



Research article

Drug repositioning of COVID-19 based on mixed graph network and ion channel

Xianfang Wang^{1,2,*}, Qimeng Li², Yifeng Liu¹, Zhiyong Du¹ and Ruixia Jin³

¹ Henan Institute of Technology, Xinxiang 453003, China

² College of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China

³ SanQuan Medical College, Xinxiang 453003, China

* **Correspondence:** Email: 2wangfang@163.com; Tel: +86 13598659781.

Abstract: Research on the relationship between drugs and targets is the key to precision medicine. Ion channel is a kind of important drug targets. Aiming at the urgent needs of corona virus disease 2019 (COVID-19) treatment and drug development, this paper designed a mixed graph network model to predict the affinity between ion channel targets of COVID-19 and drugs. According to the simplified molecular input line entry specification (SMILES) code of drugs, firstly, the atomic features were extracted to construct the point sets, and edge sets were constructed according to atomic bonds. Then the undirected graph with atomic features was generated by RDKit tool and the graph attention layer was used to extract the drug feature information. Five ion channel target proteins were screened from the whole SARS-CoV-2 genome sequences of NCBI database, and the protein features were extracted by convolution neural network (CNN). Using attention mechanism and graph convolutional network (GCN), the extracted drug features and target features information were connected. After two full connection layers operation, the drug-target affinity was output, and model was obtained. Kiba dataset was used to train the model and determine the model parameters. Compared with DeepDTA, WideDTA, graph attention network (GAT), GCN and graph isomorphism network (GIN) models, it was proved that the mean square error (MSE) of the proposed model was decreased by 0.055, 0.04, 0.001, 0.046, 0.013 and the consistency index (CI) was increased by 0.028, 0.016, 0.003, 0.03 and 0.01, respectively. It can predict the drug-target affinity more accurately. According to the prediction results of drug-target affinity of SARS-CoV-2 ion channel targets, seven kinds of small molecule drugs acting on five ion channel targets were obtained, namely SCH-47112, Dehydroaltenusin, alternariol 5-o-sulfate, LPA1 antagonist 1, alternariol, butin, and AT-9283. These drugs provide a reference for drug repositioning

and precise treatment of COVID-19.

Keywords: COVID-19; graph neural network; attention mechanism; drug-target affinity; drug repositioning

1. Introduction

COVID-19 has caused huge losses and led to great health and economic burdens to the world. Currently, drug discovery for COVID-19 caused by SARS-CoV-2 is still ongoing. Accurate prediction of drug-target interactions (DTI) is crucial for drug discovery. Recently, deep learning models have been used for DTI prediction and drug repurposing of COVID-19. Huang et al. presents DeepPurpose, a comprehensive deep learning library for DTI prediction [1]. Number of drugs such as remdesivir, favipiravir, lopinavir have shown inhibitory effects against the SARS-CoV2 in-vitro as well as in clinical conditions [2]. The mainly targets of SARS-CoV-2 include S protein, angiotensin converting enzyme 2 (ACE2), transmembrane protease serine 2 (TMPRSS2), 3CLpro, RNA-dependent RNA polymerase (RdRp), etc. Zhou's team of Westlake University analyzes the electron microcopy structure of complexes between S protein and ACE2 [3]. The ACE2 receptor recognized by the S protein and TMPRSS2 have also been identified as potential targets, they are both key molecules in the initial stage of viral invasion of the host. There are already several drugs that can target ACE2, including captopril, propofol and tiliquinol, and some polypeptides and antibody drugs, but their anti-SARS-CoV-2 activity is unknown. Egyptian scientists use high-resolution PLpro crystals of SARS-CoV as a template to predict the structure, and predict the binding ability of several anti-SARS-CoV PLpro and anti-HCV NS3 drugs to novel coronavirus PLpro using docking method [4]. Similarly, some studies have screened 2525 FDA-listed drugs in the ZINC database using the structure of homology modeling, and found 16 molecules with strong affinity [5]. Chinese scientists obtains 7 molecules that could bind PLpro from traditional natural compounds by using ADME filtration and molecular docking, and search Chinese medicines containing these ingredients [6]. However, these above methods have shortcoming in the drug-target interaction prediction Accurate drug-target interactions plays a key role in drug repositioning, which can not only deepen the understanding of drug action, but also reposition drugs from the perspective of pharmacology.

Most drug targets come from enzymes, G-protein-coupled receptors (GPCRs), ion channels and nuclear receptors, which account for 88% of the total drug targets [7,8]. Ion channels are composed of a special class of proteins, which are assembled and embedded on the cell membrane to form a pathway that allows the rapid transport of inorganic ions across the cell membrane. Ion channels are involved in a variety of important biological functions, such as cell excitation, muscle contraction, gland secretion, nervous system development, and regulation of gene expression. The occurrence and development of many diseases (such as neuropathic pain, arrhythmia, hypertension, etc.) are related to the dysfunction of ion channels. Therefore, ion channels have become one of the classic targets for pharmaceutical companies to study.

Therefore, a DTI prediction model for COVID-19 based on graph network and ion channels was proposed in this paper. Firstly, the atomic features were extracted, and the undirected graph was generated by RDKit tool. The sequence of each target protein was regarded as a string recognized by computer, and the features of the sequences were extracted by convolution neural network. Attention

mechanism and graph convolution network were used to predict the drug-target affinity and screen drugs that act on the target proteins of SARS-CoV-2.

The main contributions of this paper are as follows: (1) graph representation learning of drugs based on attention mechanism to achieve drug feature extraction; (2) a mixed graph network model was constructed to predict drug-target affinity; (3) seven drugs acting on SARS-CoV-2 target proteins were found from the prediction results, which provided a reference for drug repositioning.

2. Materials and methods

2.1. Graph representation of drugs

SMILES, developed by David Weininger is a standard language for expressing molecular structures with ASCII strings. Using this language, chemical formulas can be converted and stored in the form of text in computers [9]. For each drug, according to its SMILES code, RDKit is used to construct the molecular graph, which serves as the input of the graph convolutional network. The nodes of the graph represent the features of the drug atoms, and the edges represent the bonds between the atoms. The feature vector of a drug atom was composed of five features: atom type, atomic level, total number of hydrogen, implicit value of atom, and whether the atom is aromatic.

2.2. Ion channel target proteins of COVID-19

The amino acid sequences of SARS-CoV-2 ion channel target proteins were extracted from the whole SARS-CoV-2 genome sequences (accession NC_045512.2) of the National Center for Biotechnology Information (NCBI) database. They were 3C-like proteinase (YP_009725301.1), RNA-dependent RNA polymerase (YP_009725307), helicase (YP_009725308.1), 3'-to-5' exonuclease (YP_009725309.1) and 2'-O-ribose methyltransferase (YP_009725311.1), respectively [10].

2.3. Graph convolutional network

GCN is a graph-based convolutional network model proposed in 2017 [11,12]. The formula is as follows:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (1)$$

where $H^{(l+1)}$ is the output matrix of layer l , $W^{(l)}$ is the weight matrix of layer l , $H^{(l)}$ is the input matrix. σ is the activation function, and the ReLU function is usually used. \tilde{D} is the degree matrix, $\tilde{A} = A + I$, where A is the adjacency matrix of node i in the graph, and I is the identity matrix with the same shape as A , i.e., the physical meaning of which is to consider its own features information when calculating node features.

2.4. Graph attention network

GAT proposes an attention-based architecture to learn hidden representations of nodes in a graph by applying a self-attention mechanism [13,14]. Its principle is to obtain the attention coefficient of a

certain node and its neighbor nodes, and then perform a weighted summation of all surrounding neighbor nodes according to the attention coefficient, and then obtain the feature value of this node after the neighbor nodes are aggregated. The calculation formula of the attention coefficient is as follows:

$$a_{ij} = \frac{\exp(\text{LeakyReLU}(A^T[W\vec{H}_i \| W\vec{H}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(A^T[W\vec{H}_i \| W\vec{H}_k]))} \quad (2)$$

where vector \vec{H}_i and vector \vec{H}_j are transformed with the learnable parameter W and connected, and then multiplied by a learnable parameter A^T to obtain the attention of node i and node j . LeakyReLU function is used as an activation function and finally get the attention coefficients of node i and j . The node feature formula is as follows:

$$H'_i = \sigma(\sum_{j \in N_i} a_{ij} W\vec{H}_j) \quad (3)$$

After the training of the attention coefficient, all neighbor node vectors \vec{H}_j of node i are aggregated with the attention coefficient, and then W is used. In this way, the features of node i and all its neighbor nodes are aggregated and taken as the output features of the node. The formula of multi-head attention is as follows:

$$H'_i = \parallel_{k=1}^K \sigma(\sum_{j \in N_i} a_{ij} W\vec{H}_j) \quad (4)$$

where K is the number of attention heads. When calculating the attention coefficient of node i and node j , the relationship between them is considered for several times. The number of considerations is the head parameter and can be controlled artificially. This multi-attention mechanism is improved the correlation between nodes.

2.5. Graph isomorphism network

GIN is an improved GCN model proposed by Keyulu et al. [15] in 2019. The formula is as follows:

$$h_o^{(k)} = \text{MLP}^{(k)}((1 + \varepsilon^{(k)})h_i^{(k-1)} + \sum_{u \in n_i} h_u^{(k-1)}) \quad (5)$$

where multilayer perceptron (MLP) is a multi-layer perceptron, h_i is the input feature vector of the k -layer perceptron, h_i is the output vector, h_u is the feature vector of adjacent nodes, and ε is the offset, which can be set manually.

2.6. Drug-target affinity prediction model based on mixed graph network

In this paper, we proposed a mixed graph network model based on GAT and GCN for regression prediction of drug-target affinity. The model framework is shown in Figure 1.

Firstly, SMILES was converted to a molecular graph, then a deep learning algorithm was adopted to learn graph representation. The advantages of GCN and GAT model were used to improve the accuracy of feature extraction. Protein sequence was encoded and embedded, then several CNN layers were used to learn sequence representation. Then two representation features were cascaded and

undergo two fully connected layers, and finally the drug-target affinity was output.

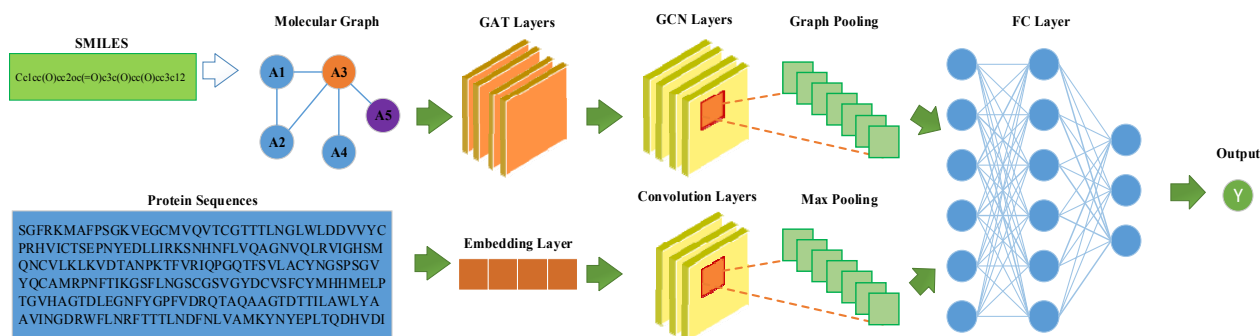


Figure 1. Framework of drug-target affinity prediction model based on mixed graph network.

3. Experiments and results

3.1. Dataset and feature extraction

To compare with DeepDTA [16] and WideDTA [17], we ran our model on the same dataset used in these works. KIBA dataset including 2116 drugs and 229 targets and 118,254 binding affinities. The affinity value ranges from 0.0 to 17.2. 98545 binding affinities were used for training and 19709 binding affinities were used for testing the models. Datasets used in this paper are publicly available at <http://www.ddcnn.wang/CovidData>. GetNumAtoms method was used to obtain the total number of atoms in drug molecules, GetBonds method was used to traverse the bonds, and then GetBeginAtom method and GetEndAtom method were used respectively to obtain chemical index values of the initial atomic and end the index values of the atom, after that the information for undirected graph nodes was created. In order to improve the computational performance and reduce the waste of computational resources, information such as drug molecular graph, target sequence and drug-target affinity was packaged and combined into a data unit. Therefore, the training process can be performed only by directly reading the packaged data unit.

3.2. Evaluation index

The same evaluation indexes as WideDTA were used in this paper to judge the prediction ability of the model, namely MSE and CI [18].

MSE is the mean of the sum of squares of the difference between the predicted value and the true value. The smaller the calculated result is, the more accurate the predicted value is, which is defined as follows:

$$\text{MSE}(y, f) = \frac{1}{m} \sum_{i=1}^m (y_i - f_i)^2 \quad (6)$$

CI is used to judge the accuracy of the model. The more the calculation result tends to 1, the higher the accuracy of the trained model is. It is defined as follows:

$$CI(y, f) = \frac{1}{z} \sum_{y_i > y_j}^m h(f_i - f_j) \quad (7)$$

where, y is the true label of drug-target affinity, f is the predicted value of drug-target affinity, and the $h(n)$ is defined as follows:

$$h(n) = \begin{cases} 1 & \text{if } n > 0 \\ 0.5 & \text{if } n = 0 \\ 0 & \text{if } n < 0 \end{cases} \quad (8)$$

3.3. Experimental environments

The experimental machine was configured as Intel(R) Core(TM) i7-6800K at 2.30GHz, with 64GB memory of CPU and NVIDIA Titan XP of GPU. PyTorch was used to implement the model. The experimental environments were Python 3.7, CUDA 10.1, Keras 2.2.5, Numpy 1.19.5, and PyTorch 1.2.0. Adaptive moment estimation (ADAM) optimizer was applied to train the model, with a default learning rate of 0.0005 and a training batch size of 512. The mixed model integrates the advantages of the GAT and GCN models in series to obtain better model performance. The output channel of the GAT layer is 78, the number of attention nodes is 10. And the output channel of the GCN layer is 780. The results were mapped to a 128D feature vectors in the output layer.

3.4. Results analysis

3.4.1. Prediction results of drug-target affinity

To verify the performance of the proposed model, DeepDTA, WideDTA, GCN, GAT and GIN models were compared with the proposed model under the same dataset and conditions. The comparison results are shown in Table 1.

Table 1. Evaluation index table of model prediction results.

Model	MSE	CI
DeepDTA	0.194	0.863
WideDTA	0.179	0.875
GCN	0.140	0.888
GAT	0.185	0.861
GIN	0.152	0.881
GAT_GCN_mix	0.139	0.891

By comparing with DeepDTA, WideDTA, GAT, GCN and GIN models, it is proved that MSE of the proposed model is decreased by 0.055, 0.04, 0.001, 0.046, 0.013, and CI is increased by 0.028, 0.016, 0.003, 0.03, 0.01, respectively. It can be seen that the performance of the mixed network model is better than that of previous model, indicating that in the process of graph feature extraction, the mixed network model combines the advantages of the GCN and the GAT to obtain a certain effect.

In the KIBA dataset, the GCN model overall is slightly better than the GAT model, indicating that convolution is a good choice to solve the problem of graphs. If the graph attention network is

introduced before the GCN model, the prediction effect will be better than that of pure convolution network. Therefore, continuous attempts and innovation are conducive to finding better solutions.

3.4.2. Visualization of prediction results of drug-target interaction network

The drug-target interaction network was drawn using Cytoscape software to realize the visualization of the prediction results of the model. There were 10555 interaction relationships in the KIBA dataset for five targets of SARS-CoV-2, and the first 87 interaction relationships were screened under the binding affinity > 13.5 condition to construct drug-target interaction network. As shown in Figure 2, red circles represent drugs and green diamonds represent targets. The degree of correlation between the drug and the target is represented by the color and size of the drug. The lighter the color and the smaller the shape, the worse the correlation; the darker the color and the larger the shape, the stronger the correlation. Through the diagram, the interaction relationships between drugs and targets can be understood more intuitively.

The drugs in Tables 2–6 are intersected to obtain 7 drugs targeting 5 ion channel target proteins of SARS-CoV-2, which are CHEMBL483525, CHEMBL291126, CHEMBL519982, CHEMBL495727, CHEMBL83790, CHEMBL520144 and CHEMBL483526, and their corresponding drug molecules are alternariol 5-O-sulfate, SCH-47112, alternariol, AT-9283, LPA1 antagonist 1, dehydroaltenusin, butin. The affinities and average values of the 7 drugs screened by the model established in this paper for 5 target proteins of SARS-CoV-2 are shown in Table 7, and the Venn diagram of the targets and the drugs is shown in Figure 3.

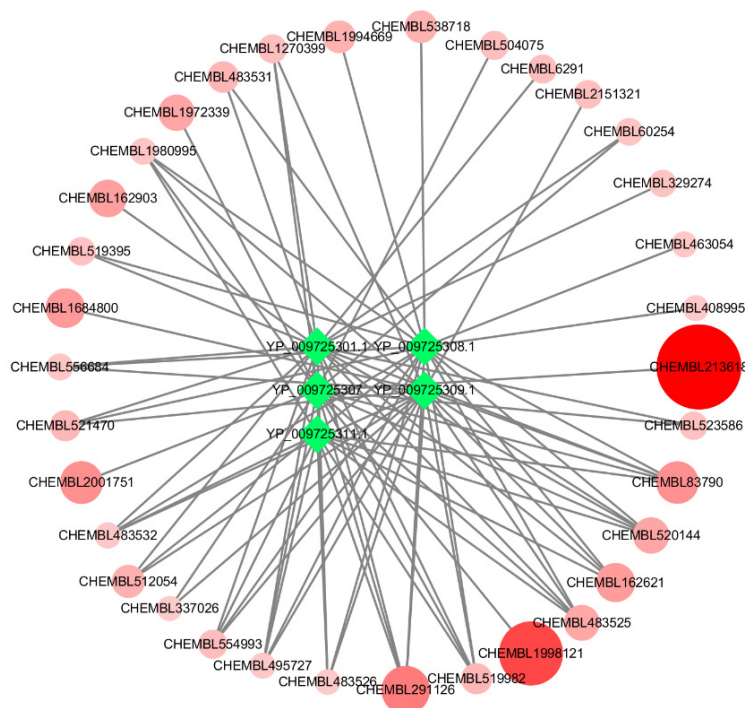


Figure 2. Drug-target interaction network.

3.4.3. Drug repositioning of COVID-19 based on ion channel targets

The prediction results for 5 ion channel targets of SARS-CoV-2 are shown in Tables 2–6 respectively, including ChEMBL ID, small molecule drugs, SMILES code, and drug-target affinity (The name of some small drug molecules have not found).

Table 2. Prediction results of drug-target interaction of 3C-like proteinase (YP_009725301.1).

ChEMBL ID	Small molecules	SMILES	Affinity
CHEMBL162621	Undefined	<chem>N#CC1=C(c2c(c3cc4ccccc4s3)n(CCCSC(=N)N)c3ccccc23)C(=O)NC1=O</chem>	14.227
CHEMBL520144	Dehydroatenu sin	<chem>COc1cc(O)c2c(=O)oc3c(O)c(O)cc(C)c3c2c1</chem>	14.207
CHEMBL291126	SCH-47112	<chem>O=C1NC(=O)c2c1c1c3ccccc3n3c1c1c2c2ccccc2n1C1OC3CC1CO</chem>	14.164
CHEMBL483525	Altenariol 5-O-Sulfate	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(OS(=O)(=O)O)cc3c12</chem>	14.086
CHEMBL519982	Altenariol	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(O)cc3c12</chem>	13.957
CHEMBL483526	Butin	<chem>COc1cc(O)c2c(=O)oc3cc(O)cc(C)c3c2c1</chem>	13.891
CHEMBL512054	Altersolanol A	<chem>COc1cc(O)c2c(c1)C(=O)C1=C(C2=O)C(O)C(O)C(C)(O)C1O</chem>	13.880
CHEMBL554993	Alterporriol G/H	<chem>COc1cc2c(c(O)c1c1c(OC)cc(O)c3c1C(=O)C1=C(C3=O)C(O)C(O)C(C)(O)C1O)C(=O)c1cc(O)c(C)cc1C2=O</chem>	13.815
CHEMBL83790	LPA1 antagonist 1	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cc(O)ccc6n2e5c31)C(=O)NC4</chem>	13.801
CHEMBL162903	Undefined	<chem>N#CC1=C(c2c(c3cc4ccccc4s3)n(CCCN)c3ccccc23)C(=O)NC1=O</chem>	13.717
CHEMBL483532	Desmethylalte nusin	<chem>Cc1cc(O)c(O)cc1-c1cc(O)cc(O)c1C(=O)O</chem>	13.702
CHEMBL556684	6-O- methylalaterni n	<chem>COc1cc(O)c2c(c1)C(=O)c1cc(C)c(O)c(O)c1C2=O</chem>	13.678
CHEMBL521470	Altertoxi I	<chem>O=C1CC(O)C2c3c(ccc(O)c31)c1ccc(O)c3c1C2(O)CCC3=O</chem>	13.648
CHEMBL483531	Altenusin	<chem>COc1cc(O)c(C(=O)O)c(c2cc(O)c(O)cc2C)c1</chem>	13.613
CHEMBL6291	Bisindolylmal eimide IX	<chem>Cn1cc(C2=C(c3cn(CCCSC(=N)N)c4ccccc3)C(=O)NC2=O)c2ccccc21</chem>	13.570
CHEMBL127039 9	Undefined	<chem>Cc1cn2c(c3cn[nH]c3)cnc2c(Nc2ccc(C(C)N3CC(C)OC(C)C3)ns2)n1</chem>	13.567
CHEMBL60254	Balanol	<chem>O=C(NC1CNCCCC1OC(=O)c1cc(O)c(C(=O)c2c(O)cccc2C(=O)O)c(O)c1)c1ccc(O)cc1</chem>	13.563
CHEMBL519395	Alterna rienonic Acid	<chem>COc1cc(O)c(C(=O)O)c(C2=C(C)CC(O)C2=O)c1</chem>	13.561
CHEMBL523586	Undefined	<chem>NCC(Cc1cccc(F)c1)NC(=O)c1cc(Br)c(c2ccnc3[nH]ccc23)s1</chem>	13.559
CHEMBL329274	QD-325	<chem>O=C1NC(=O)C(n2ccc3ccccc32)=C1n1ccc2ccccc21</chem>	13.543
CHEMBL495727	AT-9283	<chem>O=C(NC1=CNCC1=C1N=c2ccc(CN3CCOCC3)cc2=N1)NC1CC1</chem>	13.534

Table 3. Prediction results of drug-target interaction of 2'-O-ribose methyltransferase (YP_009725311.1).

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL162621	Undefined	<chem>N#CC1=C(c2c(-c3cc4ccccc4s3)n(CCCSC(=N)N)c3ccccc23)C(=O)NC1=O</chem>	13.960
CHEMBL520144	Dehydroalte nusin	<chem>COc1cc(O)c2c(=O)oc3c(O)c(O)cc(C)c3c2c1</chem>	13.927
CHEMBL291126	SCH-47112	<chem>O=C1NC(=O)c2c1c1c3ccccc3n3c1c1c2c2ccccc2n1C1OC3CC1CO</chem>	13.903
CHEMBL519982	Alternariol	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(O)cc3c12</chem>	13.856
CHEMBL83790	LPA1 antagonist 1	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6ccc(O)ccc6n2c5c31)C(=O)NC4</chem>	13.831
CHEMBL483525	Alternariol 5-O-Sulfate	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(OS(=O)(=O)O)cc3c12</chem>	13.825
CHEMBL483526	Butin	<chem>COc1cc(O)c2c(=O)oc3cc(O)cc(C)c3c2c1</chem>	13.729
CHEMBL554993	Alterporriol G/H	<chem>COc1cc2c(c(O)c1-c1c(OC)cc(O)c3c1C(=O)C1=C(C3=O)C(O)C(O)C(C)(O)C1O)C(=O)c1cc(O)c(C)cc1C2=O</chem>	13.573
CHEMBL1980995	Staurosporin e	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6ccccc6n2c5c31)C(=O)NC4</chem>	13.532
CHEMBL483532	Desmethylal tenusin	<chem>Cc1cc(O)c(O)cc1-c1cc(O)cc(O)c1C(=O)O</chem>	13.522
CHEMBL495727	AT-9283	<chem>O=C(NC1=CNCC1=C1N=c2ccc(CN3CCOCC3)cc2=N1)NC1CC1</chem>	13.519

Table 4. Prediction results of drug-target interaction of RNA-dependent RNA polymerase (YP_009725307).

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL213618	N-[(1S)-2-AMINO-1-(2,4-DICHLOROBENZYL)ETHYL]-5-[2-(METHYLAMINO)PYRIMIDIN-4-YL]THIOPHENE-2-CARBOXAMIDE	<chem>CNc1nccc(c2ccc(C(=O)NC(CN)Cc3ccc(Cl)cc3Cl)s2)n1</chem>	14.524
CHEMBL523586	N-[(1S)-2-amino-1-phenylethyl]-5-(1H-pyrrolo[2,3-b]pyridin-4-yl)thiophene-2-carboxamide	<chem>NCC(Cc1cccc(F)c1)NC(=O)c1cc(Br)c(c2ccnc3[nH]ccc23)s1</chem>	14.501

Continued on next page

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL199812 1	N-(5-{{(2S)-4-amino-2-(3-chlorophenyl)butanoyl}amino}-1H-indazol-3-yl)benzamide	<chem>NCCC(C(=O)Nc1ccc2[nH]nce2c1)c1ccc(Cl)c(Cl)c1</chem>	14.167
CHEMBL291126	SCH-47112	<chem>O=C1NC(=O)c2c1c1c3ccccc3n3c1c1c2ccccc2n1C1OC3CC1CO</chem>	13.974
CHEMBL83790	LPA1 antagonist 1	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cc(O)ccc6n2c5c31)C(=O)NC4</chem>	13.822
CHEMBL200175 1	EGFR-IN-9	<chem>NC(COc1cnccc(-c2ccc3c(c2)C(=Cc2ccco2)C(=O)N3)c1)Cc1c[nH]c2ccccc12</chem>	13.806
CHEMBL495727	AT-9283	<chem>O=C(NC1=CNNC1=C1N=c2ccc(CN3CCOCC3)cc2=N1)NC1CC1</chem>	13.796
CHEMBL198099 5	Staurosporine	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6ccccc6n2c5c31)C(=O)NC4</chem>	13.702
CHEMBL520144	Dehydratenuosin	<chem>COc1cc(O)c2c(=O)oc3c(O)c(O)cc(C)c3c2c1</chem>	13.695
CHEMBL197233 9	Undefined	<chem>Cc1[nH]nc2ccc(-c3nnn(Cc4ccccc4)c3-c3ccc(F)cc3)cc12</chem>	13.694
CHEMBL483525	Altenariol 5-O-Sulfate	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(OS(=O)(=O)O)cc3c12</chem>	13.686
CHEMBL127039 9	Undefined	<chem>Cc1cn2c(c3cn[nH]c3)enc2c(Nc2cc(C(C)N3CC(C)OC(C)C3)ns2)n1</chem>	13.653
CHEMBL519982	Altenariol	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(O)cc3c12</chem>	13.607
CHEMBL504075	PKI-166 hydrochloride	<chem>Cc1cccc(C(=O)Nc2cccc(C(C)Nc3encc(Cl)n3)c2)c1</chem>	13.570
CHEMBL60254	Balanol	<chem>O=C(NC1CNCCCC1OC(=O)c1cc(O)c(C(=O)c2c(O)cccc2C(=O)O)c(O)c1)c1ccc(O)cc1</chem>	13.546
CHEMBL483526	Butin	<chem>COc1cc(O)c2c(=O)oc3cc(O)cc(C)c3c2c1</chem>	13.512
CHEMBL337026	Staurosporine	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6ccccc6n2c5c31)C(=O)NC4=O</chem>	13.504

Table 5. Prediction results of drug-target interaction of helicase (YP_009725308.1).

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL520144	Dehydratenuosin	<chem>COc1cc(O)c2c(=O)oc3c(O)c(O)cc(C)c3c2c1</chem>	14.244
CHEMBL483525	Altenariol 5-O-Sulfate	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(OS(=O)(=O)O)cc3c12</chem>	14.185
CHEMBL519982	Altenariol	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(O)cc3c12</chem>	14.166

Continued on next page

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL291126	SCH-47112	<chem>O=C1NC(=O)c2c1c1c3ccccc3n3c1c1c2c2ccccc2n1C1O</chem> <chem>C3CC1CO</chem>	14.079
CHEMBL162621	Undefined	<chem>N#CC1=C(c2c(c3cc4ccccc4s3)n(CCCSC(=N)N)c3ccccc</chem> <chem>23)C(=O)NC1=O</chem>	14.047
CHEMBL483526	Butin	<chem>COc1cc(O)c2c(=O)oc3cc(O)cc(C)c3c2c1</chem>	14.039
CHEMBL83790	LPA1 antagonist 1	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cc(O)cc</chem> <chem>c6n2c5c31)C(=O)NC4</chem>	13.980
CHEMBL554993	Alterporriol G/H	<chem>COc1cc2c(c(O)c1c1c(OC)cc(O)c3c1C(=O)C1=C(C3=O)</chem> <chem>C(O)C(O)C(C)(O)C1O)C(=O)c1cc(O)c(C)cc1C2=O</chem>	13.924
CHEMBL512054	Altersolanol A	<chem>COc1cc(O)c2c(c1)C(=O)C1=C(C2=O)C(O)C(O)C(C)(O</chem> <chem>)C1O</chem>	13.861
CHEMBL483532	Desmethylaltenusin	<chem>Cc1cc(O)c(O)cc1-c1cc(O)cc(O)c1C(=O)O</chem>	13.826
CHEMBL521470	Altertoxin I	<chem>O=C1CC(O)C2c3c(ccc(O)c31)c1ccc(O)c3c1C2(O)CCC</chem> <chem>3=O</chem>	13.796
CHEMBL556684	6-O-methylalaternin	<chem>COc1cc(O)c2c(c1)C(=O)c1cc(C)c(O)c(O)c1C2=O</chem>	13.754
CHEMBL519395	Alternarienonic Acid	<chem>COc1cc(O)c(C(=O)O)c(C2=C(C)CC(O)C2=O)c1</chem>	13.729
CHEMBL198099	Staurosporine 5	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6ccccc6n</chem> <chem>2c5c31)C(=O)NC4</chem>	13.708
CHEMBL483531	Altenusin	<chem>COc1cc(O)c(C(=O)O)c(c2cc(O)c(O)cc2C)c1</chem>	13.674
CHEMBL199466	Undefined 9	<chem>Cn1cc(Nc2ncc(Br)c(Nc3ccccc3C(N)=O)n2)cn1</chem>	13.634
CHEMBL538718	Ro-32-0557	<chem>CN1CC2Cc3c(C4=C(c5cn(C)c6ccccc56)C(=O)NC4=O)c</chem> <chem>4ccccc4n3CC2C1</chem>	13.632
CHEMBL495727	AT-9283	<chem>O=C(NC1=CNCC1=C1N=c2ccc(CN3CCOCC3)cc2=N1</chem> <chem>)NC1CC1</chem>	13.616
CHEMBL463054	Macrosporin	<chem>COc1cc(O)c2c(c1)C(=O)c1cc(C)c(O)cc1C2=O</chem>	13.519
CHEMBL408995	Talaroflvone	<chem>COc1cc(O)c2c(c1)C1(OC2=O)C(C)=CC(=O)C1O</chem>	13.518

Table 6. Prediction results of drug-target interaction of the 3'-to-5' exonuclease (YP_009725309.1).

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL83790	LPA1 antagonist 1	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cc(O)cc</chem> <chem>c6n2c5c31)C(=O)NC4</chem>	14.441
CHEMBL483525	Altenariol 5-O-Sulfate	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(OS(=O)(=O)O)cc3c12</chem>	14.138
CHEMBL291126	SCH-47112	<chem>O=C1NC(=O)c2c1c1c3ccccc3n3c1c1c2c2ccccc2n1C1O</chem> <chem>C3CC1CO</chem>	14.129
CHEMBL520144	Dehydroal tenusin	<chem>COc1cc(O)c2c(=O)oc3c(O)c(O)cc(C)c3c2c1</chem>	14.009
CHEMBL495727	AT-9283	<chem>O=C(NC1=CNCC1=C1N=c2ccc(CN3CCOCC3)cc2=N1</chem> <chem>)NC1CC1</chem>	13.999

Continued on next page

CHEMBL ID	Small molecules	SMILES	Affinity
CHEMBL519982	Altenariol	<chem>Cc1cc(O)cc2oc(=O)c3c(O)cc(O)cc3c12</chem>	13.971
CHEMBL337026	Undefined	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cccc6n2c5c31)C(=O)NC4=O</chem>	13.883
CHEMBL483526	Butin	<chem>COc1cc(O)c2c(=O)oc3cc(O)cc(C)c3c2c1</chem>	13.812
CHEMBL1684800	Undefined	<chem>COc1ccc(Cn2ncc(NC(=O)c3cc(NC(=O)Nc4ccc(Cl)c(C(F)(F)F)c4)ccc3C)c2N)cc1</chem>	13.752
CHEMBL162621	Undefined	<chem>N#CC1=C(c2c(-c3cc4cccc4s3)n(CCCSC(=N)N)c3ccccc23)C(=O)NC1=O</chem>	13.743
CHEMBL554993	Alterporriol G/H	<chem>COc1cc2c(c(O)c1-c1c(OC)cc(O)c3c1C(=O)C1=C(C3=O)C(O)C(O)C(C)(O)C1O)C(=O)c1cc(O)c(C)cc1C2=O</chem>	13.716
CHEMBL512054	Altersolanol A	<chem>COc1cc(O)c2c(c1)C(=O)C1=C(C2=O)C(O)C(O)C(C)(O)C1O</chem>	13.640
CHEMBL521470	Altortoxin I	<chem>O=C1CC(O)C2c3c(ccc(O)c31)c1ccc(O)c3c1C2(O)CCC3=O</chem>	13.604
CHEMBL1270399	Undefined	<chem>Cc1cn2c(-c3cn[nH]c3)enc2c(Nc2cc(C(C)N3CC(C)OC(C)C3)ns2)n1</chem>	13.604
CHEMBL1980995	Staurosporine	<chem>CNC1CC2OC(C)(C1OC)n1c3ccccc3c3c4c(c5c6cccc6n2c5c31)C(=O)NC4</chem>	13.585
CHEMBL2151321	Undefined	<chem>CC(C)Oc1cc(-n2cnc3ccc(NC(C)c4ccc(F)cn4)nc32)[nH]n1</chem>	13.564
CHEMBL483532	Desmethy laltenusin	<chem>Cc1cc(O)c(O)cc1-c1cc(O)cc(O)c1C(=O)O</chem>	13.538
CHEMBL556684	6-O- methylaltenin	<chem>COc1cc(O)c2c(c1)C(=O)c1cc(C)c(O)c(O)c1C2=O</chem>	13.534

Table 7. Affinity and average values of 7 drugs targeting 5 target proteins of SARS-CoV-2.

CHEMBL ID	Small molecules	YP_0097253 01.1	YP_009725 307	YP_0097253 08.1	YP_0097253 09.1	YP_0097253 11.1	Average
CHEMBL291126	SCH- 47112	14.164	13.974	14.079	14.129	13.903	14.050
CHEMBL520144	Dehdroal- tenuisin	14.207	13.695	14.244	14.009	13.927	14.016
CHEMBL483525	Alternariol 5-O- Sulfate	14.086	13.686	14.185	14.138	13.825	13.984
CHEMBL83790	LPA1 antagonist 1	13.801	13.822	13.980	14.441	13.831	13.975
CHEMBL519982	Alternariol	13.957	13.607	14.166	13.971	13.856	13.911
CHEMBL483526	Butin	13.891	13.512	14.039	13.812	13.729	13.797
CHEMBL495727	AT-9283	13.534	13.796	13.616	13.999	13.519	13.693

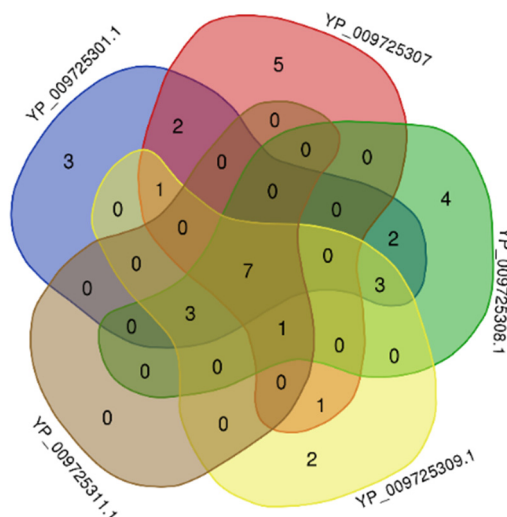


Figure 3. Drug-target Venn diagram.

SCH-47112 is a derivative of Staurosporine. Reynolds N J et al. verified that SCH-47112 had an inhibitory effect on skin inflammation and epidermal hypertrophy induced by tPA (12-O-tetradecanoylphorbol-13-acetate) in hairless mice [19]. According to the prediction results of this model, SCH-47112 should have the best effect on COVID-19. Staurosporine, a potent ATP-competitive kinase inhibitor, can effectively inhibit PKC activity. The biological activity of Staurosporine in antifungal and antihypertensive treatment makes it have great potential in anti-cancer treatment. The main functions of Staurosporine are inducing G2/M phase blocking of cancer cells, adjusting G1 phase blocking and cell apoptosis [20]. It may be possible to study whether Staurosporine is effective in the treatment of COVID-19. Dehydroaltenusin, less effective than SCH 47112 as shown in the results, is a small molecule selective inhibitor of DNA polymerase α , an antibiotic produced by fungus. It blocks the cancer cell cycle and triggers apoptosis in the S phase and has antitumor activity against human adenocarcinoma tumors in vivo [21]. LPA1 antagonist 1 is a highly selective, lysophosphatidic acid receptor antagonist, LPA interacts with GPCRs to modulate signal response. Alternariol is a mycotoxin produced by *Alternaria* that inhibits the catalytic activity of topoisomerase I and topoisomerase II. Alternariol has biological activities such as anti-HIV, anti-cancer and anti-microbial properties [22]. Butin and AT-9283 were found to be less effective in treating COVID-19. Butin is a kind of bioactive flavonoids isolated from the heartwood of sandalwood, which has strong antioxidant, anti-platelet and anti-inflammatory activities. It can significantly reduce myocardial infarction, improve cardiac function and prevent oxidative damage of the heart caused by diabetes [23]. AT-9283 is A multi-target kinase inhibitor that effectively inhibits Aurora A/B, Jak2/3 and Flt3. It can inhibit the growth and survival of multiple solid tumors in vitro and in vivo [24]. Based on the predicted results of the proposed model, these drugs act on 5 target proteins of SARS-CoV-2 and can be retargeted to investigate their potential to treat COVID-19.

4. Conclusions

Aiming at the urgent needs of COVID-19 treatment, a mixed graph network model based on graph attention network and graph convolutional network was proposed for prediction of drug-target affinity.

Compared with DeepDTA, WideDTA, GAT, GCN and GIN models, it was proved that MSE of the proposed model was decreased by 0.055, 0.04, 0.001, 0.046, 0.013, and CI was increased by 0.028, 0.016, 0.003, 0.03 and 0.01, respectively. According to the prediction results of drug-target affinity of SARS-CoV-2 ion channel targets, 7 kinds of small molecule drugs acting on 5 ion channel targets were obtained, namely SCH-47112, dehydroaltenusin, alternariol 5-o-sulfate, LPA1 antagonist 1, alternariol, butin, and AT-9283. However, the research results in this paper only provide a certain reference basis for drug repositioning and precise treatment of COVID-19, and the actual therapeutic effect still needs to be verified by a large number of clinical trials. In the future, we will further modify our model, such as extracting protein sequence features using location specific score matrix (PSSM) and other deep learning methods [25–28], to achieve higher model prediction performance.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62072157,61802116), the Natural Science Foundation of Henan province (202300410102), the Doctoral program of Henan Institute of Technology (KQ2002), and the Science and Technology Research Key Project of Henan Province (No. 192102210113).

Conflict of interest

We declare that there is no conflict of interest.

References

1. K. Huang, T. Fu, L. M. Glass, M. Zitnik, C. Xiao, J. Sun, DeepPurpose: a deep learning library for drug–target interaction prediction, *Bioinformatics*, **36** (2020), 5545–5547. <https://doi.org/10.1093/bioinformatics/btaa1005>
2. T. U. Singh, S. Parida, M. C. Lingaraju, M. Kesavan, D. Kumar, R. K. Singh, Drug repurposing approach to fight COVID-19, *Pharmacol. Rep.*, **72** (2020), 1479–1508. <https://doi.org/10.1007/s43440-020-00155-6>
3. 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. <https://doi.org/10.1126/science.abb2762>
4. A. Elfiky, N. S. Ibrahim, Anti-SARS and anti-HCV drugs repurposing against the Papain-like protease of the newly emerged coronavirus (2019-nCoV), preprint. <https://doi.org/10.21203/rs.2.23280/v1>
5. R. Arya, A. Das, V. Prashar, M. Kumar, Potential inhibitors against papain-like protease of novel coronavirus (COVID-19) from FDA approved drugs, preprint, chemRxiv. <https://doi.org/10.26434/chemrxiv.11860011>
6. D. Zhang, K. Wu, X. Zhang, S. Deng, B. Peng, In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus, *J. Integr. Med.*, **18** (2020), 152–158. <https://doi.org/10.1016/j.joim.2020.02.005>

7. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, Trends in GPCR drug discovery: new agents, targets and indications, *Nat. Rev. Drug Discovery*, **16** (2017), 829–842. <https://doi.org/10.1038/nrd.2017.178>
8. N. Wang, Y. Li, S. Liu, L. Gao, C. Liu, X. Bao, et al., Bioinformatics analysis and validation of differentially expressed microRNAs with their target genes involved in GLP-1RA facilitated osteogenesis, *Curr. Bioinf.*, **16** (2021), 928–942. <https://dx.doi.org/10.2174/1574893615999200508091615>
9. D. Weininger, SMILES: a chemical language and information system, *J. Chem. Inf. Comput. Sci.*, **28** (1988), 31–36. <https://doi.org/10.1021/ci00057a005>
10. B. R. Beck, B. Shin, Y. Choi, S. Park, K. Kang, 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.*, **18** (2020), 784–790. <https://doi.org/10.1016/j.csbj.2020.03.025>
11. T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in *Proceedings of the International Conference on Learning Representations (ICLR)*, (2017). <https://arxiv.org/abs/1609.02907>
12. H. Jiang, P. Cao, M. Xu, J. Yang, Z. Osmar, Hi-GCN: a hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction, *Comput. Biol. Med.*, **127** (2020), 104096. <https://doi.org/10.1016/j.compbiomed.2020.104096>
13. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph attention networks, in *Proceedings of the International Conference on Learning Representations (ICLR)*, (2018). <https://doi.org/10.17863/CAM.48429>
14. C. Yang, P. Wang, J. Tan, Q. Liu, X. Li, Autism spectrum disorder diagnosis using graph attention network based on spatial-constrained sparse functional brain networks, *Comput. Biol. Med.*, **139** (2021), 104963–104963. <https://doi.org/10.1016/j.compbiomed.2021.104963>
15. K. Xu, W. Hu, J. Leskovec, S. Jegelka, How powerful are graph neural networks?, in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2019. <https://arxiv.org/abs/1810.00826>
16. H. Öztürk, A. Özgür, E. Ozkirimli, DeepDTA: deep drug-target binding affinity prediction, *Bioinformatics*, **34** (2018), 821–829. <https://doi.org/10.1093/bioinformatics/bty593>
17. H. Öztürk, E. Ozkirimli, A. Özgür, WideDTA: prediction of drug-target binding affinity, preprint, arXiv:1902.04166. <https://arxiv.org/abs/1902.04166v1>
18. X. Guo, W. Zhou, B. Shi, et al., An efficient multiple kernel support vector regression model for assessing dry weight of hemodialysis patients, *Curr. Bioinf.*, **16** (2021), 284–293. <https://dx.doi.org/10.2174/1574893615999200614172536>
19. N. J. Reynolds, S. W. McCombie, B. B. Shankar, W. R. Bishop, G. J. Fisher, SCH 47112, a novel staurosporine derivative, inhibits 12-O-tetradecanoylphorbol-13-acetate-induced inflammation and epidermal hyperplasia in hairless mouse skin, *Arch. Dermatol. Res.*, **289** (1997), 540–546. <https://doi.org/10.1007/s004030050236>
20. I. Kruman, Q. Guo, M. P. Mattson, Calcium and reactive oxygen species mediate staurosporine-induced mitochondrial dysfunction and apoptosis in PC12 cells, *J. Neurosci. Res.*, **51** (1998), 293–308. [https://doi.org/10.1002/\(SICI\)1097-4547\(19980201\)51:3<293::AID-JNR3>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1097-4547(19980201)51:3<293::AID-JNR3>3.0.CO;2-B)

21. Y. Mizushina, N. Maeda, I. Kuriyama, H. Yoshida, Dehydroaltenusin is a specific inhibitor of mammalian DNA polymerase α , *Expert Opin. Invest. Drugs*, **20** (2011), 1523. <https://doi.org/10.1517/13543784.2011.619977>
22. E Fliszár-Nyúl, B. Lemli, S. Kunsági-Máté, L. Dellafiora, C. Dall'Asta, G. Cruciani, et al., Interaction of mycotoxin alternariol with serum albumin, *Int. J. Mol. Sci.*, **20** (2019), 2352. <https://doi.org/10.3390/ijms20092352>
23. J. Duan, Y. Guan, F. Mu, C. Guo, E. Zhang, Y. Yin, et al., Protective effect of butin against ischemia/reperfusion-induced myocardial injury in diabetic mice: involvement of the AMPK/GSK-3 β /Nrf2 signaling pathway, *Sci. Rep.*, **7** (2017), 41491. <https://doi.org/10.1038/srep41491>
24. S. Howard, V. Berdini, J. A. Boulstridge, M. G. Carr, D. M. Cross, J. Curry, et al., Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity, *J. Med. Chem.*, **52** (2009), 379–388. <https://doi.org/10.1021/jm800984v>
25. H. Li, Y. Gong, Y. Liu, H. Lin, G. Wang, Detection of transcription factors binding to methylated DNA by deep recurrent neural network, *Briefings Bioinf.*, **2021** (2021), <https://doi.org/10.1093/bib/bbab533>.
26. L. Zhang, Y. Yang, L. Chai, Q. Li, J. Liu, H. Lin, et al., A deep learning model to identify gene expression level using co-binding transcription factor signals, *Briefings Bioinf.*, **2021** (2021), <https://doi.org/10.1093/bib/bbab501>
27. H. Lv, L. Shi, J. W. Berkenpas, F. Y. Dao, H. Zulfiqar, H. Ding, et al., Application of artificial intelligence and machine learning for COVID-19 drug discovery and vaccine design, *Briefings Bioinf.*, **22** (2021), 1–10. <https://doi.org/10.1093/bib/bbab320>
28. Z. Chen, Q. Zhou, A. Khan, J. Jill, R. Xiong, X. Liu, New trends of deep learning in clinical cardiology, *Curr. Bioinf.*, **16** (2021), 954–962. <https://dx.doi.org/10.2174/1574893615999200719234517>



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)