Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
1,2,4 triazolo[1,5-a] pyrimidin-7-ones as novel SARS-CoV-2 Main protease inhibitors: In silico screening and molecular dynamics simulation of potential COVID-19 drug candidates

Kuppuswamy Kavitha\textsuperscript{a}, Subramaniam Sivakumara, Balasubramanian Rameshc,\textsuperscript{c}\textsuperscript{*}

\textsuperscript{a} Research Department of Microbiology, Sri Sankara Arts and Science College, Kanchipuram 631 561, India
\textsuperscript{b} Research Department of Biochemistry, Sri Sankara Arts and Science College, Kanchipuram 631 561, India
\textsuperscript{c} Research Department of Biotechnology, Sri Sankara Arts and Science College, Kanchipuram 631 561, India

\begin{abstract}
Discovery of a potent SARS-CoV-2 main protease (M\textsuperscript{pro}) inhibitor is the need of the hour to combat COVID-19. A total of 1000 protease-inhibitor-like compounds available in the ZINC database were screened by molecular docking with SARS-CoV-2 M\textsuperscript{pro} and the top 2 lead compounds based on binding affinity were found to be 1,2,4 triazolo[1,5-a] pyrimidin-7-one compounds. We report these two compounds (ZINC000621278586 and ZINC000621285995) as potent SARS-CoV-2 M\textsuperscript{pro} inhibitors with high affinity (< −9 kCal/mol) and less toxicity than Lopinavir and Nelfinavir positive controls. Both the lead compounds effectively interacted with the crucial active site amino acid residues His41, Cys145 and Glu166. The lead compounds satisfied all of the druglikeness rules and devoid of toxicity or mutagenicity. Molecular dynamics simulations showed that both lead 1 and lead 2 formed stable complexes with SARS-CoV-2 M\textsuperscript{pro} as evidenced by the highly stable root mean square deviation (< 0.23 nm), root mean square fluctuations (0.12 nm) and radius of gyration (2.2 nm) values. Molecular mechanics Poisson-Boltzmann surface area calculation revealed thermodynamically stable binding energies of −129.266 ± 2.428 kJ/mol and −116.478 ± 3.502 kJ/mol for lead1 and lead2 with SARS-CoV-2 M\textsuperscript{pro}, respectively.
\end{abstract}

\begin{keywords}
COVID-19
SARS-CoV-2 Main protease inhibitor
Molecular docking
Molecular dynamics simulation
Novel antiviral compound
1,2,4 triazolo[1,5-a] pyrimidin-7-one
\end{keywords}
1. Introduction

Coronaviruses (CoV) generally cause mild to moderate upper-respiratory tract illnesses and can infect humans and most of the animal species. However, it has recently caused severe pulmonary diseases like severe acute respiratory syndrome (SARS), middle east respiratory syndrome (MERS) and currently coronavirus disease 2019 (COVID-19). The latest pathogen is severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of July 8, 2020, this virus affected 11,869,690 people and killed 544,147 worldwide. It caused severe economic crises and impacts on financial markets. As per the UN Department of Economic and Social Affairs, the world economy could contract by 0.9% in 2020 as opposed to a previous forecast of 2.5% growth [1].

The SARS-CoV2 is comprised of a positive-strand RNA genome of size 29.7kb and encodes a viral replica that is associated with the novel genome synthesis and generation of a nested set of sub-genomic messenger RNAs, encoding both structural proteins present in all CoVs: Spike (S), Envelope (E), Membrane (M) and Nucleoprotein (N), and a group of proteins specific for SARS-CoV: 3a, 3b, 6, 7a, 7b, 8a, 8b, and 9b [2].

So far, there is neither a drug nor a vaccine for COVID-19. The rapid development and identification of efficient interventions against SARS-CoV-2 remains a major challenge. Elfiky showed that Sofosbuvir, Ribavirin, Galidesivir, Remdesivir, Favipiravir, Cefuroxime, Tenofivir, and Hydroxychloroquine could bind to the RdRp active site tightly and supposed to be good candidates for clinical trials [3]. Recently, Stilbenei analogues have been reported to be potential disruptors of the SARS-CoV-2 spike protein and human ACE2 receptor complex [4]. One study suggested hydroxychloroquine and azithromycin as a treatment for COVID-19 [5] and immediately refuted by others [6]. Remdesivir and chloroquine were shown to inhibit SARS-CoV-2 in vitro [7]. Lopinavir exhibited an anti-CoV effect in vitro and is tried for clinical treatment of COVID-19 [8,9]. Nelfinavir was shown to inhibit replication of the SARS coronavirus (SARS-CoV), which could reduce the replication of virions from Vero cells [10] and was predicted to be a potential inhibitor of SARS-CoV-2 main protease [11]. Attention has been given to the development of furin inhibitors as a potential therapeutic platform against SARS-CoV-2. However, furin-like enzymes contribute to several pathways and systemic inhibition may lead to some adverse effects [12]. Although repurposing of drugs is a good idea, when their effectiveness is not certain, novel drugs are to be designed and developed specifically for novel viruses like SARS-CoV-2. Structure-based virtual screening and molecular dynamics approaches are particularly suitable to identify novel SARS-CoV-2 inhibitors [13].

The coronavirus main protease (Mpro) is essential for the viral gene expression and replication by the proteolytic cleavage of replicase polyproteins, without which the virus replication is severely hampered and is an important target for anti-CoV drug design [14]. Mpro has emerged as the most potent antiviral target because of its main role in self-maturation and subsequent maturation of polyproteins [15]. X-ray structures of the unliganded SARS-CoV-2 Mpro and its complexes with various ligands have been reported. Since there are no human counterparts with similar cleavage specificity, inhibitors of SARS-CoV-2 Mpro are unlikely to be toxic [16].

SARS-CoV-2 Mpro is a cysteine protease containing Cys-145 and His-41 catalytic dyad in its active center. The proteolytic process is believed to be dependant on active site cysteine (Cys-145) side chain thiolate nucleophile attack on amide bond of the substrate [17]. The NH2 group of Cys145 is ion-paired with His41 forming Cys145-His41 catalytic dyad, which differs from most serine proteases that have a catalytic Ser-His-Asp triad in their active sites. In Mpro, a stable water molecule occupies the Asp position of the typical serine protease triad [18].

Both covalent and non-covalent inhibitors of Mpro are of immense value as a potent drug against SARS-CoV-2. Covalent inhibitors establish a covalent bond (C–S) with the reactive thiol group of Cys145 and form favourable interactions with residues lining the substrate-binding site [19]. Non-covalent inhibitors mainly act by binding to the active site stronger than the natural substrate by non-covalent bonds like hydrogen bonds, van der walls interactions, and electrostatic interactions. Covalent inhibitors are highly selective inhibitors. However, irreversible drug toxicity can be a real challenge related to this class of therapeutics. On the other hand, non-covalent inhibitors could never cause irreversible toxicity but might be less effective. It is evident that both classes have their merits [20].

The objectives of this study were i) to identify evolutionarily important active site amino acids by structure-based sequence alignment of SARS-CoV-2 and SARS-CoV Mpro enzymes ii) to identify potential non-covalent Mpro inhibitors by screening protease-inhibitor-like compounds available in the ZINC database by molecular docking studies iii) prediction of absorption, distribution, metabolism, excretion and toxicity properties of the top-scoring inhibitors using in silico methods iv) to validate the stable binding of the lead compounds with SARS-CoV-2 Mpro by molecular dynamics (MD) simulations and v) to calculate thermodynamic binding energies for each lead compound - SARS-CoV-2 Mpro complex using Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) calculations.

2. Materials and methods

2.1. Crystal structure of SARS-CoV-2 and SARS-CoV Mpro enzymes and small molecular library

The three-dimensional structures of SARS-CoV-2 Mpro (PDB IDs: 6LU7, 6Y84, 6Y87, 5RE4 and 6W63) were obtained from RCSB-PDB [21]. PDB structures of SARS-CoV Mpro (SNH0, 1P9S and 2ZU2) with 95% structural similarities with SARS-CoV-2 Mpro were selected using the jFATCAT-rigid algorithm [22] and retrieved. The sequences of all these structures were used for further structure-based sequence alignment of SARS-CoV-2 and SARS-CoV Mpro enzymes. The crystal structure of SARS-CoV-2 Mpro in complex with an inhibitor N3 determined by X-ray diffraction with 2.16 Å resolution (PDB ID: 6LU7) [23] was used as the drug target for molecular docking and molecular dynamics (MD) studies. Ton et al. [24] screened 1.3 billion compounds from the ZINC15 [25] library and identified 1000 probable ligands for SARS-CoV-2 Mpro protein. The compounds were made publicly available for further research by the scientific community. All these 1000 ligands for the SARS-CoV-2 Mpro protein were downloaded in SDF format and used as the small molecular library for screening.

2.2. Structure-based sequence alignment of SARS-CoV-2 and SARS –CoV Mpro enzymes

Structure-based sequence alignment was carried out to discern the amino acids that are conserved evolutionarily, particularly in the active site. The sequences of all Mpro structures were exported into FASTA format and aligned using Clustalo [26] and their evolutionary relationship was inferred by the neighbor-joining method [27] using Mega X [28]. The bootstrap consensus tree resulting from 500 replicates represented the evolutionary history [29]. The Poisson correction method was used to calculate evolutionary distances [30]. This analysis involved 8 Mpro sequences. Ambiguous positions were removed by the pairwise deletion option and finally 307 positions were included in the dataset. Structure-based alignment was performed and important features of the sequences and structures were deciphered using ESPript [31].

2.3. Preparation of SARS-CoV-2 Mpro and small molecule library for docking

The selected drug target with PDB ID 6LU7 [23] was prepared at pH (7.0), water molecules and the inhibitor were removed from the structure and incomplete residues were fixed using UCSF Chimera
Fig. 1. A) Evolutionary relationships of SARS-CoV-2 Mpro sequences extracted from PDB structures 6LU7, 6Y84, 6YB7, 6W63, 5RE4 and SARS-CoV main protease structures 5NHO, 1P95 and 2ZU2 inferred using the neighbor-joining method. The evolutionary distances are in units of the number of amino acid substitutions per site. B) Structure-based sequence alignment of SARS-CoV-2 and SARS-CoV main proteases is shown and their secondary structural features are shown above and below the alignment, respectively. Amino acids conserved in all sequences are shaded. Active site amino acid dyad His41 and Cys145 are labeled in red and blue, respectively. The dimerization site amino acid Glu166 is well conserved in all SARS-CoV-2 Mpro and SARS-CoV Mpro sequences. All other active site amino acids of SARS-CoV-2 Mpro are labeled in black.
2.6. Molecular dynamics (MD) simulations

lead compounds were predicted by the AdmetSAR 2.0 [43].

GROMACS. The PRODRG2.5 server [45] was used to build the topology parameters of lead1, lead2, lead3, Lopinavir and Nelfinavir. MD simulation of 20 ns with a time step of 2 fs at a 300 K temperature was carried out. A total of 6 systems; one SARS-CoV-2 Mpro apoenzyme (6 LU7) and 5 SARS-CoV-2 Mpro complexes viz., 6 LU7-Lead1, 6 LU7-Lead2, 6 LU7-Lead3, 6 LU7-Lopinavir and 6 LU7-Nelfinavir were prepared. The apo-protein and protein-ligand complexes were submerged in a solvent box, surrounded by 4 Na ions to maintain electro-neutrality. Energy minimization was done using the steepest descent algorithm in order to alleviate the bad van der Waals interactions strain. After the convergence of the system, equilibration was carried out with NVT and NPT ensembles to attain the system temperature and pressure of 300 K and 1 bar, respectively. The electrostatic interaction in the systems was measured with the particle mesh Ewald. The GROMACS molecular dynamics simulation engine “mdrun” program was used to carry out equilibration MD simulations. The temperature and pressure of the system were kept constant using the velocity-rescale algorithm and the Parrinello-Rahman algorithm. The LINEAR Constraint Solver algorithm was utilized to restrain all the bonding lengths [46].

2.7. Calculation of the thermodynamic parameters

Binding energy of each protein-ligand complex was calculated by the MM-PBSA method [49] using the g_mmpbsa tool [50]. The binding energy of each complex was computed from van der Waals energy, electrostatic energy, polar solvation energy and non-polar solvation energy based on the solvent accessible surface area (SASA) model using the script MmPbSaStat.py. The final contribution energy of each residue from individual energetic terms obtained from the g_mmpbsa was calculated using MmPbSaDecomp.py script.

3. Results and discussion

3.1. Structure-based sequence alignment of SARS-CoV-2 and SARS-CoV Mpro enzymes

SARS-CoV-2 Mpro structures 6 LU7, 6Y84, 6YB7, 5RE4 and 6 W63 and SARS-CoV Mpro structures with 95% similarities, viz. 5NH0, 1P9S and 2ZU2 were obtained from RCSB-PDB. The sequences of all Mpro structures were aligned and their evolutionary relationships (Fig. 1A) were inferred. A clear distinction for SARS-CoV-2 Mpro was observed when the evolutionary distances were calculated for 5 SARS-CoV-2 Mpro and 3 SARS-CoV Mpro sequences.

SARS-CoV-2 and SARS-CoV main proteases shared common secondary structural features viz. 2 α Helices and 7 β sheets in Domain I, 6 β sheets in Domain II and 5 α Helices in Domain III (Fig. 1B). Amino acid conservation is high in Domains I and II compared to Domain III. Active site amino acid dyad His41, Cys145 and the dimerization site Glu166 are conserved in both SARS-CoV-2 and SARS-CoV as reported previously [16]. The overall molecular architecture of SARS-CoV-2 Mpro was similar to SARS-CoV Mpro and consistent with previous reports [41].

Sequence alignment showed some residues that are unique for SARS-CoV-2 in comparison to SARS-CoV, which further showed a profound effect on docking studies resulting in new lead compounds, which were not reported for SARS-CoV Mpro. The evolutionary replacement of amino acids from SARS-CoV to SARS-CoV-2 were found to be Leu3Phe,
Table 1
Top 10 Lead compounds with positive controls based on docking results.

| Lead-Zinc ID/Name | Structure | Binding Affinity (kCal/mol) | Interacting amino acids | Interacting amino acids | Interacting amino acids | Amino acids unique to SARS-CoV-2 MPro |
|-------------------|-----------|-----------------------------|-------------------------|-------------------------|-------------------------|---------------------------------------|
| Lead1-ZINC000621278586 | ![Image](image1.png) | −9.3 | Phe140, Leu141, Gly143, Ser144, Cys145, His164, Glu166 | His41, Met49, Met165 | – | Met49, Leu141, Ser144, His164, Met165 |
| Lead2-ZINC000621285995 | ![Image](image2.png) | −9.1 | Phe140, Leu141, Ser144, Cys145, His164, Glu166 | His41, Met49, Met165 | – | Met49, Leu141, Ser144, His164, Met165 |
| Lead3-ZINC000566550443 | ![Image](image3.png) | −9 | Phe140, Ser144, Cys145, His164, Met165 | Glu166 | – | Met49, Leu141, Ser144, His164, Met165 |
| Lead4-ZINC000358396994 | ![Image](image4.png) | −9 | Phe140, Leu141, Gly143, Ser144, Glu166, Gln189, Arg188 | Met49, Met165, Leu167, Pro168 | – | His41, Cys145, Met49, Leu141, Ser144, His164, Met165, Pro168, Gln189, Arg188 |
| Lead5-ZINC000636416501 | ![Image](image5.png) | −8.8 | Thr25, Thr26, His41, Cys145 | Met49, His163 | Thr24 | Leu141, Gly143, Ser144, Met165, Glu166, Arg188, Gln189 | Thr24, Thr26, Met49, Leu141, Ser144, Arg188, Gln189 |
| Lead6-ZINC000621266801 | ![Image](image6.png) | −8.8 | Phe140, Ser144, Cys145, Glu166 | His41, Met165 | – | His163, Leu141, Ser144, Met165 |
| Lead7-ZINC000123269462 | ![Image](image7.png) | −8.8 | Tyr54, Leu141, Asn142, Gly143, Ser144, Cys145, Gln189, Arg188 | His41, Met165 | Asp187 | Phe140, Leu141, Ser144, Met165, Gln189, Arg188 |
| Lead8-ZINC000055656943 | ![Image](image8.png) | −8.8 | Leu141, Asn142, Gly143, Ser144, Cys145, His164 | Met165 | Glu166 | Asp187, Arg188, Gln189, Leu141, Ser144, His164, Met165, Arg188, Gln189 |
| Lead9-ZINC001627906106 | ![Image](image9.png) | −8.7 | Leu141, Gly143, Ser144, Cys145, His164 | His41, Met49, Met165 | – | Glu166, Met49, Ser144, His164, Met165 |
| Lead10-ZINC001331329001 | ![Image](image10.png) | −8.7 | Phe140, Leu141, Gly143, Ser144, Cys145, Glu166, Thr190 | His41, Met49, Met165 | – | His164, Met165, Met49, Leu141, His164, Met165, Gln189, Thr190 |

(continued on next page)
Gln8Phe, Phe12Lys, Lys15Gly, Val17Met, Arg19Gln, Cys21Thr, Tyr22Cys, Asn24Thr, Val26Thr, Gly33Asp, Ile/Thr35Val, Ala44Cys, Ser/Pro45Thr, Thr47Glu, Thr48Asp, Ser/Val49Met, Ile52Pro, Asp53Asn, Asp55Glu, Ile141Leu, Ala144Ser, Gln164His, Ile165Met, Gly168Pro, Ser169Thr, Gln188Arg, and Arg189Gln.

3.2. Binding site analysis of SARS-CoV-2 M pro as the drug target

The X-ray crystallographic structure of SARS-CoV-2 M pro (Fig. 2A) showed 3 characteristic domains I, II and III like SARS CoV M pro. The active site is made up of His41 and Cys145 dyad, consistent with previous reports of SARS CoV M pro [16,41]. Druggable binding pockets were predicted by CASTp 3.0. The largest pocket was with a solvent accessible area of 351.125 Å³ and a volume of 319.370 Å³. The second pocket had a 104.616 Å³ area and a 69.549 Å³ volume. All the other pockets are less than 100 Å³ volume and area. The high-volume pocket (Fig. 2B) is made up of Thr24, Thr25, Thr26, Leu27, His41, Cys44, Thr45, Ser46, Met49, Pro52, Tyr54, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Gln166, Leu167, Pro168, His172, Asp187, Arg188, Gln189, Thr190 and Gln192.

3.3. Screening of small-molecule library by molecular docking

The top 1000 viral protease inhibitor-like molecules identified by Ton et al. [24] were obtained in SDF format, cleaned three-dimensionally, hydrogenized and used as small-molecule library for screening. SARS-CoV-2 M pro structure (PDB ID: 6LU7) was prepared by removing water and ligand molecules. Two lead compounds Lead1-ZINC000621285959 and Lead2-ZINC000621285995 showed maximum binding affinities of −9.3 kcal/mol and −9.1 kcal/mol towards the SARS-CoV-2 M pro active site, respectively, which are far better than the positive controls Lopinavir (−6.8 kcal/mol) and Nelfinavir (−7.9 kcal/mol) (Table 1). Binding of three drugs viz. lopinavir, oseltamivir, and ritonavir simultaneously with the protein resulted in a binding affinity of −8.32 kcal/mol [15]. Three molecules of natural origin from Moroccan medicinal plants Crocin, Digitoxigenin and β-Eudesmol were docked with SARS-CoV-2 M pro and showed an interaction energy equal to −8.2 kcal/mol, −7.2 kcal/mol and −7.1 kcal/mol, respectively. Both lead1 and lead2 were found to surpass these previously reported binding energies.

Molecular docking of SARS-CoV-2 M pro with lead1 and lead2 are depicted in Fig. 3. Lead1 binds to the active site formed by Domain-I

![Fig. 3. Molecular docking of SARS-CoV-2 M pro with Lead compounds. (A) The binding of Lead1 is in the groove between Domain-I and Domain-II chymotrypsin-like β barrel, where the active site is located and binds exactly with active site dyads His41 (Red), Cys145 (Blue) and Mpro dimerization amino acid Glu166 (Cyan). (B) Binding of Lead 2 with SARS-CoV-2 M pro at the same active site.](image-url)
and Domain-II chymotrypsin-like β barrels, where the active site dyad His41 and Cys145 is located. Khan et al. proposed 5 inhibitors, all of which exhibited significant interactions with the same active site dyads [51]. The binding of lead1 was found to be stabilized by various hydrogen bonds and alkyl bonds. His41, Met49 and Met165 showed a stronger tendency to form alkyl bonds; on the other hand, Phe140, Leu141, Ser144, His164, and Met165 involved in critical bond formation. Some of these amino acids were involved in non-covalent interactions of a best five compounds interacted with either both (Cys145 and His41) or at least one catalytic residue [52]. Zhang et al. demonstrated that dimerization of SARS-CoV-2 M\textsubscript{pro} is crucial for catalytic activity because the N-finger of SARS-CoV-2 M\textsubscript{pro} is crucial for catalytic activity because the N-finger hydrogen bond acceptors/donors, topological polar surface area, lipophilicity and solubility were calculated.

### Table 2: ADME/Tox properties of lead compounds and positive controls.

| Property                                | Lead1-ZINC000621278586 | Lead2-ZINC000621285995 | Positive control-1-Lopinavir | Positive control-2-Nelfinavir |
|-----------------------------------------|-------------------------|-------------------------|-------------------------------|-------------------------------|
| Gastrointestinal Absorption             | High                    | High                    | High                          | High                          |
| Blood Brain Barrier Permeation          | No                      | No                      | No                            | No                            |
| p-Glycophosphate Substrate              | Yes                     | Yes                     | Yes                           | Yes                           |
| Cytochrome P450 Inhibitor\(^\text{a}\)   | No                      | CYP1A2                  | CYP2C19, CYP3A4               | CYP2C19, CYP3A4               |
| Skin permeation                         | −7.10                   | −7.24                   | −5.93                         | −5.74                         |
| Log \(K_{d}\) (cm/s)                   |                         |                         |                               |                               |
| Rule Based Druglikeness                 | Yes                     | Yes                     | Violations except Egan        | Violations except Egan        |
| Medicinal chemistry                     |                         |                         |                               |                               |
| PAINS\(^{b}\)                           | 0 alert                 | 0 alert                 | 0 alert                       | 0 alert                       |
| Brenk\(^{c}\)                          | 0 alert                 | 0 alert                 | 0 alert                       | 0 alert                       |
| Synthetic accessibility score\(^{d}\)   | 3.58                    | 3.46                    | 5.67                          | 5.58                          |
| Toxicity and carcinogenesis             |                         |                         |                               |                               |
| Drug-induced liver injury (Probability Value) | 0.5500           | 0.5500                  | 0.7000                        | 0.5250                        |
| Acute oral toxicity LD50 mol / kg       | 2.034                   | 2.371                   | 3.196                         | 3.004                         |
| Ames mutagenesis (Probability Value)    | −0.5200                 | −0.5500                 | −0.8300                       | −0.6900                       |
| Carcinogenesis                          | Non-Carcinogenic        | Non-Carcinogenic        | Non-Carcinogenic              | Non-Carcinogenic              |

\(^{a}\)Cytochrome P450 Inhibitors include inhibitors of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4; all the molecules showed a bioavailability score of 0.55; \(^{b}\) Pan assay interference compounds alert; \(^{c}\) 105 fragments identified by Brenk database; \(^{d}\) Synthetic accessibility score on a scale of 1–10 (1 easy to 10 difficult to synthesize).

The pharmacokinetics predictions (Table 2) showed that lead1 and lead2 were non-permeators of the blood brain barrier and skin with high gastrointestinal absorption. Lead1 was predicted to inhibit none of the Cytochrome P450 viz., CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4. Lead2 was inhibitory to CYP1A2 alone. On the other hand, both the positive controls Lopinavir and Nelfinavir were predicted to inhibit CYP2C19 and CYP3A4. All the 12 molecules showed a bioavailability score of 0.55. Hence, lead1 can be considered as a potent drug candidate followed by lead2. ADME/Tox properties of all other lead compounds are provided in supplementary tables (S-Table 1–4).

Druglikeness screening showed that all lead compounds satisfied all the druglikeness rules viz., Lipinski [54], Ghose [55], Veber [56], Egan [57] and Muegge [58], except lead10, which showed violations in the Veber and Egan rules owing to its high TPSA values. On the other hand, the positive controls showed at least one violation in all the rules except the Egan rule. Medicinal chemistry analysis showed that all the leads were passed by filters for removal of pan assay interference compounds (PAINS) [59] and a list of 105 fragments identified by Brenk et al. [60], except lead4, which showed hydantoin alert. The synthetic accessibility scores of Lopinavir and Nelfinavir were 5.67 and 5.58, respectively, while those of lead1 and lead2 were 3.58 and 3.46, which shows that these leads could be easily synthesized compared to positive controls.

Drug-induced liver injury probability values of lead1 and lead2 were found to be lower than that of Lopinavir; and the acute oral toxicity LD50 values of lead1 and lead2 were predicted to be 2.034 mol/kg and 2.371 mol/kg, respectively. Both the leads exhibited negative Ames mutagenesis probability scores and were found to be non-carcinogenic. Since, both lead1 and lead2 show exceptional drug-like properties with good medicinal chemistry properties, they can further be assessed for their in vitro SARS-CoV-2 M\textsubscript{pro} inhibitory activities. However, lead1 (ZINC000621278586) could be better than lead2 (ZINC000621285995) because of its non-inhibitory nature of cytochrome P450, while lead2 inhibits CYP1A2. Apart from this minor negative characteristic, lead2 was also found to be a good SARS-CoV-2 M\textsubscript{pro} inhibitor.

#### 3.4. ADME/Tox evaluation of lead compounds

Absorption, distribution, metabolism, excretion and toxicity predictions were carried out for all 10 lead compounds and positive controls. The physiochemical properties like molecular weight, number of hydrogen bond acceptors/donors, topological polar surface area, lipophilicity and solubility were calculated.

#### 3.5. Molecular dynamics (MD) simulations

Based on the docking scores and absorption, distribution, metabolism, excretion and toxicity predictions Lead1- ZINC00062127858, Lead2- ZINC000621285995 and Lead3-ZINC000566550443 were selected and their complexes with SARS-CoV-2 M\textsubscript{pro} were subjected to MD simulations along with complexes of SARS-CoV-2 M\textsubscript{pro} - Lopinavir and

### References

[51] K. Kavitha, et al. (2020). Biophysical Chemistry 267 (2020) 106478
Nelfinavir positive controls. The best docking conformation of each of the complexes was chosen and used as the starting point for a 20 ns simulation. The trajectories were analyzed for stability and the complete details of conformations of proteins were observed to reconfirm the results of docking. Furthermore, the time-dependent RMSD values of atoms in the unliganded SARS-CoV-2-M^apo^, SARS-CoV-2-M^apo^-Lead1 complex, SARS-CoV-2-M^apo^-Lead2 complex, SARS-CoV-2-M^apo^-Lead3 complex, SARS-CoV-2-M^apo^ - Lopinavir and SARS-CoV-2-M^apo^ - Nelfinavir complexes were plotted (Fig 4A). The complexes of all lead compounds and Lopinavir were well correlated with unliganded protein with only a few atomic fluctuations in the magnitude. Overall mean RMSD values for SARS-CoV-2-M^apo^ apo protein and SARS-CoV-2-M^apo^ complexes with Lead1, Lead2, Lead3, Lopinavir and Nelfinavir were found to be 0.28 ± 0.034 nm 0.20 ± 0.025 nm, 0.23 ± 0.039 nm, 0.20 ± 0.024 nm 0.21 ± 0.025 nm and 0.26 ± 0.060 nm, respectively. The RMSD of lead1, lead2 and lead3 complexes were less than 0.23 nm, while overall RMSD of all the complexes showed consistency within 0.3 nm over the entire trajectory, which is well within the range of previous reports [53].

Similarly, the backbone radiation of gyration (Rg) values (Fig 4B) for SARS-CoV-2-M^apo^ apo protein was found to be 2.195 ± 0.016 nm and that of SARS-CoV-2-M^apo^ complexes with Lead1, Lead2, Lead3, Lopinavir and Nelfinavir were found to be 2.218 ± 0.014 nm, 2.203 ± 0.016 nm, 2.222 ± 0.017 nm 2.212 ± 0.018 nm and 2.186 ± 0.025 nm, respectively. Rg value of 2.2 nm for all lead compounds showed that the binding of these ligands does not cause considerable stress on the backbone of SARS-CoV-2-M^apo^. The data revealed that all the systems were compact throughout the simulation, which indicates that the systems are well converged.

RMSF was calculated for SARS-CoV-2-M^apo^ apo protein and SARS-CoV-2-M^apo^ complexes with Lead1, Lead2, Lead3, Lopinavir and Nelfinavir (Fig. 5). Fluctuations were generally observed in loops, coils and at amino and carboxy terminals of the protein chain. The mean RMSF values were found to be 0.12 ± 0.052 nm, 0.12 ± 0.051 nm,
0.12 ± 0.054 nm, 0.13 ± 0.059 nm and 0.14 ± 0.055 nm for apo protein and its complex Lead1, Lead2, Lead3, Lopinavir and Nelfinavir, respectively. Only the Lopinavir complex alone showed high mean fluctuation when compared to all other complexes. Lead compounds showed fluctuations around amino acid 45 and between residues 140–170, which correlates well with the active site of SARS-CoV-2-Mpro.

Number of hydrogen bonds between target protein and ligands were calculated throughout the trajectory to predict ligand stabilization. SARS-CoV-2-Mpro-Lead1, SARS-CoV-2-Mpro-Lead2 and SARS-CoV-2-Mpro-Lead3 complexes showed 4, 2 and 2 hydrogen bonds until the end, respectively (Fig. 6). On the other hand, the SARS-CoV-2-Mpro-Lopinavir complex showed 1 stable hydrogen bond and the SARS-CoV-2-Mpro-Nelfinavir complex lost the initial hydrogen bonds before 10,000 ps. Hence, lead1, lead2 and lead3 were shown to have better binding capabilities than both the positive controls. The lead1 exhibited a maximum interaction of four out of six hydrogen bonds intact throughout the 20 ns simulation, viz. Gly-143 (3.21 Å), Ser-144 (2.18 Å), Cys145 (2.53 Å) and His172 (2.74 Å). Glu166 was shown to form a salt bridge at 4.57 Å distance with lead1. Lead2 exhibited a consistent hydrogen bond with Phe-140 (2.45 Å), salt bridge with Glu-166 (4.30 Å) and hydrophobic interaction with His-41 (3.79 Å). Lead3 showed hydrogen bonds with Phe-140 (1.81 Å) and His-164 (2.07 Å) and a salt bridge with Glu-166 (1.91 Å).

3.6. Thermodynamic studies

The binding energies for all protein-ligand complexes were calculated for the last 10 ns of MD trajectories. All 5 complexes showed negative binding energies (Table 3) indicating that all the complexes were energetically stable. Lead1 showed the lowest binding energy (−129.266 ± 2.428 kJ/mol) of the lead molecules. The binding energy of lead 1 is lower than Lopinavir (−29.410 ± 9.493) and higher than Nelfinavir (−140.785 ± 3.989) and was considered as the most stable lead molecule. Lead2 and lead3 showed binding energies of −116.478 ± 3.502 and −96.864 ± 3.820, respectively, which is better than that of Lopinavir. The Lopinavir complex showed a less favourable energy value of −29.410 ± 9.493 kJ/mol. The rigorous bootstrapping used in this study resulted in a reduced standard deviation. A previous study calculated ΔG binding energy for Remdesivir, calculated throughout the trajectory to predict ligand stabilization. SARS-CoV-2-Mpro-Lead1, SARS-CoV-2-Mpro-Lead2 and SARS-CoV-2-Mpro-Lead3 complexes showed 4, 2 and 2 hydrogen bonds until the end, respectively (Fig. 6). On the other hand, the SARS-CoV-2-Mpro-Lopinavir complex showed 1 stable hydrogen bond and the SARS-CoV-2-Mpro-Nelfinavir complex lost the initial hydrogen bonds before 10,000 ps. Hence, lead1, lead2 and lead3 were shown to have better binding capabilities than both the positive controls. The lead1 exhibited a maximum interaction of four out of six hydrogen bonds intact throughout the 20 ns simulation, viz. Gly-143 (3.21 Å), Ser-144 (2.18 Å), Cys145 (2.53 Å) and His172 (2.74 Å). Glu166 was shown to form a salt bridge at 4.57 Å distance with lead1. Lead2 exhibited a consistent hydrogen bond with Phe-140 (2.45 Å), salt bridge with Glu-166 (4.30 Å) and hydrophobic interaction with His-41 (3.79 Å). Lead3 showed hydrogen bonds with Phe-140 (1.81 Å) and His-164 (2.07 Å) and a salt bridge with Glu-166 (1.91 Å).

### Table 3

| Compound          | Van der Waals energy (kJ/mol) | Electrostatic energy (kJ/mol) | Polar salvation energy (kJ/mol) | SASA energy (kJ/mol) | Binding energy (kJ/mol) |
|-------------------|-------------------------------|-------------------------------|--------------------------------|----------------------|------------------------|
| Mpro-Lead1        | −161.521 ± 2.101              | −91.803 ± 3.518               | 139.687 ± 2.661                | −15.681 ± 0.306      | −129.266 ± 2.428       |
| Mpro-Lead2        | −162.605 ± 3.262              | −22.705 ± 4.998               | 85.050 ± 4.682                 | −16.003 ± 0.342      | −116.478 ± 3.502       |
| Mpro-Lead3        | −145.29 ± 1.942               | −25.349 ± 3.256               | 88.581 ± 2.140                 | −15.030 ± 0.208      | −96.864 ± 3.820        |
| Mpro-Lopinavir    | −63.502 ± 2.962               | −7.886 ± 1.559                | 49.413 ± 10.392                | −7.625 ± 0.362       | −29.410 ± 9.493        |
| Mpro-Nelfinavir   | −196.671 ± 2.974              | −32.067 ± 4.236               | 109.637 ± 3.147                | −21.459 ± 0.448      | −140.785 ± 3.989       |

Fig. 5. Root-Mean-Square Fluctuation (RMSF) of unliganded SARS-CoV-2-Mpro (Black), SARS-CoV-2-Mpro-Lead1 complex (Red) and SARS-CoV-2-Mpro-Lead2 complex (Magenta), SARS-CoV-2-Mpro-Lead3 complex (Blue), SARS-CoV-2-Mpro-Lopinavir complex (Green) and SARS-CoV-2-Mpro-Nelfinavir complex (Cyan) in nm plotted against the number of amino acid residues.

Fig. 6. Number of hydrogen bond interactions during simulation between protein and ligand complexes of SARS-CoV-2-Mpro-Lead1 complex (Red) and SARS-CoV-2-Mpro-Lead2 complex (Magenta), SARS-CoV-2-Mpro-Lead3 complex (Blue), SARS-CoV-2-Mpro-Lopinavir complex (Green) and SARS-CoV-2-Mpro-Nelfinavir complex (Cyan) plotted against time (ps).
Saquinavir, Darunavir, Nat-1 and Syn-16 with target protein Chymotrypsin-like protease (3CLpro) as $-45.5240$, $-36.3026$, $-48.1041$, $-41.2565$ and $-31.5581$ kJ/mol, respectively, and proposed Darunavir as the best protease inhibitor [51]. Another study reported $-4.62$ kCal/mol or $-19.33$ kJ/mol for the Mpro-ZINC000015988935 complex [61]. The lead compounds of this study, particularly lead1 and lead2, exhibited far better binding energies and hence can be expected to outperform these previously reported drugs and lead compounds.

Energy decomposition plot was calculated as the energy contribution of each residue (Fig. 7). All the lead compounds showed stabilization of the complex around residue number 40 and between residues 140 to 170, indicating that they bind to the active site of SARS-CoV-2-MPro. Lopinavir showed binding in the same region with lower affinity. Nelfinavir has been shown to bind in a different region between residues 250 and 300, which makes its usability as an inhibitor questionable. The thermodynamic calculations showed that the binding of both lead1 and lead2 to the active site of SARS-CoV-2-MPro is energetically favored followed by lead3. Hence, they can act as good inhibitors of SARS-CoV-2-MPro.

4. Conclusions and future perspective

Structure-based sequence alignment of SARS-CoV-2 and SARS-CoV main proteases showed extensive similarities in their secondary and tertiary structures. Hence, it can be construed that in silico molecular approaches used for screening SARS-CoV-2 MPro inhibitors can also be used for finding potent SARS-CoV-2 MPro inhibitors. This study screened 1000 protease-inhibitor-like molecules against SARS-CoV-2-MPro and proposes lead compounds viz., Lead1–2-amino-5-(((5R)-5-methyl-2,3,4,5-tetrahydro-1H-1-benzazepin-1-yl)methyl)-1H,7H-[1,2,4]triazolo[1,5-a]pyrimidin-7-one (ZINC000621278586) and Lead2–2-amino-5-((1′,2′-dihydrospiro[cyclobutane-1,3′-indol]-1′-yl)methyl)-1H,7H-[1,2,4]triazolo[1,5-a]pyrimidin-7-one (ZINC600621285995) as potent SARS-CoV-2 MPro inhibitors with better binding properties and less toxicity than existing protease inhibitors like Lopinavir and Nelfinavir. Both these molecules are [1,2,4]triazolo[1,5-a]pyrimidin-7-one compounds and their antiviral properties have not been reported previously. MD simulation studies showed that both lead compounds had high binding affinity towards SARS-CoV-2-MPro. Binding free energy calculations by MM/PBSA showed energetically stable negative values of $-129.266 \pm 2.428$ kJ/mol and $-116.478 \pm 3.502$ kJ/mol for lead1 and lead2, respectively. Taken all together, according to docking studies, physicochemical characterizations, ADME/Tox predictions and molecular dynamics studies, it is safer to conclude that these pyrimidin-7-one lead compounds could be considered as possible SARS-CoV-2 MPro inhibitors. However, the inhibitory activity of these lead compounds should be further tested in vitro and animal studies. Future perspective of this study could be designing a covalent inhibitor based on these pyrimidin-7-one compounds that could form favourable covalent bond with the reactive thiol group of active site Cys145.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgement

The authors express immense gratitude to the Google Cloud Platform for supporting this work with COVID HCLS Research Credit 46225154. We are grateful to Dr.K.R. Venkatesan, Principal and the Management of Sri Sankara Arts and Science College, for providing the facilities and knowledge resources to conduct this research. We are thankful to Mr.B.Rajganesh, Enstoa India Private Limited, Bangalore, India, for providing computational resources for docking studies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bpc.2020.106478.

References

[1] UN-DESA, Global economy could shrink by almost 1% in 2020 due to COVID-19 pandemic: United Nations, Econ. Time (2020), https://economictimes.indiatimes.com/news/international/business/global-economy-could-shrink-by-almost-1-in-2020-due-to-covid-19-pandemic-united-nations/articleshow/74943235.cms

[2] M.K. Gupta, S. Vermula, R. Donde, G. Gouda, L. Behera, R. Vadde, In-silico perspective, J. Biomol. Struct. Dyn. (2020) 1–9, https://doi.org/10.1080/07391102.2020.1751300

[3] A. Elfiky, SARS-CoV-2 RNA dependent RNA polymerase (RdRp) targeting: an in silico approach to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel, J. Biomol. Struct. Dyn. (2020) 1–11, https://doi.org/10.1080/07391102.2020.1751300

[4] H.M. Wahedi, S. Ahmad, S.W. Abbasi, Stilbene-based natural compounds as promising drug candidates against COVID-19, J. Biomol. Struct. Dyn. (2020) 1–10, https://doi.org/10.1080/07391102.2020.1761682

[5] P. Gautret, J.-C. Lagier, P. Parola, V.T. Haung, L. Meddeb, M. Mailhe, B. Doudier, J. Courjon, V. Giordanengo, V.E. Vieira, H.T. Dupont, S. Honoré, P. Colson, E. Chabrière, B. La Scola, J.-M. Rolain, P. Brouqui, D. Raoul, Hydroxychloroquine
A.K. Ghose, V.N. Viswanadhan, J.J. Wendoloski, A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases, J. Comb. Chem. 1 (1999) 55–68, https://doi.org/10.1021/cc9800071.

D.F. Veber, S.R. Johnson, H.-Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the Oral bioavailability of drug candidates, J. Med. Chem. 45 (2002) 2615–2623, https://doi.org/10.1021/jm020017n.

W.J. Egan, K.M.J. Merz, J.J. Baldwin, Prediction of drug absorption using multivariate statistics, J. Med. Chem. 43 (2000) 3867–3877, https://doi.org/10.1021/jm000292c.

I. Muegge, S.L. Heald, D. Brittelli, Simple selection criteria for drug-like chemical matter, J. Med. Chem. 44 (2001) 1841–1846, https://doi.org/10.1021/jm015507c.

J.B. Baell, G.A. Holloway, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, J. Med. Chem. 53 (2010) 2719–2740, https://doi.org/10.1021/jm9011371.

R. Binks, A. Schipani, D. James, A. Krasowski, I.H. Gilbert, J. Freamson, P.G. Wyatt, Lessons learnt from assembling screening libraries for drug discovery for neglected diseases, ChemMedChem. 3 (2008) 435–444, https://doi.org/10.1002/cmdc.200700139.

A.D. Elmezayen, A. Al-Obaidi, A.T. Şahin, K. Yeşilçat, Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes, J. Biomol. Struct. Dyn. (2020) 1–13, https://doi.org/10.1080/07391102.2020.1758791.