

D3AI-CoV: a deep learning platform for predicting drug targets and for virtual screening against COVID-19

Yanqing Yang[†], Deshan Zhou[†], Xinben Zhang[†], Yulong Shi, Jiaxin Han, Liping Zhou, Leyun Wu, Minfei Ma, Jintian Li, Shaoliang Peng, Zhijian Xu  and Weiliang Zhu

Corresponding authors: Weiliang Zhu, CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China. Tel: +86-21-50805020; Fax: +86 021 50807188; E-mail: wzhu@simm.ac.cn; Zhijian Xu, CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China. Tel: +86-21-50806600-1201; Fax: +86 021 50807188; E-mail: zjxu@simm.ac.cn; Shaoliang Peng, Department of Computer Science, Hunan University, Changsha 410082, China. Tel: +86-0731-88822273; Fax: +86 0731 88664153; E-mail: slpeng@hnu.edu.cn

[†]The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint first authors.

Abstract

Target prediction and virtual screening are two powerful tools of computer-aided drug design. Target identification is of great significance for hit discovery, lead optimization, drug repurposing and elucidation of the mechanism. Virtual screening can improve the hit rate of drug screening to shorten the cycle of drug discovery and development. Therefore, target prediction and virtual screening are of great importance for developing highly effective drugs against COVID-19. Here we present D3AI-CoV, a platform for target prediction and virtual screening for the discovery of anti-COVID-19 drugs. The platform is composed of three newly developed deep learning-based models i.e., MultiDTI, MPNNs-CNN and MPNNs-CNN-R models. To compare the predictive performance of D3AI-CoV with other methods, an external test set, named Test-78, was prepared, which consists of 39 newly published independent active compounds and 39 inactive compounds from DrugBank. For target prediction, the areas under the receiver operating characteristic curves (AUCs) of MultiDTI and MPNNs-CNN models are 0.93 and 0.91, respectively, whereas the AUCs of the other reported approaches range from 0.51 to 0.74. For virtual screening, the hit rate of D3AI-CoV is also better than other methods. D3AI-CoV is available for free as a web

Yanqing Yang is a postgraduate at Shanghai Institute of Materia Medica. His research interests are deep learning, molecular docking and virtual screening. His affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Deshan Zhou is a postgraduate at Hunan University. His research interest is deep learning. His affiliation is with Department of Computer Science, Hunan University, Changsha, 410082, China.

Xinben Zhang got his Master's degree at East China University of Science and Technology. His research interest is software development. His affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China.

Yulong Shi is a PhD student at Shanghai Institute of Materia Medica. His research interest is molecular docking method development. His affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Jiaxin Han is a postgraduate at Nanjing University of Chinese Medicine. His research interest is molecular docking and virtual screening. His affiliation is with School of Chinese Materia Medica, Nanjing University of Chinese Medicine, Nanjing, 210046, China.

Liping Zhou is a PhD student at Shanghai Institute of Materia Medica. Her research interest is molecular dynamics. Her affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Leyun Wu is a postgraduate at Shanghai Institute of Materia Medica. Her research interest is molecular dynamics. Her affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Minfei Ma is a postgraduate at Shanghai Institute of Materia Medica. Her research interest is deep learning. Her affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Jintian Li is a postgraduate at Shanghai Institute of Materia Medica. Her research interest is deep learning. Her affiliation is with CAS Key Laboratory of Receptor Research, Stake Key Laboratory of Drug Research; Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China; School of Pharmacy, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing, 100049, China.

Professor Shaoliang Peng got his PhD degree at National University of Defense Technology in 2008. His research interest is artificial intelligence.

Professor Zhijian Xu got his PhD degree at Shanghai Institute of Materia Medica in 2012. His research interests include computer-aided drug design, computational chemistry, computational biology and artificial intelligence. More information could be found at the website: https://www.researchgate.net/profile/Zhijian_Xu

Professor Weiliang Zhu received his PhD degree from Shanghai Institute of Materia Medica in 1998. His main research fields are computer-aided drug design, computational biology, computational chemistry and pharmaceutical chemistry, with a special focus on the theoretical research and method development of drug design.

Received: January 11, 2022. Revised: March 13, 2022. Accepted: March 31, 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

application at <http://www.d3pharma.com/D3Targets-2019-nCoV/D3AI-CoV/index.php>, which can serve as a rapid online tool for predicting potential targets for active compounds and for identifying active molecules against a specific target protein for COVID-19 treatment.

Keywords: D3AI-CoV, COVID-19, deep learning, target prediction, virtual screening

Introduction

COVID-19 caused by SARS-CoV-2 has become a global pandemic [1–3]. As of 28 December 2021, there have been more than 270 million fatalities caused by the virus [4]. Although the vaccines against COVID-19 have shown great success, immune escape is becoming a real threat as new variants of the virus are emerging from time to time [5, 6]. Besides, there are eight other coronaviruses regarded as potential health threats, viz., severe acute respiratory syndrome coronavirus (SARS) in 2003, Middle East respiratory syndrome coronavirus (MERS) in 2012, human betacoronavirus 2c EMC/2012, human coronavirus 229E, feline infectious peritonitis virus, human coronavirus OC43, human coronavirus NL63 and human coronavirus HKU1. Therefore, identifying effective drug targets and developing effective drugs accordingly to cure COVID-19 as well as other coronaviruses are of great importance.

At the beginning of the COVID-19 outbreak, we developed a web server, namely D3Targets-2019-nCoV (<http://www.d3pharma.com/D3Targets-2019-nCoV/index.php>), for target prediction and virtual screening against COVID-19. The server is composed of two modules, a structure-based module named D3Docking [7, 8] and a ligand-based module named D3Similarity [9]. Other computational tools were also developed for combating COVID-19 e.g., COVID-19 Docking Server [10], Shennong [11], DockThor-VS [12], Virus-CKB [13], MolAICal [14] and REDIAL-2020 [15]. However, structure-based approaches are in general limited by the availability of their three-dimensional structures, whereas ligand-based approaches are usually hard to reveal the ligand–protein interactions.

Artificial intelligence (AI), especially deep learning (DL), has been applied successfully to drug discovery and design, and it has shown its strength in improving the accuracy. For example, Atomwise, the first DL-based technology for discovering drugs [16], has been successfully applied to discover hit compounds for more than 80 targets. With IBM Watson [17], Pfizer carried out its immuno-oncology drug discovery program at high efficiency [18]. Stokes developed a DL-based model that has identified eight antibacterial compounds from the ZINC15 database [19]. Zhavoronkov developed a deep model (GENTRL) for discovering potent inhibitors of discoidin domain receptor 1 [20]. Likewise, DL also played a key role during the COVID-19 pandemic. For example, Deep Docking has been applied to discover hits against SARS-CoV-2 M^{PRO} from 1.3 billion compounds [21]. COVIDVS-3 DL-based model was used to screen 4.9 million drug-like molecules from the ZINC15 database, discovering a compound as the inhibitor of the 3C-like protease of SARS-CoV-2 [22]. Through DL and AI, baricitinib, atazanavir and other antiviral agents against

hepatitis C have been identified as effective anti-COVID-19 agents [23–25]. Apart from the compounds mentioned above, Zhang et al. have discovered 26 herbal plants containing anti-COVID-19 ingredients using molecular docking and network pharmacology analysis [26].

Recently, we developed a multimodal drug–target interaction (DTI) prediction model, ‘MultiDTI’ [27], which projects drug, target, side effect and disease nodes in the heterogeneous network into a common space. If a drug and a target are connected by an edge, the Euclidean distance between them in the common space is adjusted to be closer. A prediction layer is designed to predict the DTI score based on the distance between the drug and the target in the common space. In addition, the graph neural network performs well in analyzing graph- or tree-like structures and can extract the contextual information contained in graph neighborhoods. Molecules can be regarded as molecular graphs, with atoms as nodes and bonds as edges. As a kind of graph neural network, Message Passing Neural Networks (MPNNs) outperform fingerprint-based methods in predicting the properties of small molecules [28, 29]. Therefore, MPNNs might be a complementary approach to MultiDTI.

Considering both the limitations of the conventional structure-based or ligand-based approaches and the unique advantages of DL, we utilized the two approaches in this study to construct MultiDTI and MPNNs-CNN [29, 30] models for target prediction and MPNNs-CNN-R for virtual screening against COVID-19. Together with a validation using an external test, it was found that the accuracy and the efficiency of target prediction and virtual screening with the newly developed DL-based models are tremendously improved in comparison with existing methods.

Materials and methods

Preparation of database including all compounds against pathogenic coronaviruses

Through literature search, a total of 842 molecules with potential activity against nine pathogenic coronaviruses are currently collected in our database. All molecules and their related information in the database are downloadable in sdf format from the <http://www.d3pharma.com/D3Targets-2019-nCoV/CoVILigands/index.php> webpage. The database will be continuously updated in the future.

Preprocessing of small molecules and protein targets

Firstly, all compound–target pairs collected in our database were used as training set. The canonical simplified molecular-input line-entry system (SMILES) files of all small molecules in the training set were prepared by using Open Babel [31]. The sequence files of all

target proteins in the training set were downloaded from the uniprot website [32]. Secondly, all the compounds and target proteins in the training set were indexed. An interaction network was formed between all the compounds and their targets. 0 and 1 were used to represent the interaction between the compounds and the targets in the network. In detail, 0 means there is no interaction between the compound and the target, 1 means there is interaction. The interaction network between the compounds and the targets in the training set was used as the input data for model training.

MultiDTI model for target predicting

We constructed a DTI network, in which the compounds were represented by SMILES and targets were represented by sequence. The drugs and targets were projected into a multimodal common space after obtaining the embedding representation of SMILES and protein sequence. Compound–target pair connected by an edge in the DTI network would have smaller Euclidean distance in the multimodal common space. In detail, n-gram embedding technology was used to obtain ‘words’ of SMILES and protein sequence. We constructed a compound dictionary and a protein dictionary based on all SMILES and protein target sequences in DrugBank and DTI networks. SMILES and protein sequence were vectorized according to the dictionary. A three-layer convolutional neural networks (CNN) was applied to obtain the regional embedding of each ‘word’ in SMILES and protein sequence. Next, multiple down-sampling residual layers were used to extract more global information. The multilayer perceptron was used to project the representations of drugs and targets into the common space. At last, the Euclidean distance between drug and target in the common space was converted to predicting score. The model continuously adjusts the Euclidean distance between drugs and targets in the common space during the training process based on the compound–target pairs in the training set. The final model is the projection of the DTI network in a multimodal common space. The framework of MultiDTI model is illustrated in Figure 1A. With the purpose of selecting the best architecture to be used in MultiDTI, we tried to use CNN and RNN to extract features for SMILES and FASTA, respectively. As shown in Table S1, the accuracy of CNN (0.86–0.90) is significantly better than that of RNN (0.47–0.52). Next, we also trained and tested models for different CNN layers (Table S2). At the end, we selected three-layers CNN and multiple residual layers to extract features since its performance was better than others.

Classification model for target predicting

Compared with MutiDTI model, we used MPNNs to extract the features of small molecules, and then used classification algorithms to explore the potential relationship between all active compounds against pathogenic coronaviruses and their targets. In detail, as shown in Table S3, the atom characteristics including the atom type, the number of atomic bonds, the formal

charge, chirality and aromaticity as well as the bond characteristics including the bond type and *cis-trans* isomerism were obtained by RDKit package [33] and were mapped to tensors. The atoms in the molecule are regarded as nodes and the bonds are regarded as edges. Thereby, a molecule can be represented as a network graph containing a lot of chemical information. In this MPNNs model, after a message passing phase and a readout phase, a molecule as a graph could be extracted as a feature vector that can represent the molecular structure. Meanwhile, as for target protein, we used CNN to extract their feature vectors. Finally, multilayer perceptron and logistic algorithm were constructed to discover potential connections between the compounds and targets. The framework of the MPNNs-CNN model is illustrated in Figure 1B.

Regression model based on activity for virtual screening

Target prediction of the compounds with known activity is helpful for further structural modification, whereas virtual screening is useful for hit discovery. Thereby, we normalized the activity data of all compounds with known targets in the database. We set 100 μM as the threshold for compound–target interaction. The normalized function was as follows:

$$\text{score} = 2 - \lg(\text{activity}) \quad (1)$$

where activity represented the activity data of compounds against pathogenic coronaviruses in the experiment, the unit was μM . Similar to classification model for target predicting, MPNNs and CNN were used to extract features of small molecules and protein targets, respectively. Multilayer perceptron and regression algorithm were constructed to discover potential relationships between compound–target pairs and their bioactivity.

System training and test procedures

For each model, all drug–target pairs in the interaction network were used as the dataset for model training and validation. Furthermore, due to the finding of there being a drastic difference between the number of positive drug–target pairs and negative pairs, we oversampled the positive samples by 10 times so as to increase the generalization ability of the models. Next, we carried out 10-fold cross-validation on the prepared drug–target pairs. In detail, 90% of sample pairs selected by stratified sampling were used as training data, and the remaining 10% of samples were used for validation. Each model was optimized by mini-batch gradient descent method. Backpropagation strategy was used to update parameters of the models. Weight decay and dropout method were used in all our models to prevent the neural network from overfitting. To further test the generalization performance of each model, we collected 39 active compounds against COVID-19 and their targets information from the latest published literatures. In addition, we

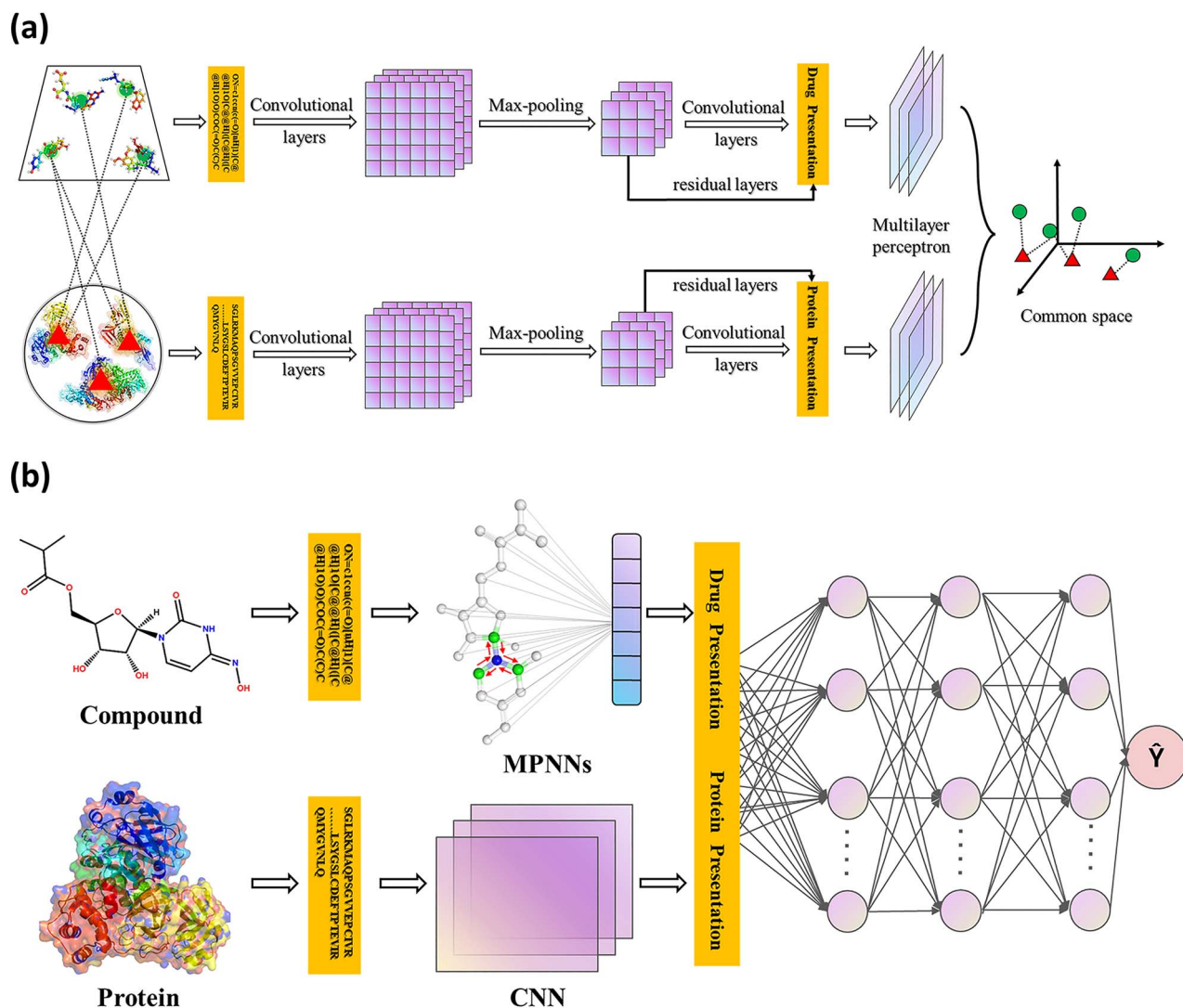


Figure 1. Frameworks of D3AI-CoV. **(A)** The framework of MultiDTI model, a three-layer convolutional neural network and multiple down-sampling residual layers is used to extract features of SMILES and protein amino acid sequence, and multilayer perceptron is used to project the representations of drugs and targets into the common space. **(B)** The framework of MPNNs-CNN model, MPNNs and CNN is used to extract features of compound SMILES and protein sequence, respectively. Multilayer perceptron and logistic algorithm were used to predict potential connections between compounds and targets.

randomly selected 39 Food and Drug Administration (FDA)-approved drugs from the DrugBank [34] based on the molecular weight range of the 39 active compounds and randomly paired them with the protein targets in the database. Accordingly, a dataset, namely Test-78, was constructed based on the 78 compounds for testing the predictive ability of the MultiDTI model and the MPNNs-CNN model, and for comparing D3AI-CoV with other methods. The structures of the 78 compounds are shown in Table S4. And Table 1 summarizes the training, validation and testing datasets.

Performance test metrics

Six performance indicators, viz., area under the curve (AUC), area under the precision-recall curve (AUPR), Acc, Pre, Recall and F1, were used to evaluate the two models for target predicting. Pearson correlation and concordance index were used to evaluate the regression model for virtual screening. AUC and AUPR were obtained based

on the area under the curve, and the other indicators are defined as follows:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (2)$$

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

where TP (True Positive) and TN (True Negative) represent the numbers of correctly predicted positive and negative samples, respectively. FP (False Positive) and FN (False Negative) represent the numbers of wrong predicted positive and negative samples.

$$\text{Pearson correlation} = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - (\sum x_i)^2} \sqrt{N \sum y_i^2 - (\sum y_i)^2}} \quad (6)$$

Table 1. Summary of datasets used for training, validation and testing

| Training and validation dataset (10-fold cross-validation) | | |
|--|--------------------|------------------------|
| | Number of the data | Percentage of the data |
| Training | 24 336 | 90% |
| Validation | 1352 | 10% |
| Test-78 | | |
| | Number of the data | Percentage of the data |
| Positive test set (from literatures) | 39 | 50% |
| Negative test set (from DrugBank) | 39 | 50% |

where N represents the number of all samples. x_i and y_i represent the labels and predicted values of samples, respectively.

$$\text{Concordance index} = \frac{\sum_{i,j} 1_{T_j < T_i} \cdot 1_{\eta_j > \eta_i} \cdot \delta_j}{\sum_{i,j} 1_{T_j < T_i} \cdot \delta_j} \quad (7)$$

where η_i represents the risk score of a unit i , $1_{T_j < T_i}$ means $1_{T_j < T_i} = 1$ if $T_j < T_i$ else 0, $1_{\eta_j > \eta_i}$ means $1_{\eta_j > \eta_i} = 1$ if $\eta_j > \eta_i$ else 0.

To evaluate the virtual screening abilities of various methods, we proposed the concept of hit rate. For example, the Test-78 contains x compounds targeting 3C-like protease, and the number of 3C-like protease inhibitors among the top x in the virtual screening result is y .

$$\text{hit rate} = \frac{y}{x} \quad (8)$$

The workflow

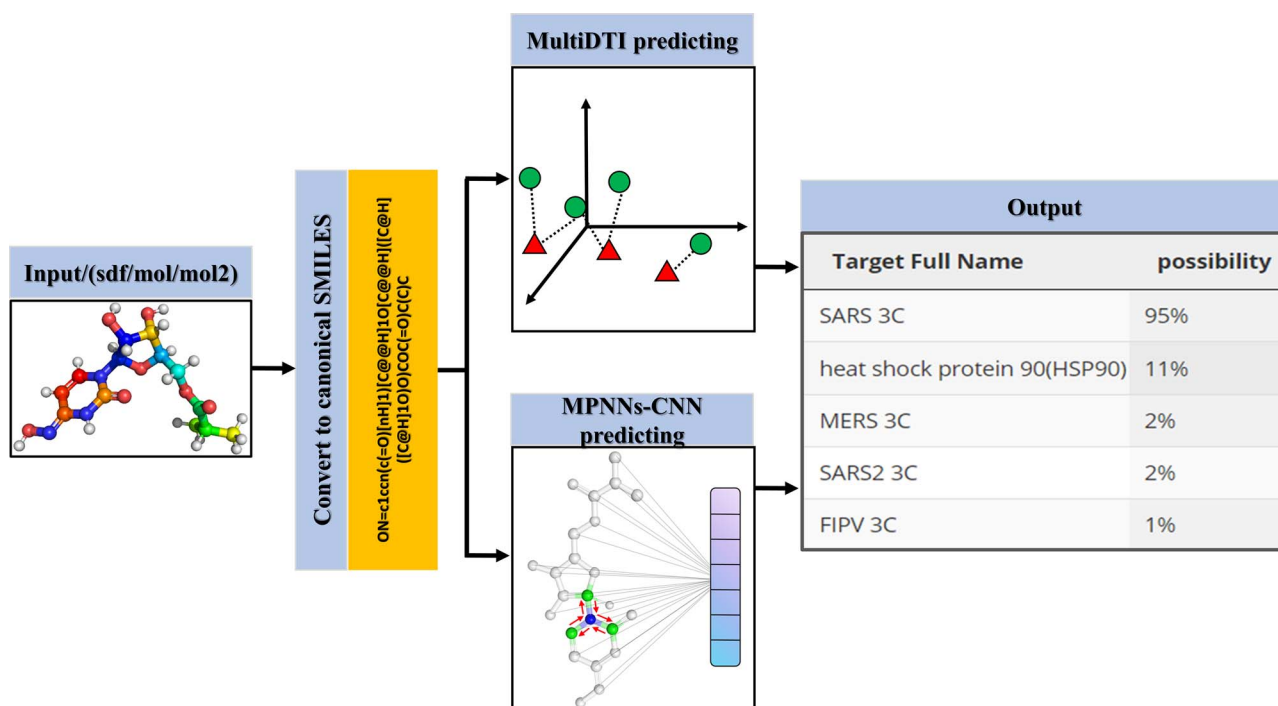
Based on all active compounds against the nine pathogenic coronaviruses in the database, we trained three DL-based models for target prediction and virtual screening. For the MultiDTI model, n-gram was used to obtain ‘words’ of SMILES. A three-layer CNN and multiple down-sampling residual layers were used to extract features of SMILES. The SMILES was then projected to the trained multimodal common space by using multilayer perceptron. Target prediction was achieved by calculating the Euclidean distance between the molecule and all targets in the multimodal common space. For classification model, we used MPNNs to extract the features of small molecule submitted by user. The features were used as classification model input for target prediction.

For user convenience, we developed a webserver, named D3AI-CoV. By submitting small molecules, canonical SMILES files will be generated with Open Babel, and the predicted target proteins of the submitted molecules will be displayed on the web page. The workflow was illustrated in Figure 2.

Results and Discussion

Expanded database of the active compounds against nine pathogenic coronaviruses

Many active compounds against various coronaviruses have been reported since the SARS outbreak in 2003. As for COVID-19, numerous active compounds at the cellular level or *in vivo* have been discovered but their targets are still unknown. For example, clofazimine, which has been approved as an antileprosy drug by the US FDA, has been found active in treating COVID-19 cases when

**Figure 2.** The workflow of D3AI-CoV for target prediction.

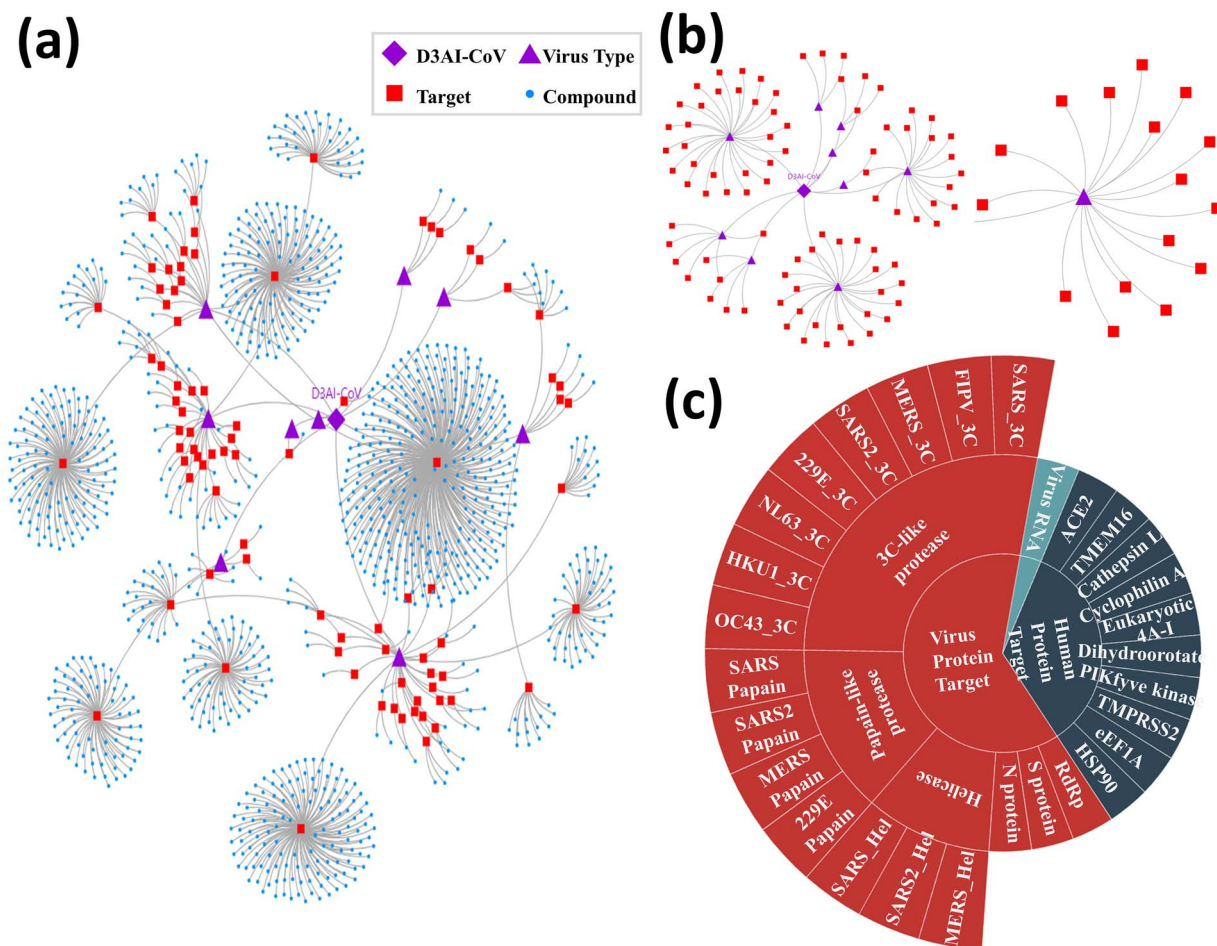


Figure 3. Overview of the database CoViLigands. (A) The interactive diagram categorizes the database according to virus types. (B) Enlarged view of (A). (C) Interactive fan diagram based on all targets in the database.

combined with remdesivir [35]. Rosenke *et al.* reported that an orally administered nucleoside analog, MK-4482, can inhibit SARS-CoV-2 *in vivo* [36]. The experiment data of MK-4482 in animals indicate that it is a promising drug to cure COVID-19. 25-Hydroxycholesterol has been reported as a potent SARS-CoV-2 inhibitor [37], its EC_{50} and the ideal safety profile show potential for further clinical development for COVID-19 treatment. Jan *et al.* have screened >3000 agents, 15 of which have been identified as inhibitors of SARS-CoV-2 in concentrations ranging from 0.1 nM to 50 μ M [38], but no clear target information is available for the 15 inhibitors. Besides, some other drugs such as AT-527 [39], X-206 [40] and ACA [41] are also promising for the treatment of COVID-19, but without clear target information. Apart from chemical drugs, extracts of *Ganoderma lucidum* (RF3), *Perilla frutescens*, and *Mentha haplocalyx* have been found effective against SARS-CoV-2 infection [38]. Glycyrrhizin, a common Chinese herbal medicine, is an efficient and safe natural compound to inhibit SARS-CoV-2 and SARS [42, 43], but without target information either.

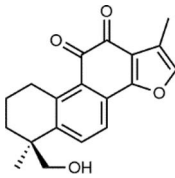
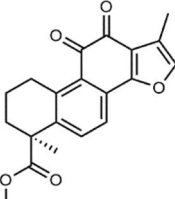
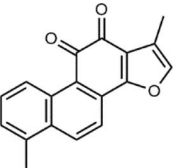
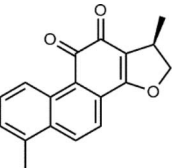
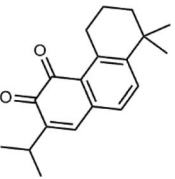
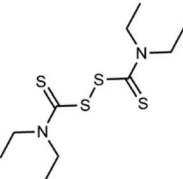
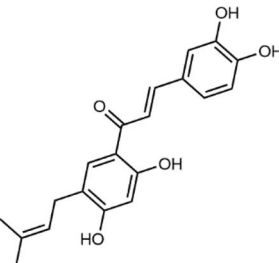
After careful review of the literature, we found 842 bioactive molecules and 29 targets against the 9

pathogenic coronaviruses. We collected the information of the 842 compounds, including molecular structures, bioactivities, target proteins, coronavirus types and crystal structures. As shown in Figure 3A, we classified all the active compounds in CoViLigands according to the virus types and made an interactive interface for easy view on the webserver. As shown in Figure 3C, all compounds were classified according to their targets. Details of ligand structures and the associated information of all compounds are provided on the webpage.

Dual inhibitors in the database

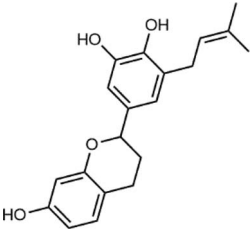
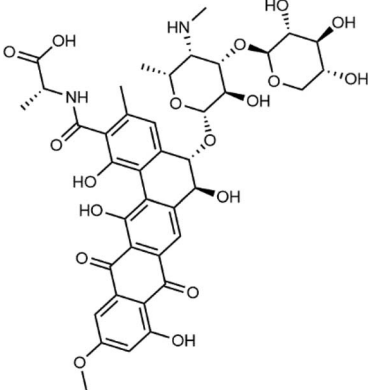
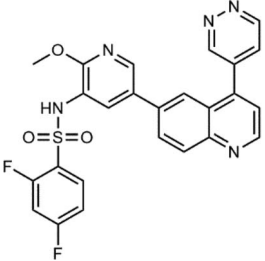
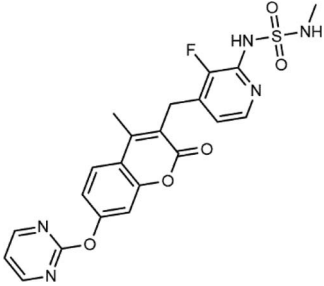
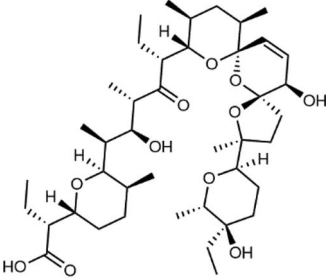
Pathogenic coronavirus invasion is a very complex biochemical process, which involves interactions between multiple viral proteins and human proteins. For example, the interactions between the spike protein of SARS-CoV-2 and the angiotensin-converting enzyme 2 of human cells, together with the cell surface serine protease TMPRSS2, play an important role in virus invasion. RNA synthesis of coronaviruses is performed by RNA-dependent RNA polymerase. 3C-like protease and papain-like protease

Table 2. Dual inhibitors in D3AI-CoV database

| Mol ID | Structure | Target and Activity | PMID |
|--------|---|--|---|
| ICV361 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 24.8 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 10.7 μM) 22 884 354 |
| ICV362 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 21.1 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 9.2 μM) 22 884 354 |
| ICV363 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 38.7 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 8.8 μM) 22 884 354 |
| ICV364 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 14.4 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 4.9 μM) 22 884 354 |
| ICV365 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 21.1 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 30 μM) 22 884 354 |
| ICV366 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 9.35 μM) | Papain-like protease (PLpro) SARS PL(IC ₅₀ = 24.1 μM); MERS PL(IC ₅₀ = 14.6 μM) 29 289 665 32 272 481 |
| ICV368 |  | 3C-like protease (3CLpro/Mpro) MERS 3CL(IC ₅₀ = 36.2 μM) | Papain-like protease (PLpro) MERS PL(IC ₅₀ = 42.1 μM) 28 112 000 |

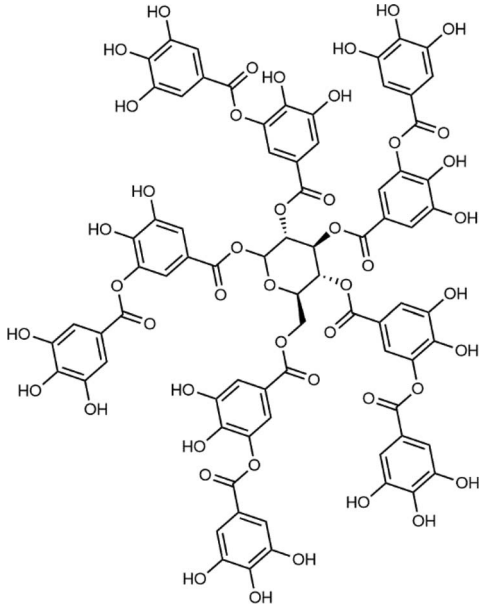
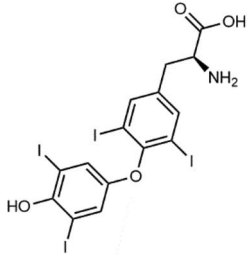
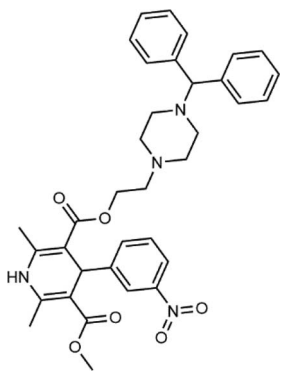
(Continued)

Table 2. Continued

| Mol ID | Structure | Target and Activity | PMID |
|--------|---|--|---|
| ICV371 |  | 3C-like protease (3CLpro/Mpro) SARS 3CL(IC ₅₀ = 30.2 μM) MERS 3CL(IC ₅₀ = 34.7 μM) | Papain-like protease (PLpro) MERS PL(IC ₅₀ = 48.8 μM) 28 112 000 |
| ICV403 |  | Spike protein (S protein) | Membrane protein (M protein) 17 704 516 17 560 666 |
| ICV646 |  | PI3K SARS-CoV-2 (IC ₅₀ = 0.014 μM) Caco2 | mTORC1/2 SARS-CoV-2 (IC ₅₀ = 0.014 μM) Caco2 32 877 642 |
| ICV648 |  | RAF SARS-CoV-2 (IC ₅₀ = 0.6 μM) Caco2 | MEK SARS-CoV-2 (IC ₅₀ = 0.6 μM) Caco2 32 877 642 |
| ICV693 |  | p-glycoprotein 1 | TMEM16 33 248 195 21 573 958 33 452 205 33 827 113 |

(Continued)

Table 2. Continued

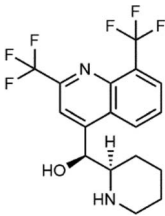
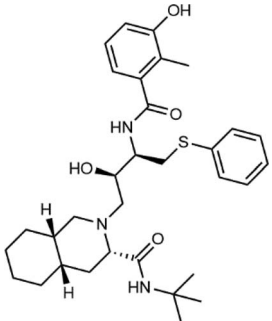
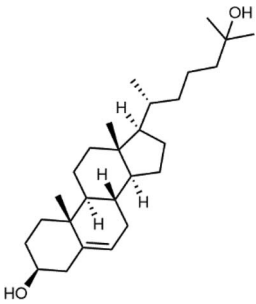
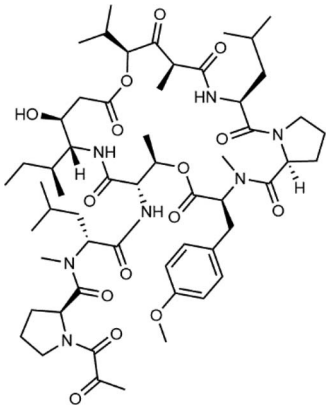
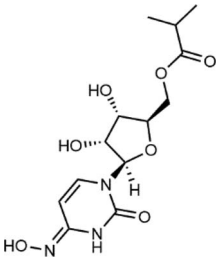
| Mol ID | Structure | Target and Activity | PMID | |
|--------|---|--|---|------------|
| ICV729 |  | 3C-like protease (3CLpro/Mpro) SARS-CoV-2 3CL(K _d = 1.1 μM, IC ₅₀ = 13.4 μM) | TMPRSS2 SARS-CoV-2 TMPRSS2(K _d = 1.77 μM, IC ₅₀ = 2.31 μM) | 33 415 017 |
| ICV734 |  | 3C-like protease (3CLpro/Mpro) SARS-CoV-2 3CL (IC ₅₀ = 19.2 μM) | Papain-like protease (PLpro) SARS-CoV-2 PL (IC ₅₀ = 15.3 μM) | 33 526 482 |
| ICV735 |  | 3C-like protease (3CLpro/Mpro) SARS-CoV-2 3CL (IC ₅₀ = 10.4 μM) | Papain-like protease (PLpro) SARS-CoV-2 PL (IC ₅₀ = 14.2 μM) | 33 526 482 |

are necessary for the reproduction and release of the coronavirus.

Dual inhibitors can target two targets in the virus life cycle, thereby inhibiting the coronaviruses more efficiently in principle. So the development of dual inhibitors is a novel strategy for the treatment of COVID-19. There are also some dual inhibitors against pathogenic coronaviruses in our database as shown in Table 2. For example, Wang *et al.* identified ICV729 as a potent dual inhibitor of both SARS-CoV-2 3C-like protease and TMPRSS2 [44], with IC₅₀ values of

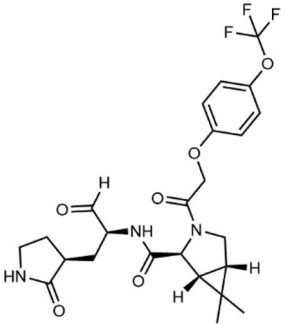
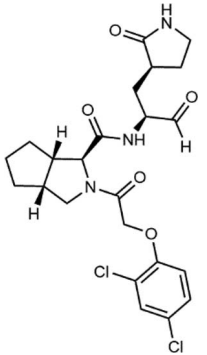
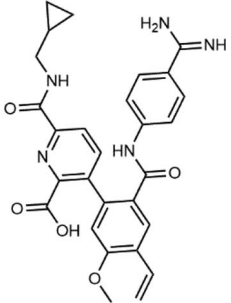
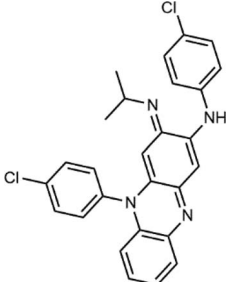
13.4 μM and 2.31 μM, respectively. ICV693 is effective against SARS-CoV-2 by inhibiting TMEM16 proteins [45], whereas previous research reported that p-glycoprotein was also the target of ICV693 [46]. So ICV693 may be a dual inhibitor against COVID-19. Besides, growth factor receptor (GFR) signaling is a central pathway necessary for SARS-CoV-2 replication. The dual phosphatidylinositol 3-kinase (PI3K)/mammalian target of rapamycin (mTOR) inhibitor ICV646 and dual rapidly accelerated fibrosarcoma (RAF)/mitogen-activated protein kinase kinase (MEK) inhibitor ICV648 can prevent SARS-CoV-

Table 3. Introduction of active compounds against SARS-CoV-2 *in vivo* in D3AI-CoV database

| Mol ID | Structure | Activity | PMID |
|--------|---|--|--------------------------|
| ICV487 |  | MERS($EC_{50} = 7.42 \mu\text{M}$) Vero E6 SARS($EC_{50} = 15.55 \mu\text{M}$) Vero E6 SARS-CoV-2($IC_{50} = 3.2 \mu\text{M}$) Vero E6 | 24 841 273 33 452 205 |
| ICV494 |  | SARS($EC_{50} = 0.048 \mu\text{M}$) Vero SARS-CoV-2($IC_{50} = 3.3 \mu\text{M}$) Vero E6 | 15 144 898 33 452 205 |
| ICV619 |  | SARS-CoV-2($EC_{50} = 3.68 \mu\text{M}$) Vero | 32 811 977 |
| ICV732 |  | SARS-CoV-2 ($IC_{50} = 1.62 \text{ nM}$) Pneumocyte-like cell SARS-CoV-2 ($IC_{50} = 0.7 \text{ nM}$) Vero E6 | 33 495 306 |
| ICV745 |  | Clinical trials (NCT04405570/NCT04405739) | 33 273 742 |

(Continued)

Table 3. Continued

| Mol ID | Structure | Activity | PMID |
|--------|---|--|------------|
| ICV754 |  | SARS-CoV-2 3CL (IC ₅₀ = 15.2 nM) SARS-CoV-2 (EC ₅₀ = 35.3 nM) Huh7 cell | 33 602 867 |
| ICV775 |  | SARS-CoV-2 3CL (IC ₅₀ = 17.2 nM) SARS-CoV-2 (EC ₅₀ = 31 nM) Huh7 cell | 33 602 867 |
| ICV835 |  | SARS-CoV-2 TMPRSS2 (IC ₅₀ = 0.19 μM) | 33 844 653 |
| ICV841 |  | SARS-CoV-2 (EC ₅₀ = 0.31 μM) Vero E6 | 33 727 703 |

2 replication by inhibiting GFR signaling [47]. ICV403 can target the spike protein and membrane protein so as to prevent the virus from invading. ICV361-ICV366, ICV368 and ICV371 can inhibit 3C-like protease and papain-like protease.

Active compounds against SARS-CoV-2 in vivo in D3AI-CoV database

Compounds that are active *in vivo* are more likely to be promising as drug candidates for the treatment

of COVID-19, which are shown in Table 3. Jan *et al.* found that ICV484, ICV497 and extracts of some herbal medicines were effective *in vivo* in the hamster model [38]. *In vivo* antiviral tests in a mouse model [37] showed that ICV619 is potent against SARS-CoV-2. ICV732 possessed anti-SARS-CoV-2 activity (IC₉₀ = 0.88 nM) *in vitro* [48]. The *in vivo* efficacy in two mouse models of SARS-CoV-2 infection and limited toxicity in cell culture of ICV732 indicate that it is a potential drug for the treatment of COVID-19. Besides, Cox *et al.* launched

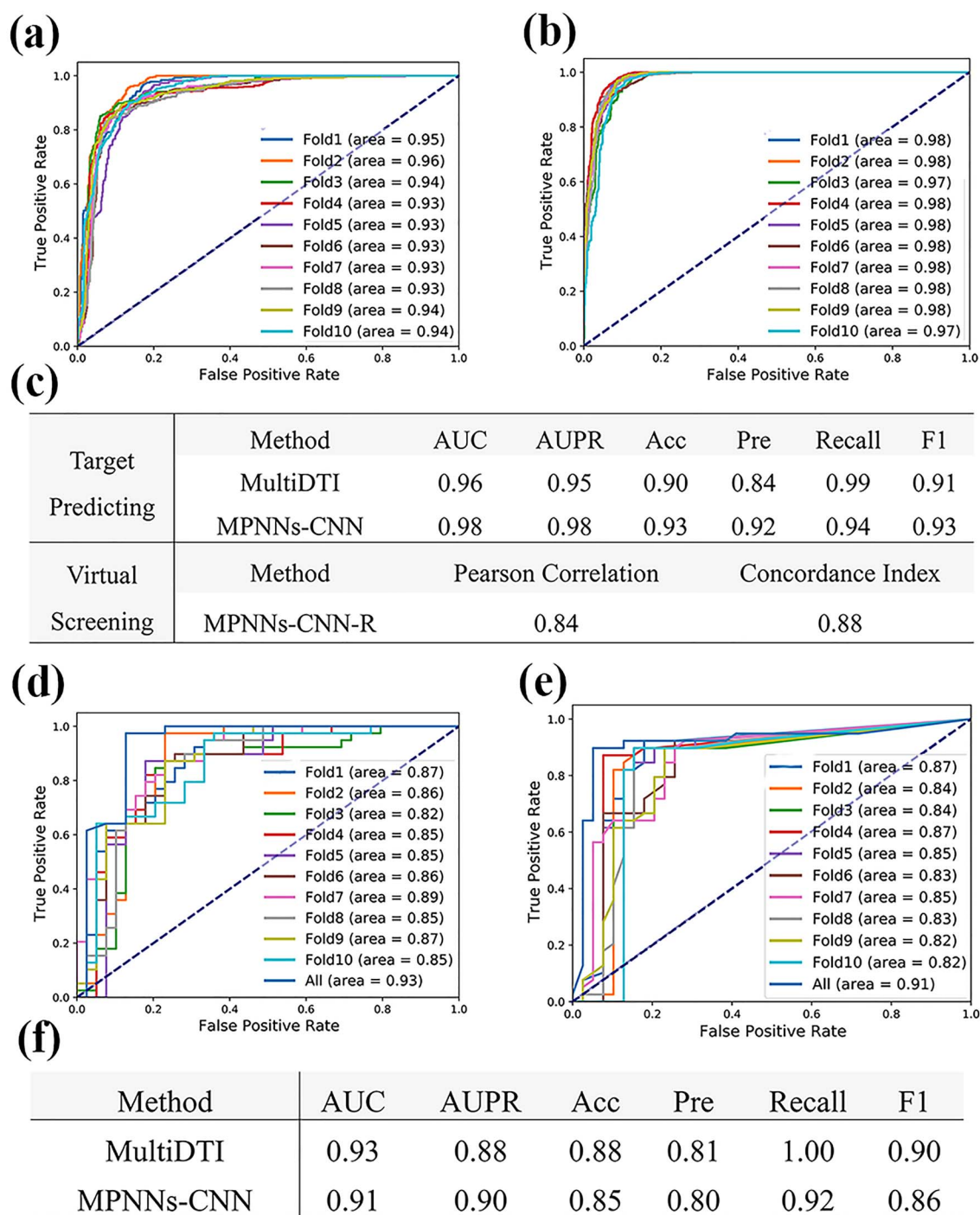


Figure 4. Model training and evaluation. (A) ROC curve of MultiDTI model by 10-fold cross-validation method. (B) ROC curve of MPNNs-CNN model by 10-fold cross-validation method. (C) Performance of the three DL-based models by 10-fold cross-validation method. (D) ROC curve of MultiDTI model on Test-78. (E) ROC curve of MPNNs-CNN model on Test-78. (F) Performance of MultiDTI and MPNNs-CNN models on Test-78.

a series of studies on ICV745, which is currently in phase II trials (NCT04405739) [49]. Qiao *et al.* designed and synthesized many 3C-like protease inhibitors [50]. ICV754 and ICV775 could reduce lung viral loads and lung lesions in a transgenic mouse model of SARS-CoV-2 infection. In addition, ICV835 and ICV841 could also inhibit SARS-CoV-2 infection in animal models [35, 51].

Training and testing of models

We developed two models, namely MultiDTI model (Figure 1A) and MPNNs-CNN model (Figure 1B), for target prediction, and one regression model based on MPNNs-CNN approach, namely MPNNs-CNN-R model, for virtual screening. About 10-fold cross-validation method was performed to train all our models. And the loss curves and accuracy curves during model

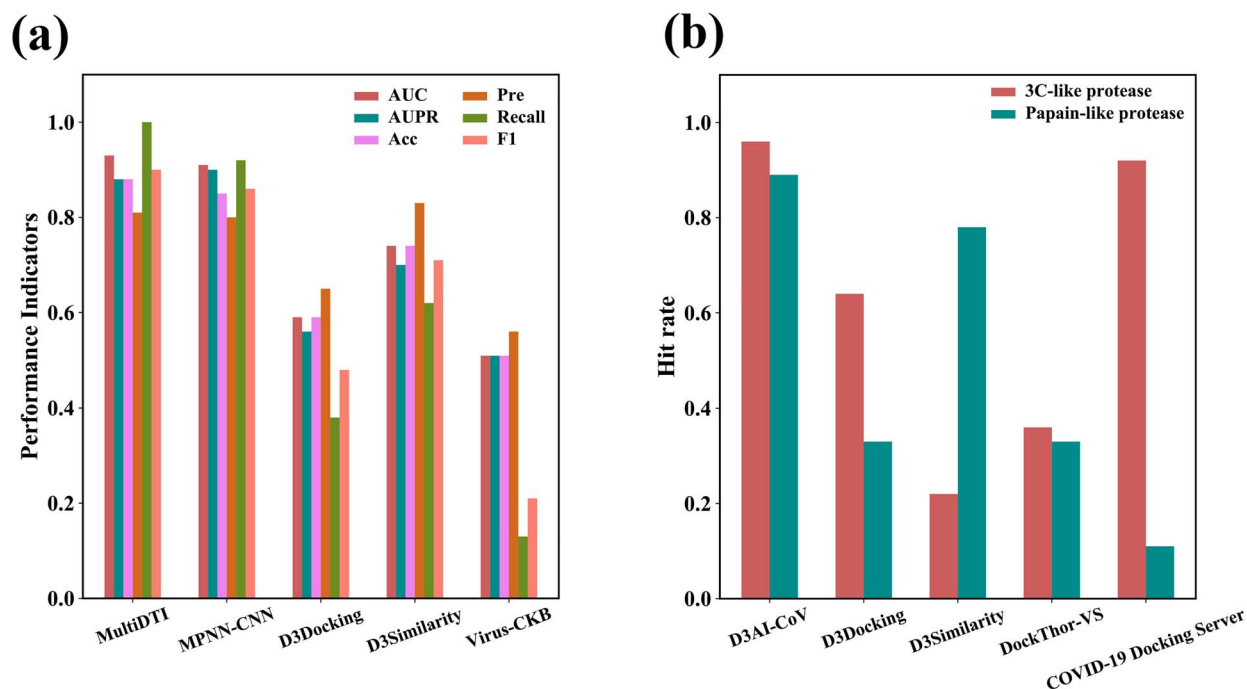


Figure 5. (A) The five performance indicators (AUC, AUPR, Acc, Pre, Recall and F1) of five methods for target prediction on Test-78. (B) The hit rate of the five methods for virtual screening against 3C-like protease and papain-like protease on Test-78.

training are shown in Figures S1–3. Finally, the model trained with all the drug–target pairs was used as the website backend of D3AI-CoV. The ROC curves are shown in Figure 4A and B. The AUCs of the two models for target prediction are 0.93–0.96 and 0.97–0.98. Other performance indicators, including AUPRs (0.88–0.95 and 0.95–0.98), accuracy (Acc) (0.86–0.9 and 0.92–0.94), precision (Pre) (0.79–0.91 and 0.89–0.93), recall (0.81–0.99 and 0.93–0.97) and F1 score (F1) (0.86–0.91 and 0.92–0.95) suggested the strong target prediction ability of the MultiDTI model and MPNNs-CNN model (Figure 4C; also see Table S5). Pearson correlation and concordance index of MPNNs-CNN-R model for virtual screening by 10-fold cross-validation method are 0.8–0.84 and 0.87–0.88, respectively (Table S6).

The external dataset, namely Test-78, was used to further test the generalization performance of the MultiDTI model and MPNNs-CNN model against COVID-19. The ROC curves are shown in Figure 4D and E. The AUCs of the two models trained by 10-fold cross-validation method for target prediction are 0.82–0.89 and 0.82–0.87. Other performance indicators, including AUPRs (0.75–0.89 and 0.72–0.79), Acc (0.72–0.82 and 0.74–0.85), Pre (0.72–0.83 and 0.81–0.9), Recall (0.64–1.00 and 0.64–0.85) and F1 (0.7–0.84 and 0.71–0.85) values of the MultiDTI and MPNNs-CNN models suggested their strong target prediction ability (Figure 4F; also see Table S7). In addition, the models trained with all data have stronger predictive performance (Table S7). Accordingly, the DL-based models have a strong predictive ability for target prediction and virtual screening against COVID-19.

Comparison of D3AI-CoV with other methods

There are three webservers publicly available for target prediction against COVID-19, which are D3Docking, D3Similarity and Virus-CKB, whereas there are four websites for virtual screening, which are D3Docking, D3Similarity, DockThor-VS and COVID-19 Docking Server. We used the external dataset Test-78 to evaluate all the webservers for target prediction and virtual screening against COVID-19. Figure 5A and B summarizes the comparison results between our newly constructed DL-based models and other methods (Tables S8 and S9).

For target predicting, a prediction is regarded as correct if the top 10 predicted targets contain the real target. In our DL-based models, a prediction is correct if the predicted probability is greater than 0.5. Next, we compared MultiDTI, MPNNs-CNN, D3Docking, D3Similarity and Virus-CKB for target prediction. All ligands of Test-78 were used for these five methods for target prediction. Based on the prediction results, we counted the correct and incorrect numbers of prediction and calculated the AUC, AUPR, Acc, Pre, Recall and F1, which were used to compare various methods. After testing, the AUCs (0.93 and 0.91), AUPRs (0.88 and 0.9), Acc (0.88 and 0.85), Pre (0.81 and 0.8), Recall (1 and 0.92) and F1 (0.9 and 0.86) of the two DL-based models outperform those of D3Docking (0.59, 0.56, 0.59, 0.65, 0.38 and 0.48), D3Similarity (0.74, 0.7, 0.74, 0.83, 0.62 and 0.71) and Virus-CKB (0.51, 0.51, 0.51, 0.56, 0.13 and 0.21) (Figure 5A; also see Table S8). Besides, DL-based models are much faster than D3Docking, D3Similarity and Virus-CKB. More importantly, the MultiDTI model correctly

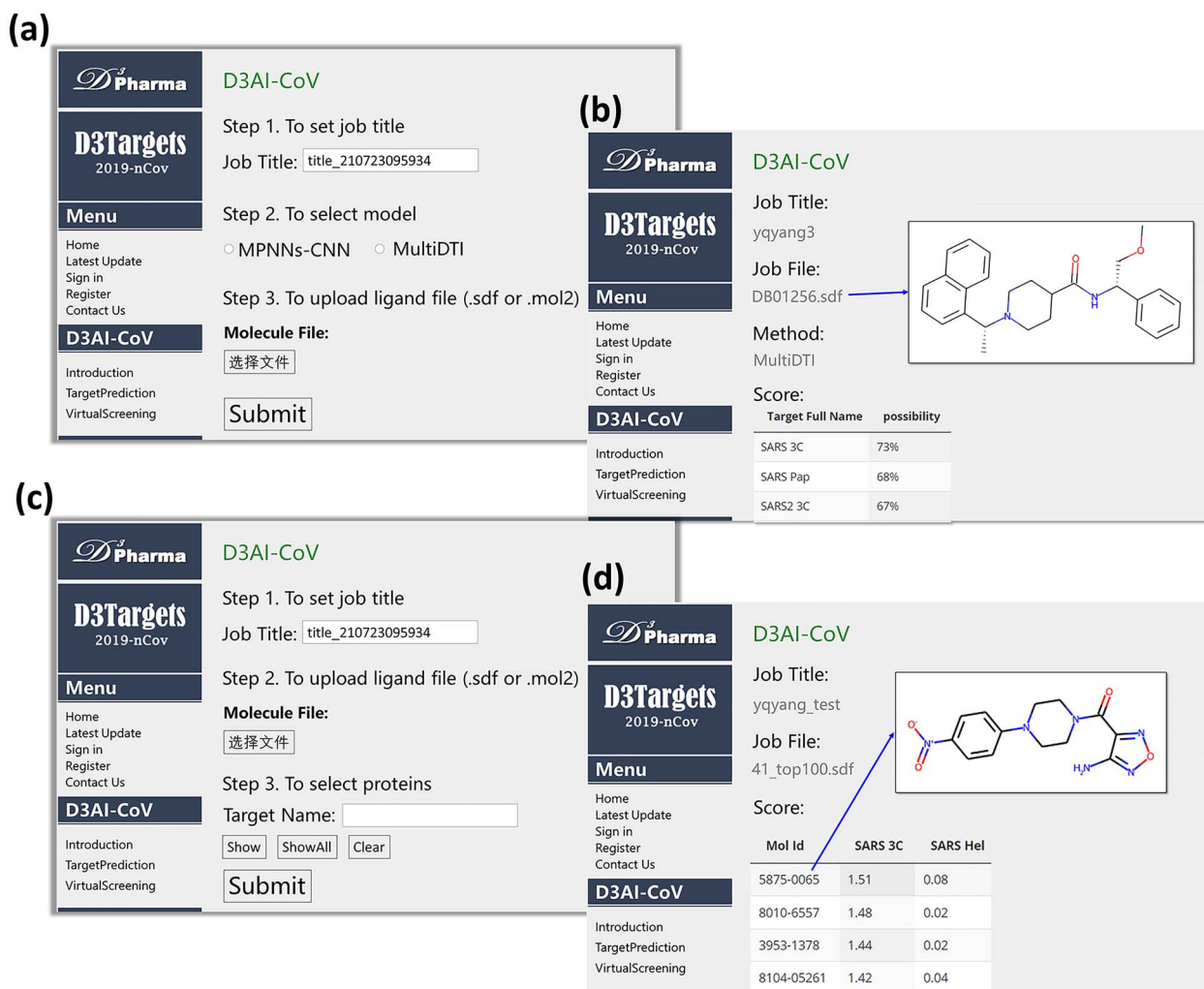


Figure 6. Graphical interface for input and output of D3AI-CoV. (A) Graphical interface for input of the target prediction module of D3AI-CoV. (B) Graphical interface for output of the target prediction module of D3AI-CoV. (C) Graphical interface for input of the virtual screening module of D3AI-CoV. (D) Graphical interface for output of the virtual screening module of D3AI-CoV.

predicts two completely new protein targets and their molecules, which indicates MultiDTI model has great expandability.

For virtual screening, inhibitors for the 3C-like and papain-like proteases account for the two largest proportions in Test-78, so we used the Test-78 to perform virtual screening against the two targets for comparison. The hit rate was used as a criterion for evaluating the performance of different virtual screening methods. After testing, the results indicate the hit rates of MPNNs-CNN-R are 0.96 and 0.89 for 3C-like protease and papain-like protease, respectively, whereas that of other methods are 0.22–0.92 and 0.11–0.78, indicating that the new MPNNs-CNN-R model is in general much better than other methods (Figure 5B; also see Table S9).

In summary, D3AI-CoV shows great predictive performance both on the validation set and on a completely independent external test set. More importantly, when compared with other anti-COVID-19 webservers, the prediction accuracy of D3AI-CoV is much higher

than other docking-based or similarity-based methods. And the efficiency of D3AI-CoV is also higher (5–10 s for a job by D3AI-CoV versus 5–15 min for a job by D3Similarity and 1–2 h for a job by D3Docking). All the results demonstrated that D3AI-CoV has great advantage in comparison with other webservers in terms of prediction accuracy and prediction speed.

Input and output

D3AI-CoV is provided free of charge for users via the web server. For target prediction, as shown in Figure 6A, the users can set the task title and select a prediction model in the target prediction interface. And then they can submit a small molecule in sdf or mol2 file format. The small molecule will be converted to canonical SMILES. Two DL-based models will be used to predict the target of the input molecule according to the SMILES. Usually, predicting process will last for a few minutes after the beginning of the calculation before the output result is

returned. Therefore, D3AI-CoV is faster than the conventional structure- and ligand-based approaches. Finally, as shown in Figure 6B, the top-ranked targets will be provided on the webpage.

For virtual screening, as shown in Figure 6C, the users can upload a small molecule library in sdf or mol2 file format and select one or two target(s). All small molecules in the library will be converted to canonical SMILES. The regression model will perform the virtual screening for the small molecule library. After finishing the task, as shown in Figure 6D, the top-ranked ligands and their scores will be presented on the webpage.

Conclusion

Target prediction and virtual screening are two important issues for discovering new drugs and lead optimization. Based on protein structure and ligand information, we have developed and continuously updated a webserver, D3Targets-2019-nCoV, for target prediction and virtual screening since the COVID-19 outbreak. In this work, with the latest updated databases of both active compounds and target proteins, we developed two classification DL-based models for target prediction and a regression DL-based model for virtual screening for discovering hits against COVID-19. The results showed that the predictive abilities of the DL-based models on the external test set are significantly stronger than D3Docking, D3Similarity and other methods. The prediction speed of the DL-based models is also much faster than other methods. We hope D3AI-CoV will be helpful to the development of anti-COVID-19 drugs.

Key Points

- Identifying effective drug targets and developing effective drugs accordingly to cure COVID-19 are of great importance.
- We developed D3AI-CoV, a DL-based platform for target prediction and virtual screening for discovering anti-COVID-19 drugs.
- The MultiDTI model and The MPNNs-CNN model can be used to predict targets for active compounds.
- The MPNNs-CNN-R model can be used to perform virtual screening.
- D3AI-CoV is available at free as a web application at <http://www.d3pharma.com/D3Targets-2019-nCoV/D3AI-CoV/index.php>

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Supplementary Data

Supplementary data are available online at <https://academic.oup.com/bib>

Funding

This work was supported by National Key Research and Development Program of China (2016YFA0502301), Natural Science Foundation of Shanghai (21ZR1475600) and Natural Science Foundation of China (U19A2067).

References

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;**579**:270–3.
2. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020;**579**:265–9.
3. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020;**382**:727–33.
4. Weekly epidemiological update on COVID-19 - 28 December 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---28-december-2021> (28 December 2021, date last accessed).
5. Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* 2021;**19**:409–24.
6. Garrett ME, Galloway J, Chu HY, et al. High-resolution profiling of pathways of escape for SARS-CoV-2 spike-binding antibodies. *Cell* 2021;**184**:2927–38.
7. Shi Y, Zhang X, Mu K, et al. D3Targets-2019-nCoV: a webserver for predicting drug targets and for multi-target and multi-site based virtual screening against COVID-19. *Acta Pharm Sin B* 2020;**10**:1239–48.
8. Chen Z, Zhang X, Peng C, et al. D3Pockets: a method and web server for systematic analysis of protein pocket dynamics. *J Chem Inf Model* 2019;**59**:3353–8.
9. Yang Y, Zhu Z, Wang X, et al. Ligand-based approach for predicting drug targets and for virtual screening against COVID-19. *Brief Bioinform* 2021;**22**:1053–64.
10. Kong R, Yang G, Xue R, et al. COVID-19 Docking server: a meta server for docking small molecules, peptides and antibodies against potential targets of COVID-19. *Bioinformatics* 2020;**36**:5109–11.
11. Xu C, Ke Z, Liu C, et al. Systemic in silico screening in drug discovery for coronavirus disease (COVID-19) with an online interactive web server. *J Chem Inf Model* 2020;**60**:5735–45.
12. Guedes IA, Costa LSC, Dos Santos KB, et al. Drug design and repurposing with DockThor-VS web server focusing on SARS-CoV-2 therapeutic targets and their non-synonym variants. *Sci Rep* 2021;**11**:5543.
13. Feng Z, Chen M, Liang T, et al. Virus-CKB: an integrated bioinformatics platform and analysis resource for COVID-19 research. *Brief Bioinform* 2020;**22**:882–95.
14. Bai Q, Tan S, Xu T, et al. MolAICal: a soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. *Brief Bioinform* 2020;**22**:bbaa161.
15. Kc GB, Bocci G, Verma S, et al. A machine learning platform to estimate anti-SARS-CoV-2 activities. *Nat Mach Intel* 2021;**3**:527–35.

16. Artificial intelligence for drug discovery. <https://www.atomwise.com/> (28 June 2021, date last accessed).
17. Ibm Watson. <https://www.ibm.com/watson> (28 June 2021, date last accessed).
18. Smalley E. AI-powered drug discovery captures pharma interest. *Nat Biotechnol* 2017;**35**:604–5.
19. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. *Cell* 2020;**180**:688–702.
20. Zhavoronkov A, Ivanenkov YA, Aliper A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol* 2019;**37**:1038–40.
21. Ton AT, Gentile F, Hsing M, et al. Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds *Mol. Inform* 2020;**39**:e2000028.
22. Wang S, Sun Q, Xu Y, et al. A transferable deep learning approach to fast screen potential antiviral drugs against SARS-CoV-2. *Brief Bioinform* 2021;**22**:bbab211. <https://doi.org/10.1093/bib/bbab211>.
23. Beck BR, Shin B, Choi Y, et al. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Comput Struct Biotechnol J* 2020;**18**:784–90.
24. Kadioglu O, Saeed M, Greten HJ, et al. Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. *Comput Biol Med* 2021;**133**:104359.
25. Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *The Lancet* 2020;**395**:e30–1.
26. Zhang DH, Wu KL, Zhang X, et al. In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus. *J Integr Med* 2020;**18**:152–8.
27. Zhou D, Xu Z, Li W, et al. MultiDTI: drug-target interaction prediction based on multi-modal representation learning to bridge the gap between new chemical entities and known heterogeneous network. *Bioinformatics* 2021;**37**:4485–92.
28. Liu K, Sun X, Jia L, et al. Chemi-net: a molecular graph convolutional network for accurate drug property prediction. *Int J Mol Sci* 2019;**20**:3389.
29. Gilmer J, Schoenholz SS, Riley PF, et al. Neural message passing for quantum chemistry. *Int Conf Mach Learn* 2017;**70**:1263–72.
30. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE* 1998;**86**:2278–324.
31. O'Boyle NM, Banck M, James CA, et al. Open babel: an open chemical toolbox. *J Chem* 2011;**3**:33.
32. UniProt C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 2021;**49**:D480–9.
33. Landrum G. et al. RDKit: open-source Cheminformatics Software (10 May 2020, date last accessed); <https://www.rdkit.org/>.
34. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018;**46**:D1074–82.
35. Yuan S, Yin X, Meng X, et al. Clofazimine broadly inhibits coronaviruses including SARS-CoV-2. *Nature* 2021;**593**:418–23.
36. Rosenke K, Hansen F, Schwarz B, et al. Orally delivered MK-4482 inhibits SARS-CoV-2 replication in the Syrian hamster model. *Nat Commun* 2021;**12**:2295.
37. Zu S, Deng YQ, Zhou C, et al. 25-Hydroxycholesterol is a potent SARS-CoV-2 inhibitor. *Cell Res* 2020;**30**:1043–5.
38. Jan JT, Cheng TR, Juang YP, et al. Identification of existing pharmaceuticals and herbal medicines as inhibitors of SARS-CoV-2 infection. *Proc Natl Acad Sci U S A* 2021;**118**:e2021579118.
39. Good SS, Westover J, Jung KH, et al. AT-527, a double prodrug of a guanosine nucleotide analog, is a potent inhibitor of SARS-CoV-2 in vitro and a promising oral antiviral for treatment of COVID-19. *Antimicrob Agents Chemother* 2021;**65**:e02479–20.
40. Svenningsen EB, Thyrsted J, Blay-Cadanet J, et al. Ionophore antibiotic X-206 is a potent inhibitor of SARS-CoV-2 infection in vitro. *Antiviral Res* 2020;**185**:104988.
41. Yuan S, Chu H, Huang J, et al. Viruses harness YxxØ motif to interact with host AP2M1 for replication: a vulnerable broad-spectrum antiviral target. *Sci Adv* 2020;**6**:eaba7910.
42. Bailly C, Vergoten G. Glycyrrhizin: an alternative drug for the treatment of COVID-19 infection and the associated respiratory syndrome? *Pharmacol Ther* 2020;**214**:107618.
43. Cinatl J, Morgenstern B, Bauer G, et al. Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus. *Lancet* 2003;**361**:2045–6.
44. Wang SC, Chen Y, Wang YC, et al. Tannic acid suppresses SARS-CoV-2 as a dual inhibitor of the viral main protease and the cellular TMPRSS2 protease. *Am J Cancer Res* 2020;**10**:4538–46.
45. Braga L, Ali H, Secco I, et al. Drugs that inhibit TMEM16 proteins block SARS-CoV-2 Spike-induced syncytia. *Nature* 2021;**594**:88–93.
46. Kim WK, Kim JH, Yoon K, et al. Salinomycin, a p-glycoprotein inhibitor, sensitizes radiation-treated cancer cells by increasing DNA damage and inducing G2 arrest. *Invest New Drugs* 2012;**30**:1311–8.
47. Klann K, Bojkova D, Tascher G, et al. Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol Cell* 2020;**80**:164–74.
48. White KM, Rosales R, Yildiz S, et al. Plitidepsin has potent preclinical efficacy against SARS-CoV-2 by targeting the host protein eEF1A. *Science* 2021;**371**:926–31.
49. Cox RM, Wolf JD, Plemper RK. Therapeutically administered ribonucleoside analogue MK-4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. *Nat Microbiol* 2021;**6**:11–8.
50. Qiao J, Li YS, Zeng R, et al. SARS-CoV-2 M(pro) inhibitors with antiviral activity in a transgenic mouse model. *Science* 2021;**371**:1374–8.
51. Sun YJ, Velez G, Parsons DE, et al. Structure-based phylogeny identifies Avoralstat as a TMPRSS2 inhibitor that prevents SARS-CoV-2 infection in mice. *J Clin Invest* 2021;**131**:e147973.