Computers in Biology and Medicine 141 (2022) 105049

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Combined deep learning and molecular docking simulations approach identifies potentially effective FDA approved drugs for repurposing against SARS-CoV-2

Muhammad U. Anwaar [a,1], Farjad Adnan [b,1], Asma Abro [c,1], Rayyan A. Khan [a,1], Asad U. Rehman [d,e], Muhammad Osama [d,e], Christopher Rainville [f], Suresh Kumar [f], David E. Sterner [f], Saad Javed [d,e], Syed B. Jamal [g], Ahmadullah Baig [d], Muhammad R. Shabbir [d,e], Waseh Ahsan [d], Tauseef R. Butt [f], Muhammad Z. Assir [d,e,h,*]

[a] Department of Electrical and Computer Engineering, Technical University Munich, Arcisstraße 21, 80333, München, Germany
[b] Paderborn University, Warburger Str. 100, 33098, Paderborn, Germany
[c] Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, 1800, Pakistan
[d] Department of Medicine, Allama Iqbal Medical College, University of Health Sciences, Lahore, 54550, Pakistan
[e] Center for Undiagnosed, Rare and Emerging Diseases, Lahore, 54550, Pakistan
[f] Progenra Inc, 271A Great Valley Parkway, Malvern, PA, 19355, USA
[g] Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan
[h] Department of Molecular Biology, Shaheed Zulfiqar Ali Bhutto Medical University, Islamabad, 44000, Pakistan

ARTICLE INFO

Keywords:
SARS-CoV-2
Drug repurposing
Machine learning
Docking
Binding affinity

ABSTRACT

The ongoing pandemic of Coronavirus Disease 2019 (COVID-19) has posed a serious threat to global public health. Drug repurposing is a time-efficient approach to finding effective drugs against SARS-CoV-2 in this emergency. Here, we present a robust experimental design combining deep learning with molecular docking experiments to identify the most promising candidates from the list of FDA-approved drugs that can be repurposed to treat COVID-19. We have employed a deep learning-based Drug Target Interaction (DTI) model, called DeepDTA, with few improvements to predict drug-protein binding affinities, represented as KIBA scores, for 2440 FDA-approved and 8168 investigational drugs against 24 SARS-CoV-2 viral proteins. FDA-approved drugs with the highest KIBA scores were selected for molecular docking simulations. We ran around 50,000 docking simulations for 168 selected drugs against 285 total predicted and/or experimentally proven active sites of all 24 SARS-CoV-2 viral proteins. A list of 49 most promising FDA-approved drugs with the best consensus KIBA scores and binding affinity values against selected SARS-CoV-2 viral proteins was generated. Most importantly, 16 drugs including anidulafungin, velpatasvir, glecaprevir, rifapentine, flavin adenine dinucleotide (FAD), terlipressin, and selinexor demonstrated the highest predicted inhibitory potential against key SARS-CoV-2 viral proteins. We further measured the inhibitory activity of 5 compounds (rifapentine, velpatasvir, glecaprevir, anidulafungin, and FAD disodium) on SARS-CoV-2 PLpro using Ubiquitin-Rhodamine 110 Gly fluorescent intensity assay. The highest inhibition of PLpro activity was seen with rifapentine (IC50: 15.18 μM) and FAD disodium (IC50: 12.39 μM), the drugs with high predicted KIBA scores and binding affinities.

1. Introduction

Coronavirus Disease 2019 (COVID-19) resulting from severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has become a global pandemic with severe effects on global public health.[1] As of 10th November 2021, more than 251 million confirmed cases of COVID-19 have been reported with an estimated 5 million deaths globally [1]. Despite the availability of multiple different vaccines against SARS-CoV-2, there is still a pressing urgency to make an expedited

* Corresponding author. Department of Medicine, Medical Unit 3, Allama Iqbal Medical College/ University of Health Sciences Lahore, Pakistan.
E-mail address: dr.zamankhan@gmail.com (M.Z. Assir).
[1] These authors contributed equally.

https://doi.org/10.1016/j.compbiomed.2021.105049
Received 5 July 2021; Received in revised form 11 November 2021; Accepted 15 November 2021
Available online 20 November 2021

List of abbreviations

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
COVID19	Novel 2019 Coronavirus Disease
SARS-CoV	Sever acute respiratory syndrome coronavirus
MERS-CoV	Middle East respiratory syndrome coronavirus
FDA	Food and Drug Administration
PDB	The Protein Data Bank
DAVIS	Densely Annotated VIdeo Segmentation
CNN	Convolutional neural network
DeepDTA	Deep Drug-Target Binding Affinity Prediction
CPU	Central Processing Unit
Ki	Inhibitory constant
Kd	Dissociation constant
IC50	Half maximal inhibitory concentration
Bi-LSTM	Bi-directional long short term memory
BPE	Byte pair encoding
ReLU	Rectified Linear Unit
MSE	Mean Squared Error
CI	Concordance index
C–I-TASSER	Contact-guided Iterative Threading ASSEmbly Refinement
S	Spike
E	Envelope
M	Membrane
N	Nucleocapsid
RNA	Ribonucleic acid
DNA	Deoxyribonucleic acid
Nsp1	Non-structural protein 1
Nsp7	Non-structural protein 7
Nsp8	Non-structural protein 8
Nsp12	Non-structural protein 12
HCV	Hepatitis C virus
HBV	Hepatitis B virus
FAD	Flavin Adenine Dinucleotide
DPP-4	Dipeptidyl peptidase 4
ACE	Angiotensin converting enzyme
HMG-CoA	β-Hydroxy β-methylglutaryl-CoA
RMSD	Root-mean-square deviation
PLpro	Papain-like protease
RdRp	RNA-directed RNA polymerase
Hel	Helicase
CPK	Corey-Pauling-Koltun
RCSB	Research Collaboratory for Structural Bioinformatics
XPO1	Exportin-1
FMN	Flavin mononucleotide

discovery of effective drugs for the treatment of COVID-19.

De-novo drug development is a time-consuming, complex, and expensive process that typically costs billions of dollars [2]. Drug repurposing, a process of investigating approved or investigational drugs for new therapeutic purposes, offers a cost and time-effective alternative approach [3]. Drug repurposing is based on a paradigm shift in our understanding that many effective drugs act via modulation of multiple proteins rather than single targets [4,5]. Both experimental and computational approaches are being employed in drug repurposing with computational approaches becoming increasingly popular, robust, and efficient [3].

Recent advancements in the field of deep learning have significantly improved the computational approaches for drug repurposing. Conventional machine learning methods are limited by the lack of ability to process the data in the raw form and hence depend on feature engineering [6] for machine learning and pattern recognition [7]. Deep learning, a novel machine learning approach utilizing deep neural networks based techniques, overcomes the issue of manual feature engineering and processes the crude multi-dimensional data (images, text sequences, etc) in their layers for algorithm training [8]. The architecture of the deep learning model is a multi-layer cascade of several modules, mapping a nonlinear relationship between input and output employing a backpropagation method to fine-tune the corresponding weights [9]. Deep learning is also being used to study Drug-Target interaction [10].

The approach of supervised deep learning can be used to predict drug-target binding affinities. In this stratagem, a deep learning model is trained on the experimentally available binding affinities of several protein-ligand complexes. Several available benchmark protein-ligand complex datasets, providing the experimental binding affinities, are PDB [11], DAVIS [12], and KIBA [13,14] datasets. These have been used in several studies to predict the binding affinities of the complexes. The training of the deep learning model is realized using physical or structural features of protein-ligand complexes on the training dataset. A properly trained model can be employed to predict the binding affinities of unseen protein-ligand complexes. A binary prediction classifier has already been proposed [15,16] which takes into account the several input representations of the protein-ligand complex. In several other deep learning network studies, a protein-ligand interaction scoring has been predicted, training the convolutional neural network (CNN) using the three-dimensional structure of the associated complexes [17,18]. The main source of these complexes is PDB [11]; however, the small number of interactions is a limitation as only 25000 protein-ligand complexes have been documented. DeepDTA [19] is another novel approach that uses SMILES and FASTA sequences of ligands and proteins, respectively. In the DeepDTA approach, two separate CNN blocks have been employed to train the protein and ligand sequences and eventually combined in a fully connected layer called DeepDTA. In this study, we have employed DeepDTA with few improvements.

Another computational approach that is employed to study drug-target interaction is virtual screening. Virtual screening can be performed by using *molecular docking* - a technique that samples the ensemble of binding modes to predict preferred pose(s) in which the ligand can bind with the receptor at a certain location, known as active/binding site, to form a stable complex [20,21]. A binding mode refers to a unique conformation along with orientation and translation of the ligand. Ranking of the preferred binding modes/poses is carried out by evaluating a mathematical function, known as scoring function, that quantifies the stability of the complex formed by a particular pose of the ligand with the receptor [22,23]. In our work, docking was performed using actual simulation as it is more realistic compared to its alternatives [23,24]. We have used the results of docking simulations to get docking scores of potential candidate drugs with COVID19 viral proteins. To reduce the time required for around 50,000 docking simulations, we made use of a recently published open-source high throughput screening platform virtualflow [25] to parallelize the docking scenarios across multiple machines and CPU cores.

In this study, we have employed a multidisciplinary, multimodal approach combining deep machine learning and large-scale molecular docking experiments sequentially to identify FDA-approved drugs that can be used as an effective treatment against SARS-CoV-2. We have then studied the inhibitory activity of selected drugs against the PLpro enzyme of SARS-CoV-2 to experimentally validate some of the findings of our *in-silico* analysis.

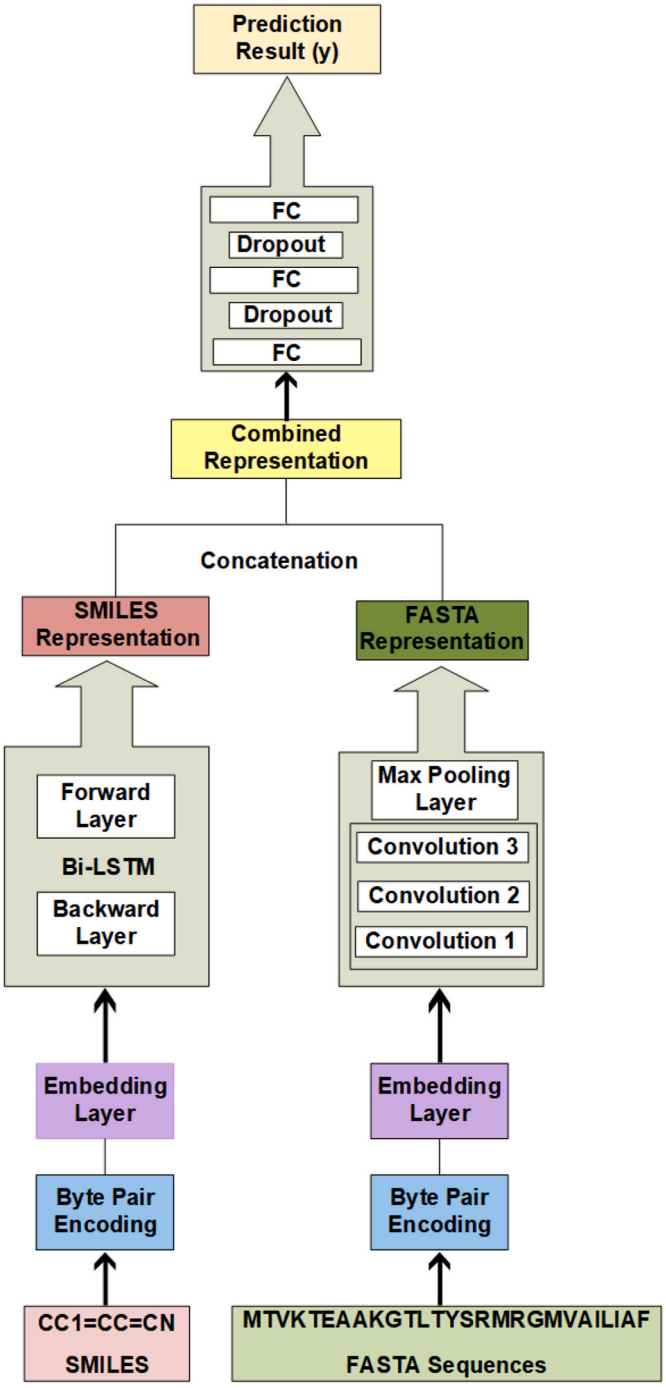


Fig. 1. Block Diagram of Modified DeepDTA model with Bi-LSTM and CNN blocks to learn from sequences - The proposed Bi-LSTM and CNN methodology is represented. It consists of two separate BI-LSTM and CNN blocks, for training the representations of ligands and proteins respectively. The output representations from both blocks have been concatenated and fed to the fully connected layers, which eventually predict the drug.-protein binding affinities at their output.

2. Methods

2.1. Deep learning model overview

In our work, we have employed a deep learning-based DTI-model, called DeepDTA[20], with few improvements. Our deep learning model automatically incorporates useful and required features from raw ligands and proteins into the model to predict the drug and protein interactions. We have utilized 1-D sequences of ligands (SMILES)[28] and proteins (FASTA)[29] to train our model. SMILES (Simplified Molecular Input Line Entry System) representation of molecules has been exploited rather than physical and chemical properties associated with the ligands. We have employed the Bi-LSTM[30] blocks of neural networks instead of CNN used by DeepDTA[20] to learn the SMILE representation of molecules whereas fully connected CNN has been engaged to learn the FASTA representation of proteins.

2.2. Dataset

We have used a benchmark KIBA dataset for training our model and prediction evaluation of binding affinities, which has been tapped previously in a handful of studies. The KIBA dataset comprises selectivity assays of the kinase proteins family and the associated inhibitors[15]. It predominantly embodies the corresponding KIBA scores. The KIBA values have been contrived by combining K_i, K_d, and IC_{50} values obtained from several sources. This Dataset has been created from the original 52,498 drugs[14], which has been filtered to 2111 unique drugs. The pool of all shortlisted drugs has at least 10 measured interactions, yielding a total of 229 proteins out of all 467 targets and a total of 118,254 interactions[20]. In our work, we have employed the same KIBA dataset used by DeepDTA[20].

2.3. Input representation

In our approach, we have employed byte pair encoding[30] to represent the SMILES of ligands and FASTA sequences of the proteins. SMILES have a varied length based upon the number of atoms and type of bonds in the drug. Since we used Bi-LSTM[31] to learn the representation of drugs, we used an end token to indicate the end of the SMILES. We utilized CNN for learning protein representation, so we set the maximum length of the FASTA sequence to 1000 characters. We truncated larger sequences and padded zeros to smaller sequences.

2.4. Proposed model

Our prediction model is based upon two sub-models. One part takes SMILES as input and the second part takes FASTA as input. SMILES are encoded via byte pair encoding (BPE) and then passed to a novel Bi-LSTM[31] learning approach for sequence analysis. It has been successfully used in recent studies for text sequence analysis. This part yields an effective representation of drugs. FASTA sequences of proteins are also encoded with BPE and then trained with Convolutional neural networks (CNN). This part learns the representation of the proteins. These separate representations are then concatenated and passed through three fully connected layers. Our CNN model greatly relies on the number and size of filters to define dependencies of the sequences and we have chosen appropriate values to fit the scenario. As the size of filters and the number of filters are varied, the performance of CNN also varies with it. All layers of networks are shown in the block diagram (Fig. 1).

The interaction pairs of the KIBA dataset have been split into training and testing datasets. As an activation function, we used Rectified Linear Unit (ReLU)[32] mathematically represented as $g(z) = \max$ [26], which has been widely used in deep learning studies[8]. A learning model is trained to minimize the difference between the expected and actual value. We have formulated it as a regression task, mean squared error (MSE) has been chosen as an appropriate loss function. The learning has been accomplished with the 70 epochs (initial training parameters set at a maximum number of 100 epochs with early stopping criteria of training or validation error more than 10% of the concordance index) and a mini-batch size of 512 is chosen to update the weights and hyperparameters of our networks. The chosen optimization algorithm to train the networks is ADAM[33] which had the default learning rate of 0.001.

2.5. Prediction of KIBA scores for FDA approved and investigational drugs

After training the model, we have predicted the KIBA scores of FDA-approved and investigational drugs. The FDA-approved and investigational drugs have been retrieved from Canadian DrugBank[34]. The DrugBank database is a handy pool of drugs that includes detailed drug information and their corresponding interactions. The DrugBank is composed of approved 3546 drugs, including 2630 small molecule drugs and 1372 approved biologics (proteins, peptides, vaccines, and allergenics), and over 9000 drugs that are either under investigation or experimentation[34]. From DrugBank, we have retrieved SMILE sequences for both FDA-approved and investigational drugs. The SMILE sequences of various drugs were not available on the DrugBank and ultimately based upon available SMILE sequences we have incorporated 2440 FDA-approved drugs whereas 8168 investigational drugs have been analyzed. The FASTA sequences of 24 viral proteins have been acquired from the published genome of SARS-CoV-2 available at C–I-TASSER [27]. Both of them, in combination, have been supplied to the trained deep learning model, which predicted all corresponding KIBA scores.

2.6. Selection of FDA approved drugs for virtual screening

FDA-approved drugs were ranked according to their KIBA scores for each viral protein. Two mutually non-exclusive subsets of drugs were prepared. Subset A comprised of drugs making it to the top 50 drugs with the highest KIBA score for any of 24 viral proteins. Subset B comprised of drugs with KIBA scores of greater than or equal to 11.5 for all 24 viral proteins. Two subsets were then combined, duplicates were removed and drugs with significant interaction with less than 3 viral proteins were removed. The resulting set of FDA-approved drugs were then subjected to Virtual Screening.

2.7. Molecular docking simulations

We made use of recently published structures of 24 viral proteins for COVID19 [27] and the three-dimensional structures of these proteins were retrieved. We performed an extensive literature search for the availability of binding site data for these viral proteins. If the structures are available and their binding site data is elucidated, it is made use in this research work. In the other case, the predicted binding sites for all the proteins are provided as determined by several binding site prediction algorithms e.g. COACH, S-Site, FINDSITE, ConCavity etc [28]. In our simulations for the sake of uniformity, we made use of the predicted binding sites given by C–I-TASSER [27] from which we have obtained our protein structures. For every site, we computed the search box mean by averaging the coordinates of residues forming the site. The size of the search box was set to be the difference of maximum and minimum along all three dimensions. Afterward, we discarded the search boxes that were contained in other search boxes. This resulted in 284 potential sites to be tested in total. In the case of published active site data, we relied on that data completely. Autodock tool was used to preprocess docking files for Autodock Vina docking algorithm [29]. The structure files for ligands were obtained from PubChem [30] and converted to pdbqt. Various tools like Chimera [31] and ChemOffice [32] were used for the refinement and proper energy minimization of the structures of ligands. The refined and best stereochemical quality structures having the suitable number of minimization steps were then docked into the active site of the target proteins using AutoDock and AutoDock Vina. The simulations were executed in parallel on eight compute nodes with eight CPU cores each. To parallelize the whole procedure, we made use of virtualflow [25], an open source platform that automates docking simulations across multiple machines in a scalable manner. Every docking scenario was run once on one CPU core with an exhaustiveness value fixed to 8 for all scenarios. VMD [33] (Visual Molecular Dynamics) and Chimera were used for the analysis of best-docked conformations and interaction of drugs with active residues.

2.8. Molecular Operating Environment

We further validated the results of our AutoDock molecular docking simulations on MOE (Molecular Operating Environment) for selected ligand-protein interactions. By removing water molecules and performing 3D protonation on the protein molecule, the structure was refined. The energy minimization of the 3D protonated protein molecule was carried out using the MOE (Molecular Operating Environment) (http://www.chemcomp.com/) software package's energy minimizing method. Gradient: 0.05, Force Field: MMFF94X + Solvation, Chiral Constraint: Current Geometry were the energy minimization parameters. The energy minimization was automatically terminated when the root mean square gradient fell below 0.05. The final product (minimized protein structure) was used to simulate molecular docking. Like the protein, the energy minimizing algorithm of the MOE tool was used to carry out energy minimization of all the ligand molecules. All the minimized ligand molecules were saved in.mol2 format and exported in a single database of.mdb file format. Finally, the prepared database of ligand molecules was subjected to molecular docking via MOE-Dock. The MOE-Dock methodology was used for molecular docking. The ligX tool within the MOE package was used to predict the ligand-protein interactions.

2.9. Measurement of IC50 of selected drugs against PLpro

SARS Cov-2 PLpro activity was measured in a fluorescence intensity assay with the use of Ubiquitin-Rhodamine Gly110 as a substrate. Experiments were performed in 384 shallow well, black polypropylene (Thermo Scientific Nunc 267461) assay plates. The assay buffer contained 20 mM TRIS (pH8), 0.05% CHAPS, and 2 mM Beta Mercaptoethanol. PLpro at a concentration of 100 nM was preincubated with compounds at a defined concentration for 20 min. Later, 100 nM Ubiquitin-Rhodamine 110 was added to initiate the DUB cleavage assay. Fluorescence intensities were monitored every 5 min interval at excitation 485 nm and emission 535 nm. Increased fluorescence intensities were recorded as a result of enzyme-catalyzed cleavage of quenched Rhodamine 110-Glycine. Inhibitory activity of 5 compounds (rifapentine, glecaprevir, and FAD disodium with high predicted binding affinities; velpatasvir with intermediate binding affinity; and Anidulafungin with low binding affinity) on SARS-CoV-2 PLpro was assessed in these assays.

3. Results and discussion

3.1. Modified DeepDTA identified key FDA approved drugs targeting SARS-CoV-2 viral proteins

Using the DeepDTA based model, we predicted drug-protein binding affinities, represented as KIBA scores, for 2440 FDA approved and 8168 investigational drugs against 24 SARS-CoV-2 viral proteins, yielding KIBA scores for a total of 254,592 interactions (Figure S1 and Table S1). These proteins include four structural proteins; spike (S), envelope (E), membrane (M), and nucleocapsid (N) [34]. N protein forms the capsid that protects the viral RNA while E, M, and S proteins make the outer coat of the virus that surrounds the capsid [27,34]. Spike protein plays a crucial role in viral attachment, entry, and fusion into the target host cell [35–37]. Two essential proteins constituting the viral replication-transcription complex are helicase and non-structural protein 12 (nsp12) [38–40]. Main protease (M^{pro}, also called $3CL^{pro}$) [41,42] and papain-like protease(s) process the huge polyproteins encoded by the SARS-CoV-2 genome. These proteins are key targets for an effective antiviral therapy [38,39,41–44]. Predicted KIBA scores for all these interactions are provided in the supporting Table S1.

By applying our predefined filtering strategy (See Section 2.6 of

Table 1
FDA approved drugs with best predicted KIBA and AutoDock vina binding affinity values against selected SARS-CoV-2 Viral Proteins.

Sr #	Drug Name	PubChem Drug ID	Drug Action	Protein Name	KIBA Score	Docking Binding Affinity Values (kcal/mol)
A. Antimicrobials						
1	Anidulafungin	166548	Antifungal	RNA-directed RNA polymerase (RdRp)	11.7602	−13.4
				Helicase (Hel)	11.7768	−9.0
				Guanine-N7 methyltransferase (ExoN)	11.7008	−13.2
				Uridylate-specific endoribonuclease	11.7605	−12.3
				Surface glycoprotein (S)	11.7623	−15.3
				N	11.6237	−14.3
2	Isavuconazonium	6918606	Antifungal	Non-structural protein 2 (nsp2)	11.5459	−11.1
				N	11.6357	−10.9
3	Procaine Penicillin G	5903	Antibiotic	Proteinase 3CL-PRO	11.2148	−9.3
4	Quinupristin	5388937	Antibiotic	RNA-directed RNA polymerase (RdRp)	11.7965	−12.2
				Surface glycoprotein (S)	11.5520	−12.1
				N	11.6461	−11.8
5	Rifapentine	135659016	Antibiotic	Papain-like proteinase	11.7672	−8.6
				Surface glycoprotein (S)	11.6241	−11.4
				N	11.7584	−11.6
6	Rifabutin	135415564	Antibiotic	Non-structural protein 2 (nsp2)	11.8004	−11
				RNA-directed RNA polymerase (RdRp)	11.8209	−11.5
				Helicase (Hel)	11.8829	−10.5
				Surface glycoprotein (S)	11.7498	−12.7
				N	11.8253	−11.7
7	Polymyxin B	49800004	Antibiotic	RNA-directed RNA polymerase (RdRp)	11.8949	−14.2
				Surface glycoprotein (S)	11.8941	−12.6
				N	11.7840	−12.7
8	Cobicistat	25151504	Antiviral	Non-structural protein 2 (nsp2)	11.8630	−10.7
				RNA-directed RNA polymerase (RdRp)	11.9491	−11.4
				Guanine-N7 methyltransferase (ExoN)	11.8669	−14.1
9	Elbasvir	71661251	Antiviral	RNA-directed RNA polymerase (RdRp)	11.9396	−13.6
				Guanine-N7 methyltransferase (ExoN)	11.8797	−13.5
				Uridylate-specific endoribonuclease	11.9112	−11.7
				Surface glycoprotein (S)	11.8991	−13.7
				N	11.8131	−13.8
10	Velpatasvir	67683363	Antiviral	Non-structural protein 2 (nsp2)	11.4887	−10.8
				RNA-directed RNA polymerase (RdRp)	11.5925	−11.9
				Guanine-N7 methyltransferase (ExoN)	11.3040	−13.5
				Uridylate-specific endoribonuclease	11.5538	−11.2
				Surface glycoprotein (S)	11.7373	−14.5
				N	11.5965	−11.7
11	Pibrentasvir	58031952	Antiviral	RNA-directed RNA polymerase (RdRp)	11.8812	−12.7
				Guanine-N7 methyltransferase (ExoN)	11.7442	−12.3
				Uridylate-specific endoribonuclease	11.8894	−11.4
				Surface glycoprotein (S)	11.8692	−13.1
				N	11.7755	−12.9
12	Glecaprevir	66828839	Antiviral	Papain-like proteinase	11.7672	−9.3
				Helicase (Hel)	11.9120	−10.5
				RNA-directed RNA polymerase (RdRp)	11.9551	−11.7
				2′-O-methyltransferase (2′-O-MT)	11.8422	−11.8
				Surface glycoprotein (S)	11.7496	−11.8
				N	12.0036	−11.9
B. Hormonal						

(*continued on next page*)

Table 1 (*continued*)

Sr #	Drug Name	PubChem Drug ID	Drug Action	Protein Name	KIBA Score	Docking Binding Affinity Values (kcal/mol)
13	Abiraterone	132971	Antiandrogen	M	11.5199	−7.7
14	Amcinonide	443958	Corticosteroid	Helicase (Hel)	11.8256	−9.1
				M	11.6823	−7.5
15	Atosiban	5311010	Tocolytic	RNA-directed RNA polymerase (RdRp)	11.8447	−12.5
				Surface glycoprotein (S)	11.9298	−11.7
16	Carbetocin	16681432	Uterotonic	Helicase (Hel)	11.6491	−9.8
				Guanine-N7 methyltransferase (ExoN)	11.6623	−13.8
				N	11.5698	−12.3
17	Cortisone	5745	Corticosteroid	M	11.6792	−7.3
18	Danazol	28417	Androgen	M	11.8513	−7.2
19	Deoxycorticosterone	5952	Corticosteroid	M	11.6867	−7.6
20	Desmopressin	5311065	ADH Analog	RNA-directed RNA polymerase (RdRp)	11.9439	−13.6
				Helicase (Hel)	11.9303	−10.6
				Guanine-N7 methyltransferase (ExoN)	11.9756	−12.3
				Uridylate-specific endoribonuclease	11.9761	−11
				Surface glycoprotein (S)	12.0064	−13.1
				N	11.8459	−12.4
21	Ethynodiol diacetate	9270	Progesterone Receptor Agonist	M	11.7367	−7.5
22	Pentetreotide	72128	Octreotide Analog	RNA-directed RNA polymerase (RdRp)	12.0372	−11.9
				Guanine-N7 methyltransferase (ExoN)	11.9211	−12.9
				Uridylate-specific endoribonuclease	12.0956	−12.7
				Surface glycoprotein (S)	11.8687	−14.5
				N	11.8935	−12.4
23	Somatostatin	16129706	Octreotide Analog	RNA-directed RNA polymerase (RdRp)	11.8702	−15.3
				Helicase (Hel)	11.9113	−9.7
				Guanine-N7 methyltransferase (ExoN)	11.7610	−12
				Uridylate-specific endoribonuclease	11.8051	−12
				Surface glycoprotein (S)	11.8407	−14.4
				N	11.7700	−12.5
24	Progesterone	5994	Sex Steroid	M	11.6636	−7.3
25	Prednisone	5865	Corticosteroid	M	11.7184	−7.3
26	Terlipressin	72081	Vasopressin Analog	RNA-directed RNA polymerase (RdRp)	11.5954	−13.8
				Helicase (Hel)	11.6692	−9.8
				Uridylate-specific endoribonuclease	11.7842	−11.9
				Surface glycoprotein (S)	11.5950	−14.5
				N	11.5939	−12.4
C. Anti-neoplastic						
27	Vindesina	11643449		Guanine-N7 methyltransferase (ExoN)	11.4910	−12.8
				Surface glycoprotein (S)	11.7376	−11.6
				N	11.7234	−10.9
28	Nilotinib	644241	Tyrosine-Kinase Inhibitor	Host translation inhibitor nsp1	11.7249	−8
				Non-structural protein 2 (nsp2)	11.7242	−12.2
				Papain-like proteinase	11.7398	−6.7
				Proteinase 3CL-PRO	11.4559	−8.5
				E	11.4128	−7.5
29	Exemestane	60198	Aromatase Inhibitor	M	11.7052	−7.5
30	Etoposide	36462	Topoisomerase Inhibitor	Proteinase 3CL-PRO	11.7101	−8.6
31	Epirubicin	41867	Anthracycline antineoplastic antibiotic	Proteinase 3CL-PRO	11.5659	−8.5
32	Enasidenib	89683805	Isocitrate Dehydrogenase Inhibitor	Host translation inhibitor nsp1	11.9167	−8.5
				Non-structural protein 2 (nsp2)	11.9669	−10.7
				Proteinase 3CL-PRO	11.8091	−8.7
				Papain-like proteinase	11.8572	−6.9
33	Daunorubicin	30323	Anthracycline antineoplastic antibiotic	Proteinase 3CL-PRO	11.5225	−8.5
34	Cabazitaxel	9854073	Microtubule Inhibitors	RNA-directed RNA polymerase (RdRp)	11.9236	−12.3
				Uridylate-specific endoribonuclease	12.0602	−11.7
				N	11.9556	−11.3
35	Docetaxel	148124	Microtubule Inhibitors		11.7887	−10.1

(*continued on next page*)

Table 1 (*continued*)

Sr #	Drug Name	PubChem Drug ID	Drug Action	Protein Name	KIBA Score	Docking Binding Affinity Values (kcal/mol)
				Non-structural protein 10 (nsp10)		
				RNA-directed RNA polymerase (RdRp)	11.8086	−11.8
				N	11.7737	−11.4
36	Brigatinib	68165256	Tyrosine Kinase Inhibitor	Non-structural protein 2 (nsp2)	11.7242	−10.7
37	Selinexor	71481097	Antineoplastic	Papain-like proteinase	11.5764	−7.1
D. Vitamins						
38	Cholecalciferol	5280795	Vitamin D- Steroid	M	11.7619	−7.2
39	Flavin adenine dinucleotide	643975	Coenzyme	Papain-like proteinase	12.0973	−11
				RNA-directed RNA polymerase (RdRp)	11.6978	−12.8
				Helicase (Hel)	11.8304	−11.2
				Guanine-N7 methyltransferase (ExoN)	11.3726	−13.6
				Uridylate-specific endoribonuclease	11.8032	−12.3
				Surface glycoprotein (S)	12.2218	−12.9
				N	11.6575	−13.3
40	Riboflavin 5′-monophosphate	643976	Vitamin B2 derivative	Papain-like proteinase	11.8168	−7
E. Miscellaneous						
41	Dabigatran etexilate	135565674	Anticoagulant	Helicase (Hel)	11.8329	−9.7
				Guanine-N7 methyltransferase (ExoN)	11.7765	−12.5
				Surface glycoprotein (S)	11.6961	−12.7
				N	11.7036	−11.2
42	Dihydroergotamine	10531	Ergot Derivative	Host translation inhibitor nsp1	11.9048	−8.2
				Helicase (Hel)	11.9629	−9.2
43	Edoxaban	10280735	Anticoagulant	Proteinase 3CL-PRO	11.5352	−8.7
44	Elexacaftor	134587348	Corrector of the CFTR protein	Non-structural protein 2 (nsp2)	11.8181	−11.1
45	Ergotamine	8223	Ergot alkaloid with vasoconstrictor and analgesic property.	Host translation inhibitor nsp1	11.9048	−8
				Proteinase 3CL-PRO	11.3610	−8.4
				Helicase (Hel)	11.9629	−9.3
46	Manidipine	4008	Calcium Channel Blocker	Non-structural protein 2 (nsp2)	12.0029	−11.1
47	Mivacurium	5281042	Neuromuscular Blocker	Guanine-N7 methyltransferase (ExoN)	11.7171	−12.1
				Surface glycoprotein (S)	11.5691	−13.2
48	Tadalafil	110635	PDE-5 Inhibitor	Host translation inhibitor nsp1	11.7874	−8.5
				Non-structural protein 2 (nsp2)	11.6870	−11.2
49	Oxatomide	4615	First-generation H1-antihistamine	Papain-like proteinase	12.1627	−7.3

Methods for details), 184 out of 2440 FDA-approved drugs (top 7.5%) with high predicted binding affinity scores were identified. Sixteen drugs with either severe toxicity or topical use only (eyes and skin) were excluded. The remaining 168 drugs included 48 antimicrobial, 18 antineoplastic, 22 central nervous system acting, 35 hormonal, 9 vitamin derivatives, and 36 other agents from miscellaneous classes of drugs. (Table S2). Antimicrobial agents included antibiotics, antiviral, antifungal, and antimycobacterial drugs. The two most frequently observed classes of antibiotics were beta-lactam agents (including penicillin derivatives and cephalosporins) and quinolones. Antiviral drugs included 3 anti-retroviral drugs (didanosine, nelfinavir and cobicistat), 5 anti-HCV drugs (sofosbuvir, elbasvir, pibrentasvir, glecaprevir and velpatasvir), 2 neuraminidase inhibitor (peramivir, oseltamivir) and 1 anti-HBV drug (adefovir dipivoxil). The three highest predicted KIBA score antifungal agents included anidulafungin, isavuconazonium, and natamycin. Antineoplastic drugs included tyrosine and BRAF kinase inhibitors, anthracyclins, and growth factor inhibitors amongst others. Vitamin derivatives included vitamin D derivatives and analogs, flavin derivatives (Riboflavin monophosphate and FAD), and biotin. Hormonal drugs included glucocorticoids, androgen antagonists, and analogs of estrogen, progesterone, oxytocin, vasopressin, and somatostatin. Drugs acting on the central nervous system included dopamine agonists, selective serotonin reuptake inhibitors, and antipsychotic agents. A miscellaneous group of drugs included dipeptidyl peptidase 4 (DPP-4) inhibitors (alogliptin and linagliptin), anticoagulants (edoxaban, ticagrelor, and dabigatran etexilate), calcium channel blockers (manidipine and diltiazem), angiotensin-converting enzyme (ACE) inhibitors (cilazapril, perindopril, trandolapril, enalaprilat, and reserpinine), angiotensin II receptor blocker (eprosartan) and HMG-CoA reductase inhibitors (rosuvastatin and cerivastatin). (For a complete list of drugs with the predicted KIBA scores against each selected SARS-CoV-2 protein see Table S2).

3.2. Molecular docking simulations identified the most promising inhibitors of selected SARS-CoV-2 proteins

We obtained the protein structures for docking from the Zhang lab server (https://zhanglab.ccmb.med.umich.edu/COVID-19/). To verify the accuracy of these structures, we performed a comparison with available crystal structures. We found that the root-mean-square deviation (RMSD) of backbone atomic positions was in the range of 0.5–1.9 Å, establishing the reliability of the Zhang lab structures. It is worth mentioning that these structures were determined experimentally [35–37,39,41] and were published when the manuscript was in the process of compilation.

This server has reported 24 proteins or peptides encompassing the complete genome of SARS-CoV-2. We aim to disrupt the pathways where these proteins are involved to inhibit the normal functioning and

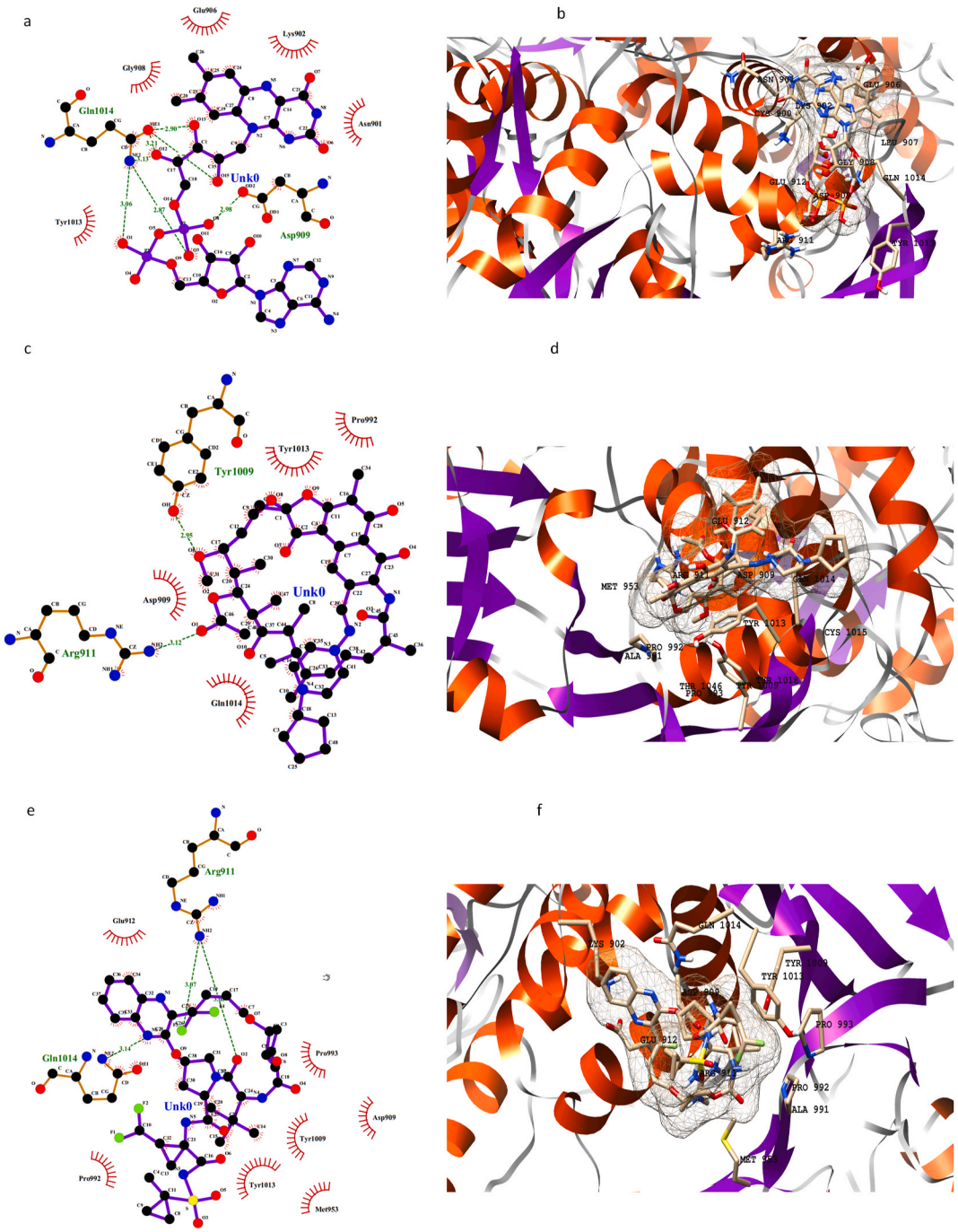


Fig. 2. 2D and 3D representation of PLpro active residues and three selected drugs in the active pocket: FAD disodium (a, b), rifapentine (c, d), and glecaprevir (e, f).

replication cycle of the virus. We ran docking simulations for 168 selected drugs with high predicted KIBA scores against 285 total predicted and/or experimentally proven active sites of all 24 SARS-CoV-2 viral proteins [27,36,37,39,41]. This yielded binding affinity values (kcal/mol) for 47,880 docking simulations (For details of all docking sites and docking energies, see supporting Table S3). AutoDock vina binding affinity values were plotted against KIBA scores of these drugs and drugs with the best *consensus* KIBA scores and binding affinity values were selected. (For visualization of KIBA scores and docking binding affinity values of shortlisted drugs, see our interactive plot in Supporting Figure S2 available online). This provided a list of the top 49 FDA-approved drugs that are predicted to effectively inhibit selected SARS-CoV-2 viral proteins through 134 key drug-protein interactions with a high degree of confidence. These include antimicrobials (n = 12), hormonal agents (n = 14), anti-neoplastic drugs (n = 11), vitamin derivatives (n = 3) and miscellaneous drugs (n = 9) (Table 1). Details of these drugs with the predicted KIBA scores and docking affinity values are provided in Table 1. The results of ligand-protein interactions on molecular docking using AutoDock are summarized in supporting Table S4.

We have shortlisted eighteen docked complexes after extensive analysis and interaction mapping based on their significance in the viral pathways. Complete results for docking energies and active site details are provided in Table S3. The shortlisted complexes include proteins host translation inhibitor (NSP1), papain-like protease (PLpro), proteinase 3CL-pro, RNA-directed RNA polymerase (RdRp), helicase (Hel), surface glycoprotein (S), and nucleocapsid phosphoprotein (N) with key drugs anidulafungin, velpatasvir, glecaprevir, rifabutin, procaine

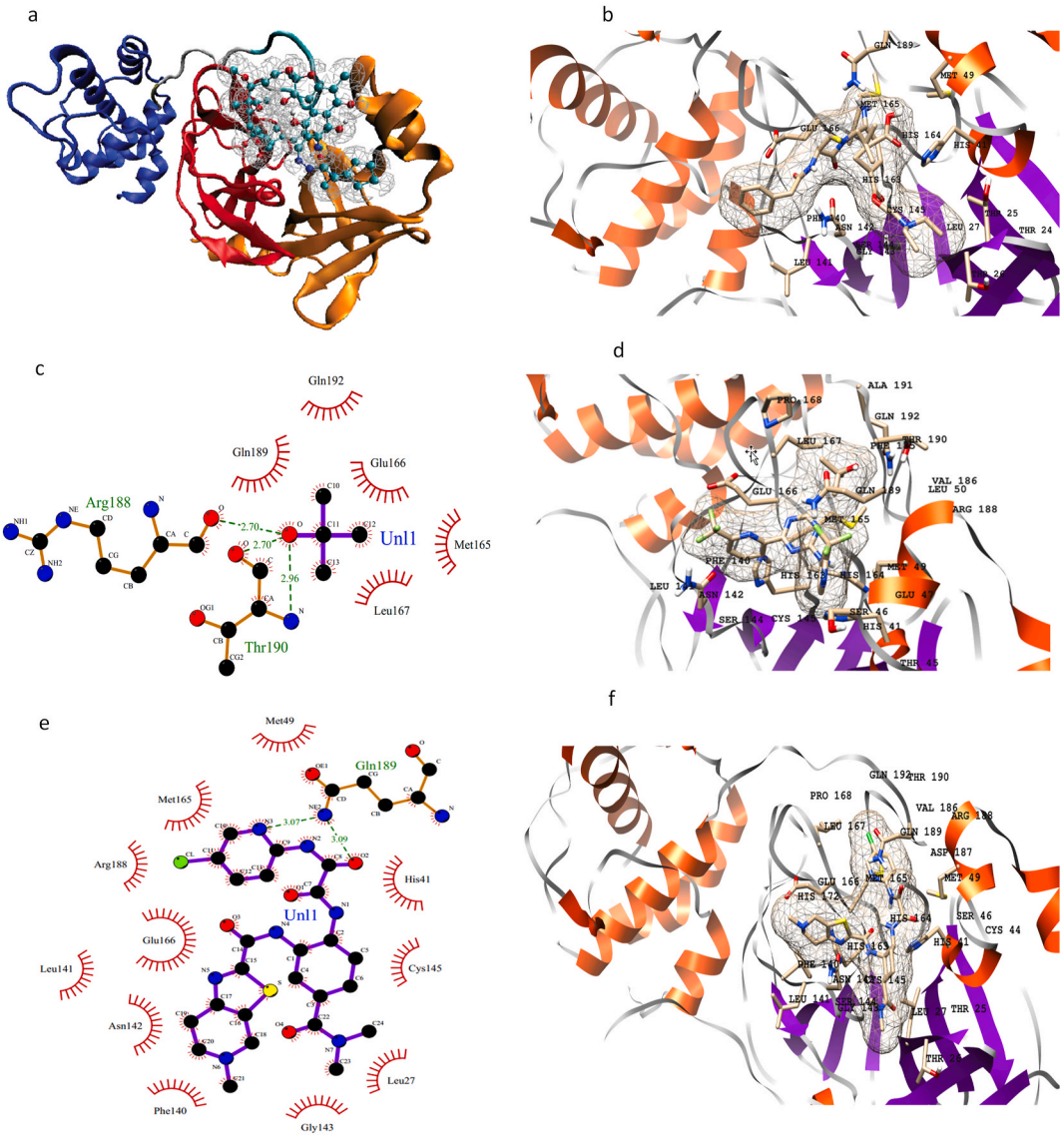


Fig. 3. 3D structure of 3CLpro highlighting three domains, the two β-domains are shown in orange and red colors respectively, blue shows α-helical domain and drug is shown with CPK representation (a); 3D representation of active site residues of 3CLpro surrounding procaine penicillin G (b); 2D and 3D representation of active site residues of 3CLpro surrounding enasidenib (c, d), and edoxaban (e, f).

penicillin G, tadalafil, riboflavin 5′-monophosphate, flavin adenine dinucleotide, terlipressin, desmopressin, elbasvir, oxatomide, enasidenib, edoxaban, and selinexor. The energy values and other information for the selected complexes are provided in Table 1.

In a broader context, PLpro, 3CL-pro, RdRp, and Hel are involved in virus RNA synthesis and replication. Henceforth, more research is being carried out on these targets due to their biological significance. Their structures are mostly available and well elucidated. The fifth protein, NSP1, is the virulence factor that is related to assisting the virus in gene expression and interfering host immune response. The remaining two proteins, that is, S and N are the structural proteins of the virus assisting it in attachment and binding to the host. Results of molecular docking and screening experiments clearly showed that certain drugs have higher affinities for a particular protein target. Here we briefly discuss the results of docking experiments for these drug-protein complexes.

The active site of PLpro is deduced through superimposition with a crystal structure having PDB code 6W9C.A. The RMSD value for backbone atoms came out to be 0.76 Å. There are 4 domains in the monomer of the PLpro enzyme, i) an extended ubiquitin-like domain, ii) thumb domain, ii) palm domain and iv) fingers domain. The binding site for inhibitors is present in between the thumb and palm domains. Three key drugs namely rifapentine, glecaprevir, and FAD disodium have shown better binding affinity values than others (Fig. 2). Other drugs with good binding affinities include riboflavin 5′-monophosphate, oxatomide, velpatasvir, and selinexor. The active pocket of the enzyme majorly consists of hydrophilic residues as Tyr857, Asp909, Arg911, Glu912, Tyr1009, Asn1012, Tyr1013, Gln1014, Cys1015, His1017, and Tyr1018. Other residues present on the interface of thumb and palm domains are Leu907, Gly908, Val910, Thr913, Met953, Pro992, Pro993, Gly1016, and Thr1046. It is noted that riboflavin 5′-monophosphate binds to more hydrophilic residues as compared to other drugs. However, the best binding affinity is shown by oxatomide having both hydrophilic and hydrophobic interactions in a better fit (Fig. 2 and S2a). The results of MOE docking are given in Supporting Table S5 and Figure S2b.

3CL-pro contains two chymotrypsin-like β-domains and an α-helical domain [41]. Domains I and II have the substrate-binding domain in between them and it is represented in figure (Fig. 3a). Residues spanning the active site around the drugs procaine penicillin G, enasidenib, and edoxaban are Thr24, Thr25, Thr26, Leu27, His41, Cys44, Thr45, Ser46,

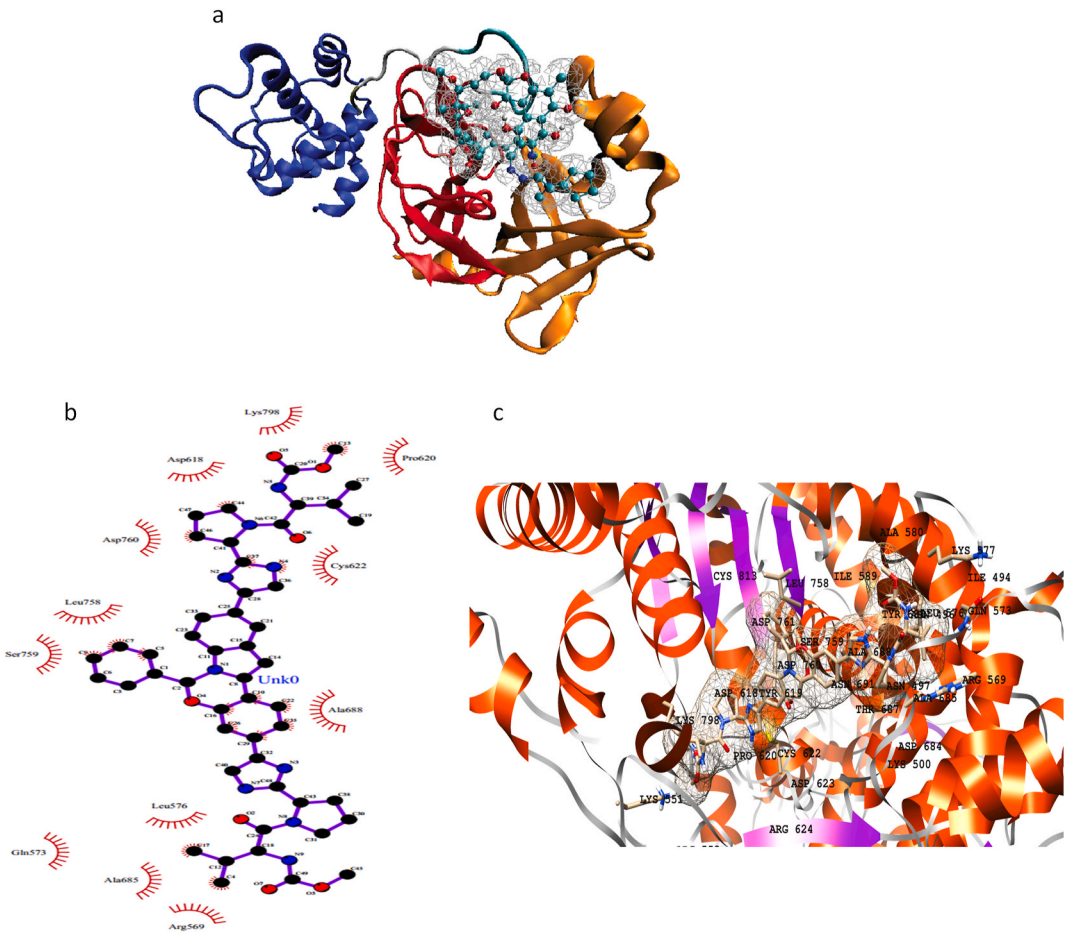


Fig. 4. Structure of RdRp enzyme highlighting different domains: N-terminal domain (red), RdRp domain (blue), interface domain (orange) an additional β-hairpin (green) (a); 2D and 3D representation of active site residues surrounding the drug elbasvir (b, c).

Glu47, Met49, Leu50, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, Met165, Glu166, Leu167, Pro168 and Gln189 (Fig. 3b–f). The results of MOE docking for these interactions are given in Supporting Table S5 and Figure S3.

Residues His41, Ser46, Met49, leu141, Asn142, Glu166, Pro168, and Gln189 are involved in hydrogen bonding with the inhibitors while Asn142 in all three drugs is forming a salt-bridge interaction as well.

RdRp contains an RdRp domain, a nidovirus N-terminal extension domain, both connected by an interface domain [39] (Fig. 4a). The inhibitor binding site is present on the RdRp domain for RdRp inhibition.

Elbasvir, which is a direct antiviral medication, shows good binding with RdRp (Fig. 4b and c). Several hydrophobic residues are involved in the binding of elbasvir and anidulafungin into the binding site of RdRp. Major contributing residues in drug binding are Ile548, Ser549, Arg569, Ile589, Ala685, Ser759, Leu758, Ala688, Gln573, Leu576, Asp760, Asp761, and Cys622. Hydrogen bonding interactions are formed by Ile548, Ser549, Ser759, Cys622, Asp760, Ala550, Lys551, Tyr689, Lys798, Lys577, Cys813, and Ser814. The results of MOE docking for these interactions are given in Supporting Table S5 and Figure S4.

Helicase contains five domains namely N-terminal zinc-binding domain, stalk domain, 1A, 2A, and 1B with inhibitor binding site between domain 1A and 1B (Fig. 5a). The drugs flavin adenine dinucleotide, desmopressin, glecaprevir, and rifabutin inhibit viral helicase protein with the lowest binding affinities. These drugs mainly form hydrogen bonding and electrostatic interactions due to the presence of majority hydrophilic amino acid residues such as Asp260, Glu261, Asn265, His290, Glu319, Lys320, Arg442, Arg443, His464, Lys465, Ser539, Glu540, Asp542, Arg567, and Lys569. Both Arg442 and Arg443 form salt-bridge interactions with the drugs contributing to the tight binding. Fig. 5(b–i) shows a snapshot of all the residues involved in binding. The results of MOE docking for these interactions are given in Supporting Table S5 and Figure S5.

In the second category of targets, Host translation inhibitor or NSP1 is selected for detailed analysis. Tadalafil is a drug that improves exercise capacity by relaxing muscles of the blood vessels and increase of blood flow. Tadalafil and enasidenib show promising results and the best energy values in our docking experiments with NSP1. The binding site contains hydrophobic residues such as Leu39, Leu88, Val89, Leu123, Phe143, and Pro153. Several hydrogen bonds are formed by amino acids: Arg43, Lys72, Glu87, Lys125, Asp144, Tyr145, and Asp144.

Our third category of targets includes S and N proteins. S protein consists of two functional subunits (S1 and S2) (Fig. 6a). it is reported that S protein binds to human ACE2 to enter the cell correlating with its speedy dissemination [36]. A structure for spike protein is reported [35] and deposited in RCSB protein databank (PDB) having PDB ID 6VXX. It was matched with the structure from the Zhang lab server and the RMSD value for backbone atoms came out to be 1.88 Å. We have observed three different drugs to show strong potency for the S protein target: anidulafungin, velpatasvir, and terlipressin. The key residues that are present in the active binding pocket include: Ser45, Ser46, Leu48, Glu281, Leu303, Ser305, Phe306, Thr307, Glu309, Lys310, Arg815, Phe823. Asn824, Thr827, Leu828, Ala829, Ala831, Ala846, Arg847, Leu849, Lys854, Leu945, Gly946, Leu1203, Lys1205, Tyr1209, Ile1210, Pro1213, Ile1216, and Trp1217. Terlipressin is a large compound covering a wide area over the protein surface. It is involved in majority hydrogen bonds and electrostatic interactions. The results of MOE

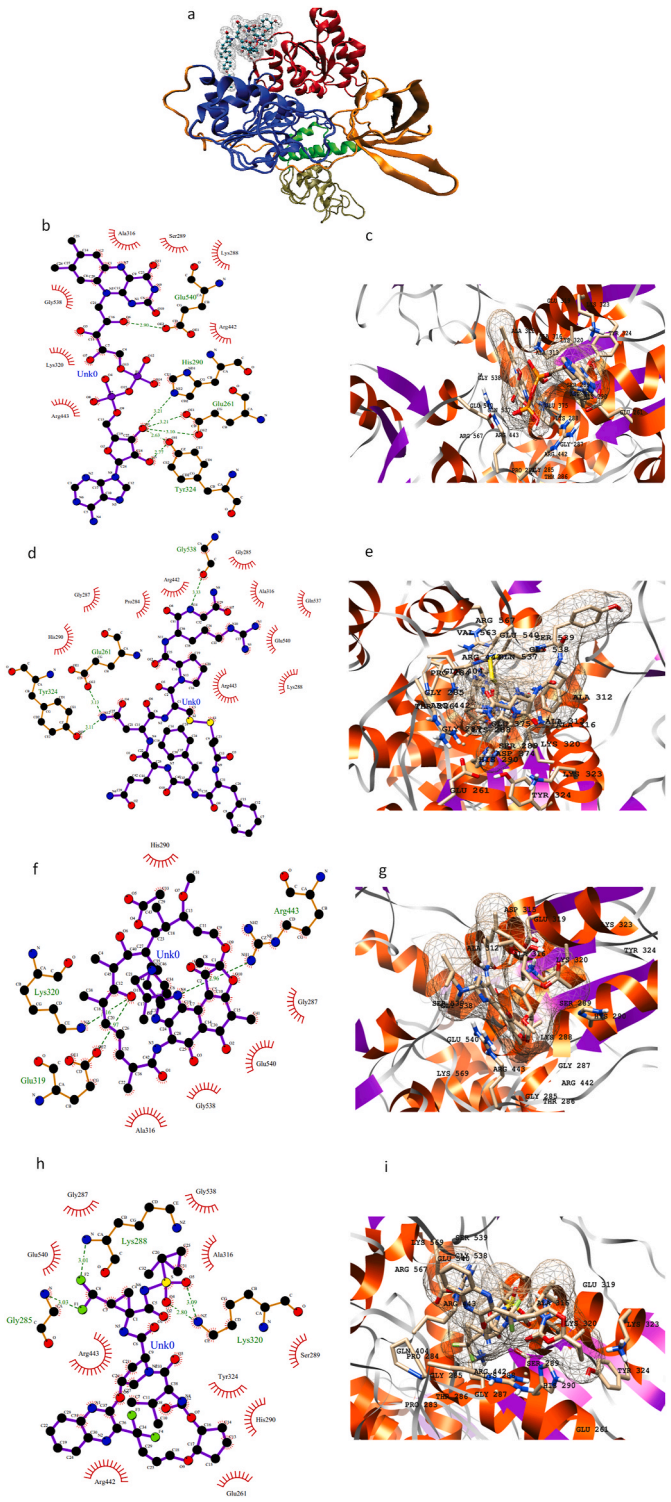


Fig. 5. 3D representation of helicase structure highlighting N-terminal zinc-binding domain (tan), stalk domain (green), 1A (blue), 2A (red), and 1B (orange) with inhibitor binding site between domain 1A and 1B (a); 2D and 3D representation of active site residues surrounding the drugs flavin adenine dinucleotide (b, c), desmopressin (d, e), rifabutin (f, g), and glecaprevir (h, i).

docking for these interactions are given in Supporting Table S5 and Figure S6.

The model of the SARS-CoV-2 nucleocapsid phosphoprotein was superimposed with the NMR structure of the RNA-binding domain of this protein [45] (PDB ID 6YI3) and the RMSD value was 1.14 Å. It is observed that the drug anidulafungin binds at the interface of the protein and interacts with many polar residues naming a few: Arg92, Tyr109, Asn150, Arg259, Gln272, Gln283, and Asp399.

Anidulafungin, a member of the echinocandin class of antifungals, appears to be a promising candidate against SARS-CoV-2 and has high predicted KIBA scores and low docking energies against key viral proteins including RdRp, helicase, exonuclease, S, and N. Anidulafungin is used for the treatment of mucosal and invasive fungal infections [46]. Isavuconazonium, another triazole antifungal approved for the treatment of invasive aspergillosis and invasive mucormycosis, showed high binding affinities for nsp2 and N proteins (Table 1). Interestingly, anidulafungin was shown in a recent study to exhibit strong in-vitro antiviral activity against the SARS-CoV-2 virus with an IC50 value of 4.64 μmol [47]. Although rare, cases of invasive fungal infections have been reported in the literature in association with severe influenza [26] and severe acute respiratory syndrome (SARS) virus infection [48]. Whether COVID-19 patients are also at increased risk of invasive fungal infections is yet to be investigated. Anidulafungin with its potential antiviral activity against SARS-CoV-2 and proven antifungal activity against invasive fungal infections appears to be a promising candidate.

To experimentally validate the inhibitory potential of identified drugs against SARS- CoV-2, we ran a fluorescent intensity assay. SARS-Cov-2 PLpro activity was measured in this fluorescence intensity assay with the use of Ubiquitin-Rhodamine GLY110 as a substrate. IC50 of 5 compounds (rifapentine, velpatasvir, glecaprevir, anidulafungin, and FAD disodium) against SARS-CoV-2 PLpro was measured (see methods). FAD disodium, the drug with the highest predicted KIBA score and binding affinity value, demonstrated the highest inhibitory activity against PLpro with an IC50 of 15.18 μM. Other drugs with high inhibitory activity were Rifapentine (IC50 of 15.18 μM) and glecparevir (IC50 of 36.85 μM) (Fig. 7). Anidulafungin with a relatively low KIBA score and an unfavorable predicted binding affinity for PLpro demonstrated low inhibitory activity (Fig. 7).

Our study highlights the potential of HCV protease inhibitors (elbasvir, velpatasvir, glecraprevir, and pibrentasvir) in inhibiting SARS-CoV-2 proteins. Elbasvir, an HCV NS5A inhibitor, demonstrated the best docking energies with RdRp, a key viral enzyme and an important therapeutic target for COVID-19. In other studies, elbasvir [49] and velpatasvir [50] have been reported to dock well with 3CLpro. Glecraprevir, an HCV NS3/4A protease inhibitor, is often given in combination with pibrentasvir, an NS5A inhibitor, for treatment-experienced cases of HCV infection. While the safety profile of this combination has been well-established [51], the potential of this combination in treating COVID-19 has not been examined so far. Rifabutin and rifapentine belong to the rifamycin group of antibiotics. Both rifabutin and rifapentine inhibit mycobacterial DNA-dependent RNA polymerase, thereby suppressing the initiation of RNA formation, and are used in combination with other drugs for the treatment of tuberculosis [52]. With the demonstrated inhibitory activity against PLpro in our study, rifapentine can be a promising therapeutic option that needs further investigation.

Selinexor, an FDA-approved drug for the treatment of relapsed or refractory multiple myeloma [53], is predicted to inhibit SARS-CoV-2 PLpro in our study. Selinexor is an inhibitor of Exportin-1 (XPO1) which is an important protein involved in the transport of multiple proteins across the nuclear envelope. In addition to its role in cancer, XPO1 also plays an important role in facilitating the transport of viral proteins across the host cell nuclear envelope [54]. Studies have shown that XPO1 plays a critical role in SARS-CoV viral replication by controlling the export of certain SARS-CoV proteins out of the nucleus [55]. We hypothesize that selinexor has both direct antiviral activity through inhibiting PLpro and indirect activity through modulation of host target proteins. A multinational clinical trial has recently been launched to study the efficacy of selinexor in patients with severe COVID-19 disease (ClinicalTrials.gov Identifier: NCT04349098). Enasidenib, another anti—neoplastic drug, inhibits mutant forms of *isocitrate dehydrogenase* 2

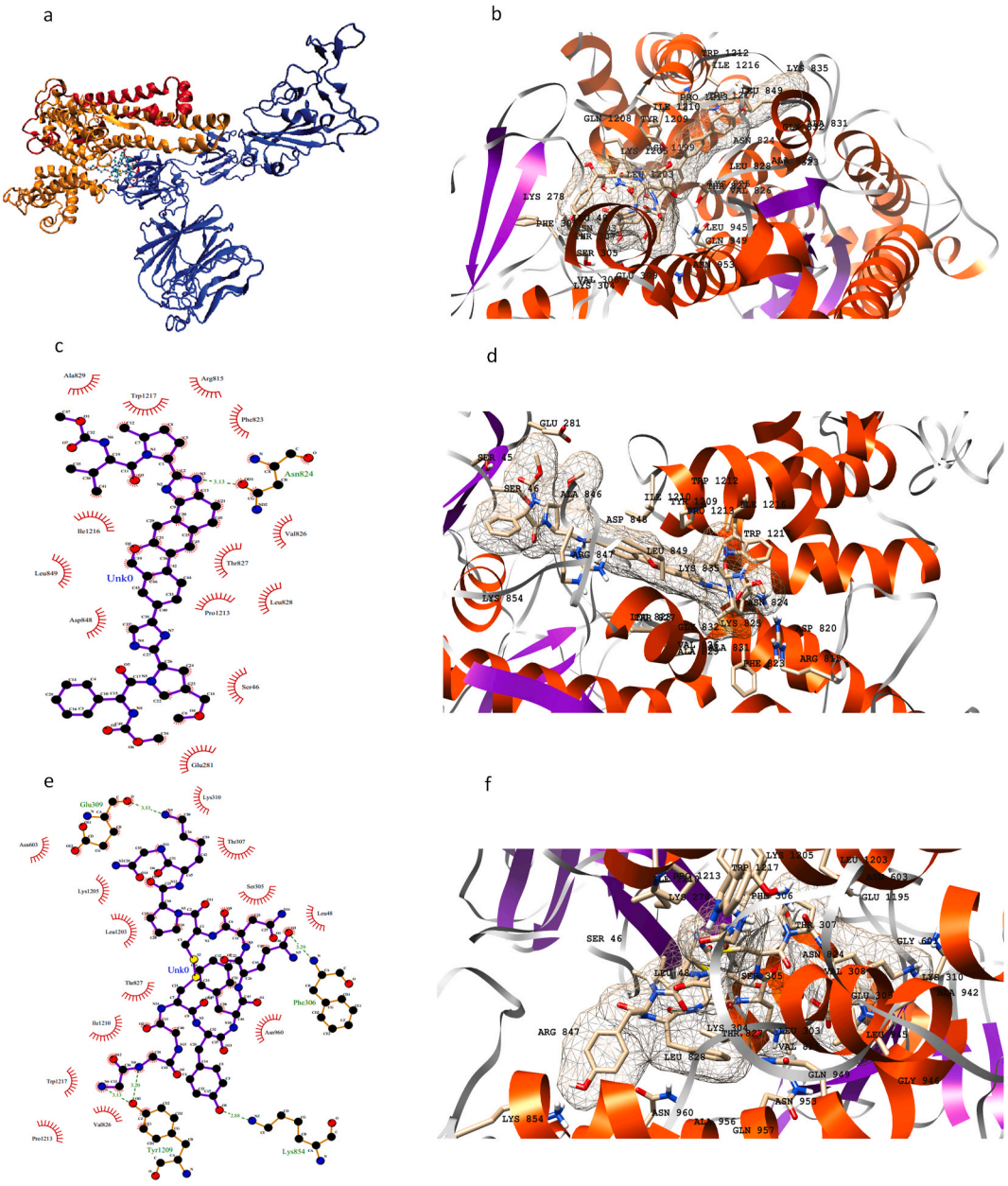
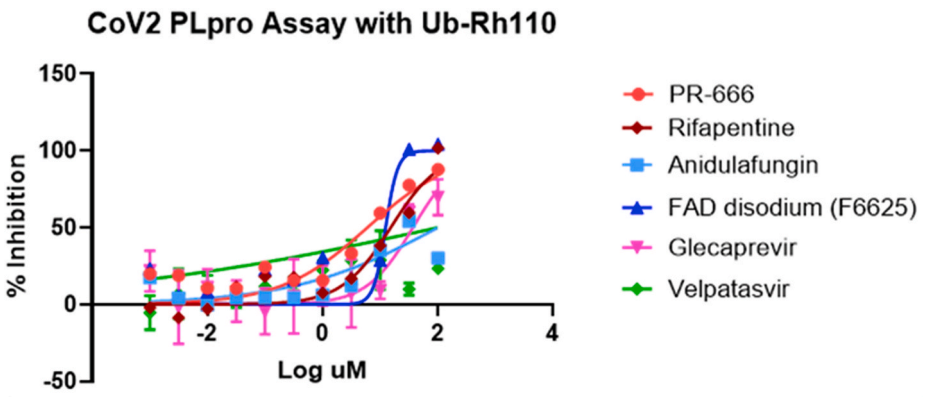


Fig. 6. 3D representation of S protein structure highlighting S1 (blue), S2 (red) and S2' (orange) with inhibitor binding site (a); 3D representation of active site residues surrounding anidulafungin (b); 2D and 3D representation of active residues surrounding velpatasvir (c, d), and terlipressin (e, f).

(IDH2) and is approved for the treatment of refractory form of acute myeloid leukemia (AML) [56]. Here, we have shown that enasidenib demonstrates high binding affinity with two key SARS-CoV-2 viral proteins: 3CLpro and nsp1.

Edoxaban is a rapidly acting selective factor Xa inhibitor and belongs to the Novel Oral Anti-Coagulant (NOACs) class of drugs. In our study, edoxaban has demonstrated the best binding affinity with 3CLpro that is a key enzyme involved in SARS-CoV-2 viral replication and an emerging drug target. Recently, a wealth of clinical data has suggested that COVID-19 is a hypercoagulable state associated with an increased incidence of thrombosis in critically ill patients. Therefore, anti-coagulation is being recommended for prophylactic and therapeutic [57]. Given the potential to inhibit 3CLpro, ease of oral administration, and anticoagulant activity, edoxaban appears to be a promising candidate drug for treating COVID-19. Further in vitro and clinical studies are warranted.

Flavin mononucleotide (FMN; also known as riboflavin-5′-phosphate) and flavin adenine dinucleotide (FAD) are two coenzymes produced from riboflavin (vitamin B2) and function as a prosthetic group of various oxidoreductases. FMN is predicted to have a high binding affinity with PLpro while FAD has high predicted affinities for RdRp, helicase, S, and N amongst other SARS-CoV-2 viral proteins. FAD is a redox-active cofactor that is essential for the functioning of flavoenzymes that play a critical role in many biochemical processes such as oxidative metabolism of macromolecules and electron transport chain [58]. Studies have shown that the intracellular redox state may play an important role in inhibiting viral replication [59]. In one study FAD was shown to enhance the antiviral activity of interferons against herpes virus-1 and influenza virus type A [59]. FAD can also increase the intracellular activity of glutathione and nitric oxide synthase, both of which may play important roles in inhibiting viral replication. FAD has shown binding affinity with spike protein of SARS-CoV-2 in another docking-based study (vina score −7.3) [60]. In another study using molecular docking, riboflavin was found to interact with Papain-like



	PR-666	Rifapentine	Anidulafungin	FAD disodium (F6625)	Glecaprevir	Velpatasvir
IC50	5.968	15.18	100.0	12.39	36.85	100.0
KIBA Score	NA	11.7672	11.6684	12.0973	11.9030	11.9046
Binding affinity (Kcal/mol)		-8.6	16	-11	-9.3	-7.2

Fig. 7. Inhibitory activity of selected drugs against SARS-CoV-2 PLpro. The SARS-Cov-2 PLpro activity was measured in a fluorescence intensity assay with the use of Ubiquitin-Rhodamine GLY110 as a substrate. IC50, KIBA scores, and binding affinity values of 5 compounds (rifapentine, velpatasvir, glecaprevir, anidulafungin, and FAD disodium) are given. PR-666 was used as a positive control.

proteinase (PLpro) and flavin mononucleotide (FMN) interacted with 3C-like main protease (3CLpro) of SARS- CoV-2 virus [43]. Other studies have shown that FAD can decrease lung injury in influenza A H5N1 infected mice by altering the levels of immune response-related genes [61]. In conclusion, the results of our study coupled with evidence from literature dictate that FAD may play an important role against the SARS-CoV-2 virus by directly targeting the virus as well as the host response to the viral replication. However, further evidence from in-vitro and in-vivo studies is required.

We believe that our work can pave way for further studies. We did not include topological information and Physico-chemical properties of the ligands in the training of the model. Further studies incorporating this information are recommended. Additionally, the results of our molecular docking analysis need further validation through molecular dynamic simulations.

4. Conclusion

In conclusion, we have combined deep learning and molecular docking simulations to identify the most promising candidates from the list of FDA-approved drugs that can be repurposed to treat COVID-19. These drugs include anidulafungin, velpatasvir, glecaprevir, rifabutin, rifapentine, procaine penicillin G, tadalafil, riboflavin 5′-monophosphate, flavin adenine dinucleotide, terlipressin, desmopressin, elbasvir, oxatomide, enasidenib, edoxaban, and selinexor amongst others. FAD and rifapentine demonstrate significant inhibitory activity against PLpro in a fluorescence intensity assay with the use of Ubiquitin-Rhodamine GLY110 as a substrate. Further in vitro studies are indicated to investigate the antiviral potential of some of these drugs. For drugs with proven in vitro antiviral activity against SARS-CoV-2, clinical trials are warranted.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contribution (CRediT author statement)

Muhammad Umer Anwaar: Methodology, Software, Formal analysis, Resources Farjad Adnan: Methodology, Software, Resources, Data Curation, Writing - Original Draft Asma Abro: Methodology, Software, Validation, Formal analysis, Writing - Original Draft Rayyan Ahmad Khan: Methodology, Software, Formal analysis, Resources Asad ur Rehman: Formal analysis, Resources, Data Curation Muhammad Osama: Formal analysis, Data Curation Christopher Rainville: Methodology, Formal analysis Suresh Kumar: Methodology, Formal analysis David Sterner: Methodology, Formal analysis Saad Javed: Formal analysis, Data Curation Syed B Jamal: Formal analysis, Visualizations Ahmadullah Baig: Data Curation, Visualizations Muhammad Raffey Shabbir: Formal analysis, Data Curation Waseh Ahsan: Data Curation Tauseef R. Butt: Methodology, Validation, Supervision Muhammad Zaman Assir: Conceptualization, Methodology, Validation, Writing - Original Draft, Supervision.

Author contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript. ‡These authors contributed equally.

Online resources

C–I-TASSER: https://zhanglab.ccmb.med.umich.edu/COVID-19/
PubChem**:** https://pubchem.ncbi.nlm.nih.gov/
DrugBank: https://www.drugbank.ca/
RCSB PDB: https://www.rcsb.org/

Declaration of competing interest

None declared.

Acknowledgments

We are very thankful to Riazuddin S for his guidance and review of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://d

oi.org/10.1016/j.compbiomed.2021.105049.

References

[1] https://www.worldometers.info/coronavirus/last visited on April 28, 2020.
[2] J.A. DiMasi, H.G. Grabowski, R.W. Hansen, Innovation in the pharmaceutical industry: new estimates of R&D costs, J. Health Econ. 47 (2016) 20–33.
[3] S. Pushpakom, F. Iorio, P.A. Eyers, K.J. Escott, S. Hopper, A. Wells, A. Doig, T. Guilliams, J. Latimer, C. McNamee, A. Norris, P. Sanseau, D. Cavalla, M. Pirmohamed, Drug repurposing: progress, challenges and recommendations, Nat. Rev. Drug Discov. 18 (2019) 41–58.
[4] A.L. Hopkins, Network pharmacology: the next paradigm in drug discovery, Nat. Chem. Biol. 4 (2008) 682–690.
[5] J.L. Medina-Franco, M.A. Giulianotti, G.S. Welmaker, R.A. Houghten, Shifting from the single to the multitarget paradigm in drug discovery, Drug Discov. Today 18 (2013) 495–501.
[6] F. Seide, G. Li, X. Chen, D. Yu, Feature engineering in context-dependent deep neural networks for conversational speech transcription, in: 2011 IEEE Workshop on Automatic Speech Recognition & Understanding, 2011, pp. 24–29.
[7] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[8] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press2016.
[9] T.P. Lillicrap, D. Cownden, D.B. Tweed, C.J. Akerman, Random synaptic feedback weights support error backpropagation for deep learning, Nat. Commun. 7 (2016) 1–10.
[10] T. Pahikkala, A. Airola, S. Pietilä, S. Shakyawar, A. Szwajda, J. Tang, T. Aittokallio, Toward more realistic drug–target interaction predictions, Briefings Bioinf. 16 (2015) 325–337.
[11] P.W. Rose, A. Prlić, A. Altunkaya, C. Bi, A.R. Bradley, C.H. Christie, L.D. Costanzo, J.M. Duarte, S. Dutta, Z. Feng, R.K. Green, D.S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A.S. Rose, C. Shao, Y.-P. Tao, Y. Valasatava, M. Voigt, J.D. Westbrook, J. Woo, H. Yang, J.Y. Young, C. Zardecki, H.M. Berman, S.K. Burley, The RCSB protein data bank: integrative view of protein, gene and 3D structural information, Nucleic Acids Res. 45 (2016) D271–D281.
[12] M.I. Davis, J.P. Hunt, S. Herrgard, P. Ciceri, L.M. Wodicka, G. Pallares, M. Hocker, D.K. Treiber, P.P. Zarrinkar, Comprehensive analysis of kinase inhibitor selectivity, Nat. Biotechnol. 29 (2011) 1046–1051.
[13] T. He, M. Heidemeyer, F. Ban, A. Cherkasov, M. Ester, SimBoost: a read-across approach for predicting drug-target binding affinities using gradient boosting machines, J. Cheminf. 9 (2017) 24, 24.
[14] J. Tang, A. Szwajda, S. Shakyawar, T. Xu, P. Hintsanen, K. Wennerberg, T. Aittokallio, Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis, J. Chem. Inf. Model. 54 (2014) 735–743.
[15] K. Tian, M. Shao, Y. Wang, J. Guan, S. Zhou, Boosting compound-protein interaction prediction by deep learning, Methods 110 (2016) 64–72.
[16] K.C. Chan, Z.-H. You, Large-scale prediction of drug-target interactions from deep representations, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016, pp. 1236–1243.
[17] M. Wen, Z. Zhang, S. Niu, H. Sha, R. Yang, Y. Yun, H. Lu, Deep-learning-based drug–target interaction prediction, J. Proteome Res. 16 (2017) 1401–1409.
[18] J. Gomes, B. Ramsundar, E.N. Feinberg, V.S. Pande, Atomic Convolutional Networks for Predicting Protein-Ligand Binding Affinity, arXiv preprint arXiv: 1703.10603, 2017.
[19] H. Öztürk, A. Özgür, E. Ozkirimli, DeepDTA: deep drug–target binding affinity prediction, Bioinformatics 34 (2018) i821–i829.
[20] N. Brooijmans, I.D. Kuntz, Molecular recognition and docking algorithms, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 335–373.
[21] P.D. Lyne, Structure-based virtual screening: an overview, Drug Discov. Today 7 (2002) 1047–1055.
[22] A.N. Jain, Scoring functions for protein-ligand docking, Curr. Protein Pept. Sci. 7 (2006) 407–420.
[23] S.-Y. Huang, S.Z. Grinter, X. Zou, Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions, Phys. Chem. Chem. Phys. 12 (2010) 12899–12908.
[24] J. Carneiro, J. Stewart, Rethinking" shape space": evidence from simulated docking suggests that steric shape complementarity is not limiting for antibody–antigen recognition and idiotypic interactions, J. Theor. Biol. 169 (1994) 391–402.
[25] C. Gorgulla, A. Boeszoermenyi, Z.-F. Wang, P.D. Fischer, P. Coote, K.M.P. Das, Y. S. Malets, D.S. Radchenko, Y.S. Moroz, D.A. Scott, An open-source drug discovery platform enables ultra-large virtual screens, Nature (2020) 1–8.
[26] K. Alshabani, A. Haq, R. Miyakawa, M. Palla, A.O. Soubani, Invasive pulmonary aspergillosis in patients with influenza infection: report of two cases and systematic review of the literature, Expet Rev. Respir. Med. 9 (2015) 89–96.
[27] C. Zhang, W. Zheng, X. Huang, E.W. Bell, X. Zhou, Y. Zhang, Protein structure and sequence re-analysis of 2019-nCoV genome refutes snakes as its intermediate host or the unique similarity between its spike protein insertions and HIV-1, J. Proteome Res. (2020).
[28] J. Yang, A. Roy, Y. Zhang, Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment, Bioinformatics 29 (2013) 2588–2595.
[29] O. Trott, A.J. Olson, AutoDock Vina, Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.
[30] PubChem, https://pubchem.ncbi.nlm.nih.gov/. (Accessed 6 April 2020).
[31] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
[32] K.R. Cousins, ChemDraw Ultra 9.0. CambridgeSoft, 100 CambridgePark Drive, Cambridge, MA 02140. www. cambridgesoft.com. See Web site for pricing options, Journal of the American Chemical Society, 127 (2005) 4115-4116.
[33] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[34] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273.
[35] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, Cell, 2020.
[36] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, Science 367 (2020) 1444–1448.
[37] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science 367 (2020) 1260–1263.
[38] A. Shannon, N.T. Tuyet Le, B. Selisko, C. Eydoux, K. Alvarez, J.-C. Guillemot, E. Decroly, O. Peersen, F. Ferron, B. Canard, Remdesivir and SARS-CoV-2: Structural Requirements at Both Nsp12 RdRp and Nsp14 Exonuclease Active-Sites, Antiviral Research, 2020, p. 104793.
[39] Y. Gao, L. Yan, Y. Huang, F. Liu, Y. Zhao, L. Cao, T. Wang, Q. Sun, Z. Ming, L. Zhang, J. Ge, L. Zheng, Y. Zhang, H. Wang, Y. Zhu, C. Zhu, T. Hu, T. Hua, B. Zhang, X. Yang, J. Li, H. Yang, Z. Liu, W. Xu, L.W. Guddat, Q. Wang, Z. Lou, Z. Rao, Structure of the RNA-dependent RNA polymerase from COVID-19 virus, Science (2020), eabb7498.
[40] R.N. Kirchdoerfer, A.B. Ward, Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors, Nat. Commun. 10 (2019) 2342.
[41] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science (2020), eabb3405.
[42] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L.W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of Mpro from COVID-19 virus and discovery of its inhibitors, Nature (2020).
[43] C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods, Acta Pharmaceutica Sinica B, 2020.
[44] T.P. Sheahan, A.C. Sims, S. Zhou, R.L. Graham, A.J. Pruijssers, M.L. Agostini, S. R. Leist, A. Schäfer, K.H. Dinnon, L.J. Stevens, J.D. Chappell, X. Lu, T.M. Hughes, A.S. George, C.S. Hill, S.A. Montgomery, A.J. Brown, G.R. Bluemling, M. G. Natchus, M. Saindane, A.A. Kolykhalov, G. Painter, J. Harcourt, A. Tamin, N. J. Thornburg, R. Swanstrom, M.R. Denison, R.S. Baric, An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice, Sci. Transl. Med. (2020), eabb5883.
[45] P. Sarma, N. Sekhar, M. Prajapat, P. Avti, H. Kaur, S. Kumar, S. Singh, H. Kumar, A. Prakash, D.P. Dhibar, In-silico homology assisted identification of inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain), J. Biomol. Struct. Dyn. (2020) 1–11.
[46] J.A. Vazquez, Anidulafungin: a new echinocandin with a novel profile, Clin. Therapeut. 27 (2005) 657–673.
[47] S. Jeon, M. Ko, J. Lee, I. Choi, S.Y. Byun, S. Park, D. Shum, S. Kim, Identification of Antiviral Drug Candidates against SARS-CoV-2 from FDA-Approved Drugs, Antimicrobial Agents and Chemotherapy, 2020. AAC.00819-00820.
[48] H. Wang, Y. Ding, X. Li, L. Yang, W. Zhang, W. Kang, Fatal aspergillosis in a patient with SARS who was treated with corticosteroids, N. Engl. J. Med. 349 (2003) 507–508.
[49] T. Sekhar, Virtual Screening Based Prediction of Potential Drugs for COVID-19, 2020. Preprints.
[50] Y.W. Chen, C.-P.B. Yiu, K.-Y. Wong, Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL pro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates F1000Research (2020) 9.
[51] X. Forns, S.S. Lee, J. Valdes, S. Lens, R. Ghalib, H. Aguilar, F. Felizarta, T. Hassanein, H. Hinrichsen, D. Rincon, Glecaprevir plus pibrentasvir for chronic hepatitis C virus genotype 1, 2, 4, 5, or 6 infection in adults with compensated cirrhosis (EXPEDITION-1): a single-arm, open-label, multicentre phase 3 trial, Lancet Infect. Dis. 17 (2017) 1062–1068.
[52] C. Grassi, V. Peona, Use of rifabutin in the treatment of pulmonary tuberculosis, Clin. Infect. Dis. 22 (Suppl 1) (1996) S50–S54.
[53] A. Chari, D.T. Vogl, M. Gavriatopoulou, A.K. Nooka, A.J. Yee, C.A. Huff, P. Moreau, D. Dingli, C. Cole, S. Lonial, M. Dimopoulos, A.K. Stewart, J. Richter, R. Vij, S. Tuchman, M.S. Raab, K.C. Weisel, M. Delforge, R.F. Cornell, D. Kaminetzky, J. E. Hoffman, L.J. Costa, T.L. Parker, M. Levy, M. Schreder, N. Meuleman, L. Frenzel, M. Mohty, S. Choquet, G. Schiller, R.L. Comenzo, M. Engelhardt, T. Illmer, P. Vlummens, C. Doyen, T. Facon, L. Karlin, A. Perrot, K. Podar, M.G. Kauffman, S. Shacham, L. Li, S. Tang, C. Picklesimer, J.R. Saint-Martin, M. Crochiere, H. Chang, S. Parekh, Y. Landesman, J. Shah, P.G. Richardson, S. Jagannath, Oral selinexor-dexamethasone for triple-class refractory multiple myeloma, N. Engl. J. Med. 381 (2019) 727–738.

[54] D.G. Widman, S. Gornisiewicz, S. Shacham, S. Tamir, In vitro toxicity and efficacy of verdinexor, an exportin 1 inhibitor, on opportunistic viruses affecting immunocompromised individuals, PLoS One 13 (2018).
[55] E.C. Freundt, L. Yu, E. Park, M.J. Lenardo, X.-N. Xu, Molecular determinants for subcellular localization of the severe acute respiratory syndrome coronavirus open reading frame 3b protein, J. Virol. 83 (2009) 6631–6640.
[56] E.M. Stein, C.D. DiNardo, D.A. Pollyea, A.T. Fathi, G.J. Roboz, J.K. Altman, R. M. Stone, D.J. DeAngelo, R.L. Levine, I.W. Flinn, H.M. Kantarjian, R. Collins, M. R. Patel, A.E. Frankel, A. Stein, M.A. Sekeres, R.T. Swords, B.C. Medeiros, C. Willekens, P. Vyas, A. Tosolini, Q. Xu, R.D. Knight, K.E. Yen, S. Agresta, S. de Botton, M.S. Tallman, Enasidenib in mutant IDH2 relapsed or refractory acute myeloid leukemia, Blood 130 (2017) 722–731.
[57] B. Bikdeli, M.V. Madhavan, D. Jimenez, T. Chuich, I. Dreyfus, E. Driggin, C. Der Nigoghossian, W. Ageno, M. Madjid, Y. Guo, COVID-19 and thrombotic or thromboembolic disease: implications for prevention, antithrombotic therapy, and follow-up, J. Am. Coll. Cardiol. (2020).
[58] T.A. Giancaspero, G. Busco, C. Panebianco, C. Carmone, A. Miccolis, G.M. Liuzzi, M. Colella, M. Barile, FAD synthesis and degradation in the nucleus create a local flavin cofactor pool, J. Biol. Chem. 288 (2013) 29069–29080.
[59] S. Yokochi, Y. Ishiwata, H. Saito, H. Ebinuma, M. Tsuchiya, H. Ishii, Stimulation of antiviral activities of interferon by a liver extract preparation, Arzneim. Forsch. 47 (1997) 968–974.
[60] J. Zhang, G. Wang, H. Li, T. Wei, H. Wang, X. Wu, Y. Lu, S. Guan, F. Dong, C. Dong, In Silico Screening of Potential Spike Glycoprotein Inhibitors of SARS-CoV-2 with Drug Repurposing Strategy, 2020.
[61] F. Huang, C. Zhang, Q. Liu, Y. Zhao, Y. Zhang, Y. Qin, X. Li, C. Li, C. Zhou, N. Jin, Identification of amitriptyline HCl, flavin adenine dinucleotide, azacitidine and calcitriol as repurposing drugs for influenza A H5N1 virus-induced lung injury, PLoS Pathog. 16 (2020), e1008341.