicity parameters namely mutagenicity and hepatoxicity were assessed using pkCSM server (Pires et al., 2015) which makes use of graph-based signatures to calculate properties. The predicted toxicity parameters viz.; Drug induced liver injury (DILI) and AMES toxicity for the selected compounds are listed in Table 3.

Molecular dynamic simulations

Molecular dynamics simulations are essential to confirm the stability of the predicted ligand binding mode (Liu et al., 2017). We performed a 100 ns MD simulation for two of the best molecules to inspect their mechanical stability in complex and ensure stable binding under dynamic conditions. The degree of variation in RMSD is inversely related to the stability of complex: the smaller the variation, the greater the stability. The RMSD of both complexes showed minimum difference (RMSD < 1 Å) in reference to their initial coordinates which states that the complexes were stable throughout the simulation, as shown in Figure 5.

Later we focused on the local atomic level fluctuations of the protein in dynamic conditions. The RMSF of complex 1 (ZINC31167921) showed a higher degree of fluctuation
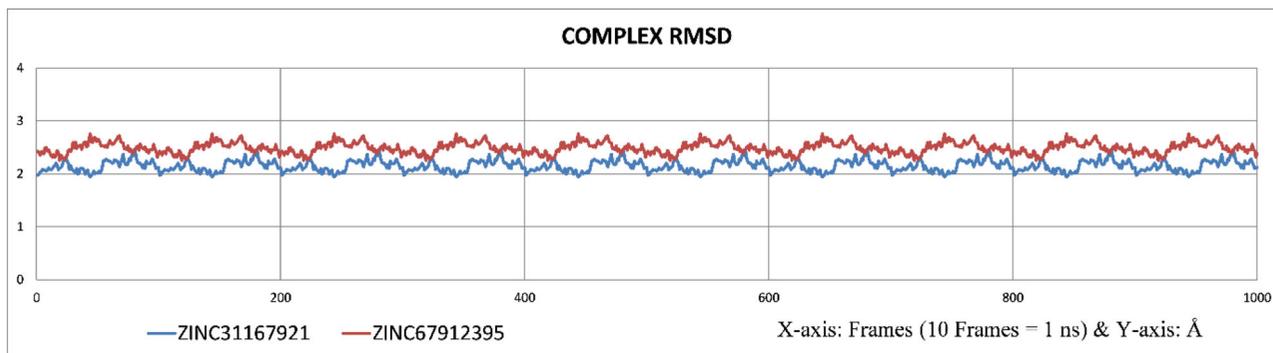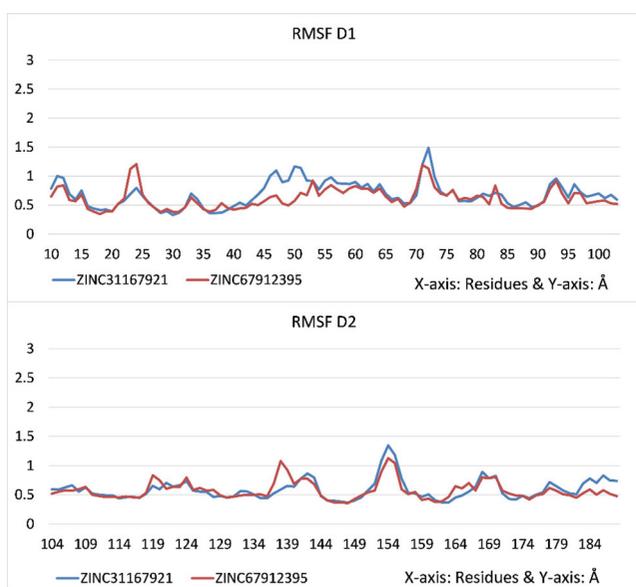

Figure 5. All atom RMSD of the protein complexes with ZINC31167921 and ZINC67912395.

compared to complex 2 (ZINC67912395). The RMSF analysis of each complex revealed a high fluctuation region which corresponds to domain III (between residues position 201-303) of the protein. This portion is far from the complex's interface or the active site and is associated with protein dimerization. Most of the binding site residues in domain I and domain II have lesser fluctuations with RMSF values below 1 Å (shown as D1 and D2 in Figure 6). Although not involved directly in the ligand binding, residues like THR24 form a part of the catalytic site as well. It has slightly higher fluctuation, which may be attributed to its location in the turn region between the two β-sheets. Another residue GLY138 with a little high RMSF value in complex 2 is located in the loop region at the interface of domain II and domain III. The fluctuations in all the three domains were more or less same in both complexes with slightly higher values in complex 1. Our interaction analysis showed that small-molecule ZIN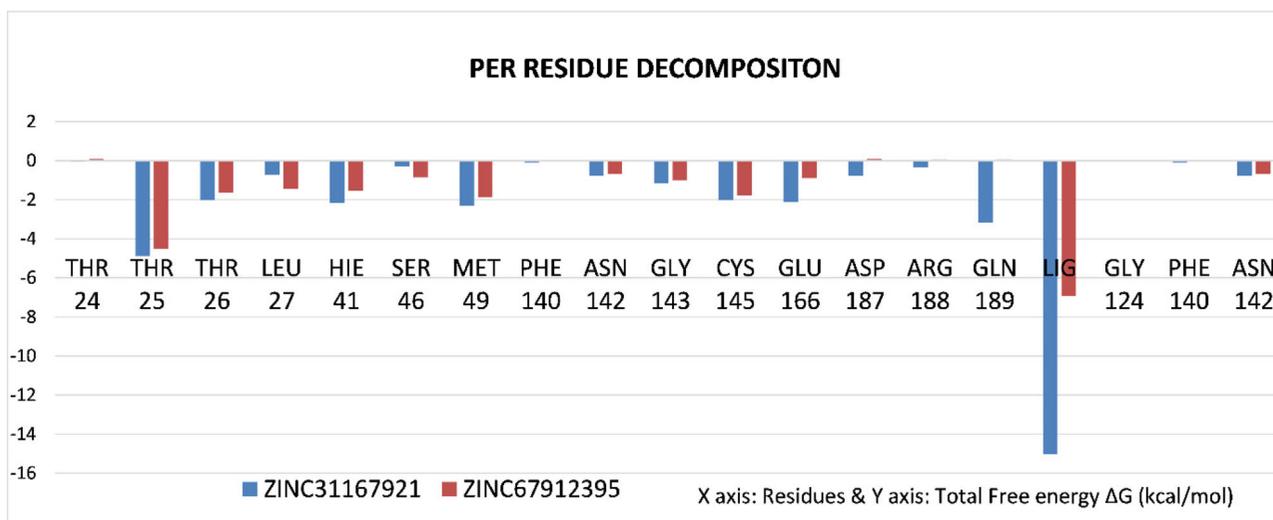C31167921 had formed a more number of intermolecular H-bonds with the protein in complex 1 compared to ZINC67912395 in complex 2. The oxygen atom of THR25 can form H-bond with oxygen of ZINC31167921 with an occupancy of 90.8%. Meanwhile oxygen atom as well as nitrogen of GLU166 displayed two hydrogen bonds with ZINC31167921 during the course of simulation with average occupancy of 54.5 and 43.2 percent. Another H-bond with an occupancy of 48 percent was formed with the side chain nitrogen at delta position which was absent earlier. Other residues like GLN189, THR26 and GLN192 displayed some propensity to form H-bonds with ZINC31167921. The gamma oxygen of THR25 as well as oxygen and nitrogen atoms of THR26 displayed hydrogen bonds with ZINC67912395 with an occupancy of 51.6, 65.6 and 46.4 percent. Oxygen atoms of ASP187 and GLU166 displayed H-bonds with ZINC67912395 with occupancy of 43.5 and 34.1 percent. The ligand ZINC67912395 also displayed H-bond interactions with side chains of ASN142 and GLN189 with average occupancy of 34.1 and 23.2 percent. The graphs highlighting the hydrogen bond interaction data are presented in supplementary data Figure S2.

MM-GBSA analysis revealed a high van der Waals contribution in the protein and small molecule binding. Overall binding free energy (ΔG) for ZINC31167921 was found to be −50.2062 kcal/mol and −24.6658 kcal/mol for ZINC67912395. Residues THR25, THR26, HIE41, MET49, CYS145, GLU166 and GLN189 contributed to ligand binding. Catalytic residues of M^{pro} (HIE41 and CYS145) contributed to the ligand-binding in both cases. Per residue scores indicated that THR25 contributed maximum to the binding free energy, as evident from Figure 7. Contribution of all the above-mentioned residues was higher for ZINC31167921 than ZINC67912395, which is apparent from the overall binding free energy (ΔG) values.

Conclusion

The COVID-19 pandemic calls for a quick response from the scientific realm and with no specific remedy available, computational approach is the key for identification of potential

Figure 6. The observed RMSF for C-α atoms of protein residues from domain I (Residues 10-103) and domain II (Residues 104-188).

Figure 7. Energy contribution of single residues in the system to the overall binding free energy (ΔG).

inhibitors that can be repurposed for effective therapeutics. The main protease involvement in viral replication makes it an important target to inhibit the virus growth. The present study has offered insights into the possible route of drug design against SARS-CoV-2 M^{pro}.

An organized approach involving three-tiered molecular docking and water thermodynamics was followed to perform this study. We report 15 candidates in this paper that show optimal binding characteristics to the viral main protease. Potential leads obtained from the binding affinity screening studies can be further utilized for *in vitro*, *in vivo* analysis. From the results obtained we can expect that the identified molecules ZINC31167921 and ZINC67912395 may exhibit good antiviral activity against the new coronavirus. Molecular dynamics and MM-GBSA results verified that the complexes are not only thermodynamically stable but structurally robust. Although this data doesn't confirm the antiviral activity of the selected molecules, yet it provides a basis for further *in vitro/ in vivo* studies. The information derived from water thermo-dynamic studies offers possible routes to design and optimize these hits to increase their selectivity towards SARS-CoV-2. Our subsequent studies will concentrate on understanding the cova-lent interaction abilities of identified molecules possessing reactive warhead groups (Z3040398264, Z2063273911, ZINC67911229, ZINC5675289 and ZINC44305556) and establish-ing their stability in the complex.

Acknowledgements

Disclosure statement

The authors declare no competing interests.

Author contributions

M.E.S designed the study. All the authors contributed to computations, analysis and preparation of the manuscript.

References

Abd El-Aziz, T. M., & Stockand, J. D. (2020). Recent progress and challenges in drug development against COVID-19 coronavirus (SARS-CoV-2) - an update on the status. *Infection, Genetics and Evolution : journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 83*, 104327 https://doi.org/10.1016/j.meegid.2020.104327

Abel, R., Salam, N. K., Shelley, J., Farid, R., Friesner, R. A., & Sherman, W. (2011). Contribution of explicit solvent effects to the binding affinity of small-molecule inhibitors in blood coagulation factor serine proteases. *Chemmedchem., 6*(6), 1049–1066. https://doi.org/10.1002/cmdc.201000533

Agbowuro, A. A., Huston, W. M., Gamble, A. B., & Tyndall, J. D. A. (2018). Proteases and protease inhibitors in infectious diseases. *Med Res Rev, 38*(4), 1295–1331. https://doi.org/10.1002/med.21475

Al-Kassmy, J., Pedersen, J., & Kobinger, G. (2020). Vaccine candidates against coronavirus infections. Where does COVID-19 stand? *Viruses, 12*(8), 861. https://doi.org/10.3390/v12080861

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., & Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science (New York, N.Y.), 300*(5626), 1763–1767. https://doi.org/10.1126/science.1085658

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry, 26*(16), 1668–1688. https://doi.org/10.1002/jcc.20290

Chen, H., Zhang, Z., Wang, L., Huang, Z., Gong, F., Li, X., Chen, Y., & Wu, J. J. (2020). First Clinical Study Using HCV Protease Inhibitor Danoprevir to Treat Naive and Experienced COVID-19 Patients. medRxiv: 2020.2003.2022.20034041.

Chen, J., Xia, L., Liu, L., Xu, Q., Ling, Y., Huang, D., Huang, W., Song, S., Xu, S., Shen, Y., & Lu, H. (2020). Antiviral activity and safety of daruna-vir/cobicistat for the treatment of COVID-19. *Open Forum Infectious Diseases, 7*(7), ofaa241 https://doi.org/10.1093/ofid/ofaa241

Elfiky, A. A. (2020). Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sciences, 248*, 117477 https://doi.org/10.1016/j.lfs.2020. 117477

Felgenhauer, U., Schoen, A., Gad, H. H., Hartmann, R., Schaubmar, A. R., Failing, K., Drosten, C., & Weber, F. (2020). Inhibition of SARS-CoV-2 by type I and type III interferons. *Journal of Biological Chemistry. 295*(41), 13958-13964. https://doi.org/10.1074/jbc.ac120.013788

Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., & Mainz, D. T. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry, 49*(21), 6177–6196. https://doi.org/10.1021/jm051256o

Fu, L., Ye, F., Feng, Y., Yu, F., Wang, Q., Wu, Y., Zhao, C., Sun, H., Huang, B., Niu, P., Song, H., Shi, Y., Li, X., Tan, W., Qi, J., & Gao, G. F. (2020). Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by target-ing its main protease. *Nature Communications, 11*(1), 4417 https://doi. org/10.1038/s41467-020-18233-x

Ghosh, A. K., Xi, K., Grum-Tokars, V., Xu, X., Ratia, K., Fu, W., Houser, K. V., Baker, S. C., Johnson, M. E., & Mesecar, A. D. (2007). Structure-based design, synthesis, and biological evaluation of peptidomimetic SARS-CoV 3CLpro inhibitors. *Bioorg Med Chem Lett, 17*(21), 5876–5880. https://doi.org/10.1016/j.bmcl.2007.08.031

Ghosh, A. K., Xi, K., Ratia, K., Santarsiero, B. D., Fu, W., Harcourt, B. H., Rota, P. A., Baker, S. C., Johnson, M. E., & Mesecar, A. D. (2005). Design and synthesis of peptidomimetic severe acute respiratory syndrome chymotrypsin-like protease inhibitors. *Journal of Medicinal Chemistry, 48*(22), 6767–6771. https://doi.org/10.1021/jm050548m

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R., & Friesner, R. A. (2016). OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation, 12*(1), 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Harrach, M. F., & Drossel, B. (2014). Structure and dynamics of TIP3P, TIP4P, and TIP5P water near smooth and atomistic walls of different hydroaffinity. *Journal of Chemical Physics., 140*(17), 174501. https://doi. org/10.1063/1.4872239

Huggins, D. J. (2020). Structural analysis of experimental drugs binding to the SARS-CoV-2 target TMPRSS2. *Journal of Molecular Graphics & Modelling, 100*, 107710 https://doi.org/10.1016/j.jmgm.2020.107710

Ivashchenko, A. A., Dmitriev, K. A., Vostokova, N. V., Azarova, V. N., Blinow, A. A., Egorova, A. N., Gordeev, I. G., Ilin, A. P., Karapetian, R. N., Kravchenko, D. V., Lomakin, N. V., Merkulova, E. A., Papazova, N. A., Pavlikova, E. P., Savchuk, N. P., Simakina, E. N., Sitdekov, T. A., Smolyarchuk, E. A., Tikhomolova, E. G., Yakubova, E. V., & Ivachtchenko, A. V. (2020). AVIFAVIR for Treatment of Patients with Moderate COVID-19: Interim Results of a Phase II/III Multicenter Randomized Clinical Trial. *Clinical Infectious Diseases*. https://doi.org/ 10.1093/cid/ciaa1176

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., … Yang, H. (2020). Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature, 582*(7811), 289–293. https://doi.org/10.1038/s41586-020-2223-y

Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., Duan, Y., Yu, J., Yang, X., Yang, X., Yang, K., Liu, X., Guddat,

L. W., Xiao, G., Zhang, L., Yang, H., & Rao, Z. (2020). Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat Struct Mol Biol*, 27(6), 529–532. https://doi.org/10.1038/s41594-020-0440-6

Kumar, V., Shin, J. S., Shie, J. J., Ku, K. B., Kim, C., Go, Y. Y., Huang, K. F., Kim, M., & Liang, P. H. (2017). Identification and evaluation of potent Middle East respiratory syndrome coronavirus (MERS-CoV) 3CLPro inhibitors. *Antiviral Research*, 141, 101–106. https://doi.org/10.1016/j.antiviral.2017.02.007

Kumar, V., Tan, K. P., Wang, Y. M., Lin, S. W., & Liang, P. H. (2016). Identification, synthesis and evaluation of SARS-CoV and MERS-CoV 3C-like protease inhibitors. *Bioorganic & Medicinal Chemistry*, 24(13), 3035–3042. https://doi.org/10.1016/j.bmc.2016.05.013

Lei, C., Qian, K., Li, T., Zhang, S., Fu, W., Ding, M., & Hu, S. (2020). Neutralization of SARS-CoV-2 spike pseudotyped virus by recombinant ACE2-Ig. *Nature Communications*, 11(1), 2070 https://doi.org/10.1038/s41467-020-16048-4

Li, G., & De Clercq, E. (2020). Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nature Reviews Drug Discovery*, 19(3), 149–150. https://doi.org/10.1038/d41573-020-00016-0

Li, Y., Xie, Z., Lin, W., Cai, W., Wen, C., Guan, Y., Mo, X., Wang, J., Wang, Y., Peng, P., Chen, X., Hong, W., Xiao, G., Liu, J., Zhang, L., Hu, F., Li, F., Zhang, F., Deng, X., & Li, L. (2020). Efficacy and Safety of Lopinavir/Ritonavir or Arbidol in Adult Patients with Mild/Moderate COVID-19: An Exploratory Randomized Controlled Trial. *Med (N Y)*, 1(1), 105–113.e4. https://doi.org/10.1016/j.medj.2020.04.001

Liu, K., Watanabe, E., & Kokubo, H. (2017). Exploring the stability of ligand binding modes to proteins by molecular dynamics simulations. *Journal of Computer-Aided Molecular Design*, 31(2), 201–211. https://doi.org/10.1007/s10822-016-0005-2

Liu, L., Wang, P., Nair, M. S., Yu, J., Rapp, M., Wang, Q., Luo, Y., Chan, J. F., Sahi, V., Figueroa, A., Guo, X. V., Cerutti, G., Bimela, J., Gorman, J., Zhou, T., Chen, Z., Yuen, K. Y., Kwong, P. D., Sodroski, J. G., … Ho, D. D. (2020). Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature*, 584(7821), 450–456. https://doi.org/10.1038/s41586-020-2571-7

Ma, C., Sacco, M. D., Hurst, B., Townsend, J. A., Hu, Y., Szeto, T., Zhang, X., Tarbet, B., Marty, M. T., Chen, Y., & Wang, J. (2020). Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Research*, 30(8), 678–692. https://doi.org/10.1038/s41422-020-0356-z

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation*, 11(8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255

Menendez, C. A., Bylehn, F., Perez-Lemus, G. R., Alvarado, W., & de Pablo, J. J. (2020). Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease. *Science Advances*, 6(37), eabd0345. https://doi.org/10.1126/sciadv.abd0345

Nicolini, L. A., Mikulska, M., Signori, A., Di Biagio, A., Portunato, F., Vena, A., Giacomini, M., & Bassetti, M. (2020). Reply to: "Antiviral activity and safety of darunavir/cobicistat for treatment of COVID-19". *Open Forum Infectious Diseases*, 7(8), ofaa321.https://doi.org/10.1093/ofid/ofaa321

Paules, C. I., Marston, H. D., & Fauci, A. S. (2020). Coronavirus infections-more than just the common cold. *JAMA*, 323(8), 707–708. https://doi.org/10.1001/jama.2020.0757

Pires, D. E., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *Journal of Medicinal Chemistry*, 58(9), 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

Riniker, S., Barandun, L. J., Diederich, F., Kramer, O., Steffen, A., & van Gunsteren, W. F. (2012). Free enthalpies of replacing water molecules in protein binding pockets. *Journal of Computer-Aided Molecular Design*, 26(12), 1293–1309. https://doi.org/10.1007/s10822-012-9620-8

Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of Chemical Theory and Computation*, 9(9), 3878–3888. https://doi.org/10.1021/ct400314y

Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: A software program for pK(a) prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design*, 21(12), 681–691. https://doi.org/10.1007/s10822-007-9133-z

Shivanyuk, A., Ryabukhin, S., Tolmachev, A., Bogolyubsky, A., Mykytenko, D., Chupryna, A., Heilman, W., & Kostyuk, A. (2007). Enamine real database: Making chemical diversity real. *Chemistry Today*, 25, 58–59.

Shu, T., Huang, M., Wu, D., Ren, Y., Zhang, X., Han, Y., Mu, J., Wang, R., Qiu, Y., Zhang, D. Y., & Zhou, X. (2020). SARS-coronavirus-2 Nsp13 Possesses NTPase and RNA helicase activities that can be inhibited by bismuth salts. *Virologica Sinica*, 35(3), 321–329. https://doi.org/10.1007/s12250-020-00242-1

Smolders, E. J., Te Brake, L. H., & Burger, D. M. (2020). SARS-CoV-2 and HIV protease inhibitors: Why lopinavir/ritonavir will not work for COVID-19 infection. *Antiviral Therapy*, https://doi.org/10.3851/IMP3365

Sterling, T., & Irwin, J. J. (2015). ZINC 15-ligand discovery for everyone. *Journal of Chemical Information and Modeling*, 55(11), 2324–2337. https://doi.org/10.1021/acs.jcim.5b00559

Ullrich, S., & Nitsche, C. (2020). The SARS-CoV-2 main protease as drug target. *Bioorganic & Medicinal Chemistry Letters*, 30(17), 127377 https://doi.org/10.1016/j.bmcl.2020.127377

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9), 1157–1174. https://doi.org/10.1002/jcc.20035

WHO (2020). *Coronavirus disease 2019 (COVID-19): weekly epidemiological update*. Report no. Geneva. https://apps.who.int/iris/handle/10665/334188

Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C., & Zhang, Y. Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), 265–269. https://doi.org/10.1038/s41586-020-2008-3

Yin, W., Mao, C., Luan, X., Shen, D. D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., Chang, S., Xie, Y. C., Tian, G., Jiang, H. W., Tao, S. C., Shen, J., Jiang, Y., Jiang, H., Xu, Y., … Xu, H. E. (2020). Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science (New York, N.Y.)*, 368(6498), 1499–1504. https://doi.org/10.1126/science.abc1560

Yuen, C. K., Lam, J. Y., Wong, W. M., Mak, L. F., Wang, X., Chu, H., Cai, J. P., Jin, D. Y., To, K. K., Chan, J. F., Yuen, K. Y., & Kok, K. H. (2020). SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerging Microbes & Infections*, 9(1), 1418–1428. https://doi.org/10.1080/22221751.2020.1780953

Zhang, L., Lin, D., Kusov, Y., Nian, Y., Ma, Q., Wang, J., von Brunn, A., Leyssen, P., Lanko, K., Neyts, J., de Wilde, A., Snijder, E. J., Liu, H., & Hilgenfeld, R. (2020). α-Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: Structure-based design, synthesis, and activity assessment. *Journal of Medicinal Chemistry*, 63(9), 4562–4578. https://doi.org/10.1021/acs.jmedchem.9b01828

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science (New York, N.Y.)*, 368(6489), 409–412. https://doi.org/10.1126/science.abb3405

Ziebuhr, J., Snijder, E. J., & Gorbalenya, A. E. (2000). Virus-encoded proteinases and proteolytic processing in the Nidovirales. *The Journal of General Virology*, 81(Pt 4), 853–879. https://doi.org/10.1099/0022-1317-81-4-853