AttentionSiteDTI: Attention Based Model for Predicting Drug-Target Interaction Using 3D Structure of Protein Binding Sites

Mehdi Yazdani-Jahromi *[1,4], Niloofar Yousefi *[1,5], Aida Tayebi[1,6], Ozlem Ozmen Garibay[1,7], Sudipta Seal[2, 3, 8], Elayaraja Kolanthai[2,9], and Craig J. Neal[2,10]

[1]Industrial Engineering and Management Systems, University of Central Florida, Orlando, FL
[2]Advanced Materials Processing and Analysis Center, Dept. of Materials Science and Engineering, Nanoscience Technology Center, University of Central Florida,Orlando, FL
[3]College of Medicine, Bionix Cluster, University of Central Florida,Orlando, FL
Email ids: [4]Yazdani@knights.ucf.edu, [5]Niloofar.Yousefi@ucf.edu, [6]Aida.Tayebi@knights.ucf.edu, [7]Ozlem@ucf.edu, [8]Sudipta.Seal@ucf.edu, [9]Elayaraja.Kolanthai@ucf.edu, [10]Craig.Neal@ucf.edu

December 8, 2021

Abstract

Investigating drug-target interactions plays a critical role in drug design and discovery. The vast chemical and proteomic space, along with the cost associated with in-virto experiments motivate the use of computational methods to narrow down the search space for novel interaction of drug–target pairs. Among all computational methods, deep learning algorithms have gained increased attention due to their power in automatically learning and extracting feature representations, and therefore identifying, processing and extrapolating complex hidden interactions between drugs and targets. In this study, we introduce and implement a new graph-based prediction model called AttentionSiteDTI. Our proposed model utilize the binding sites (pockets) of the proteins as the input for the target protein, and it uses a self-attention mechanism to make the model learn which binding sites of the protein interact with a given ligand. This, indeed, complements the black-box nature of deep learning-based methods and enables interpretability, while achieving state-of-the-art results in drug-target interaction prediction task on three datasets. The AttentionSite DTI achieves AUC of 0.97 (for seen proteins), 0.94 (for unseen proteins) in the customized BindingDB dataset, 0.971 in the DUD-E dataset, and 0.991 in the human dataset. In general, the prediction results on these datasets show the superiority of our AttentionSiteDTI compared to previous graph-based models, and our ablation studies proves the effectiveness of our proposed model in prediction of drug-target interactions. In addition, through multidisciplinary collaboration in this work, we further experimentally evaluate the

*These authors contributed equally.

practical potential of our proposed approach. To achieve this, we first computationally predict binding interaction of some candidate compounds with a target protein, and then experimentally validate the binding interactions for these pairs in the laboratory. The high agreement between the computationally-predicted and experimentally-observed (measured) drug-target interactions illustrates the potential of our AttentionSiteDTI as effective pre-screening tool in drug repurposing applications.

1 Introduction

Drug-target interaction characterizes the binding between a drug and its target, which helps the problem of identifying candidate drug compounds that interact with particular target proteins. Therefore, the accurate identification of such interactions is critical to discovery of novel drug species, and/or repurposing of existing drugs (through identifying novel interacting pairs). High-Throughput Screening (HTS) remains the most reliable approach to examine the affinity of a drug toward its targets. However, the experimental characterization of every possible compound–protein pair quickly becomes impractical, due to the immense space of chemical compounds, targets and mixtures. This motivates the use of computational approaches to predict whether candidate drugs are able to inhibit the target/protein. Drug-Target Interaction (DTI) prediction can help speed up the process of drug development, and reduce the risks of experiments, which can be of critical importance in finding safe and effective treatments to novel challenging diseases in the time of outbreaks.

Molecular simulation and molecular docking are among earlier computational approaches, which typically require 3D structures of the target proteins to assess the drug-target interaction. Although these structure-based methods can be very informative, their application is very limited, as there are many proteins with unknown 3D structures, beside that they involve an expensive process. Artificial Intelligence (AI)-based approaches, including deep learning (DL) and machine learning (ML) algorithms have then emerged to overcome some of these challenges in the process of drug design and discovery. Traditional shallow ML-based models translate knowledge about known drugs and targets into features that are then used as input to train predictive models, predicting interaction between unknown drug–target pairs. Examples of such models are Kronecker Regularized Least Squares (KronRLS) [21] and boosting machines (SimBoost) [9] that were proposed to address the problem of DTI prediction. The performance of these models, however, highly depends on hand-crafted features captured from drug data and protein sequences. Therefore, these models are usually unable to perform well in modeling complex interactions between drug-target pairs. DL has advanced these traditional computational models due to their ability in automatically capturing useful latent features (that are difficult to be hand-crafted using human experience), leading to highly flexible models with extensive power in identifying, processing and extrapolating complex patterns in molecular data.

Deep learning models for DTI can be mainly categorized into two classes. One class is designed to work with sequence-based representations for the input data. Example of this type of models are Convolutional Neural Nettworks (CNNs), which utilize the string representations of drug, usually, in the form of Simplified Molecular-Input Line-entry System (SMILES), and sequence representation of protein, typically in the form of amino acid sequence. CNNs are often inefficient in learning long-term dependencies, or the order relationship in the amino-acid sequence. However, this limitation can be avoided by using more suitable architec-

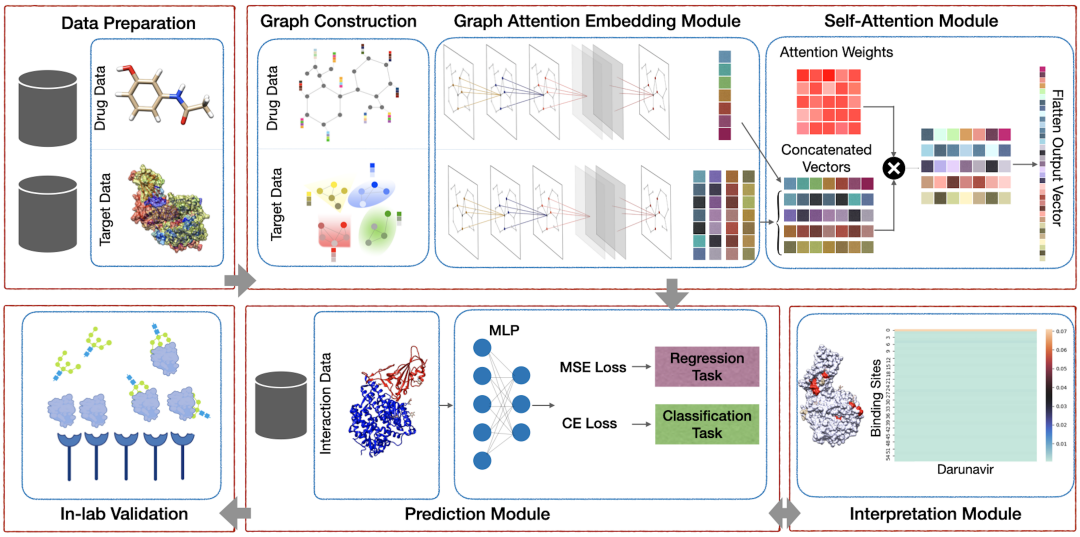


Figure 1: Our proposed framework includes 5 main modules: (1) Preprocessing module that consists of finding the binding sites of proteins; (2) AttentionSiteDTI deep learning module, where we construct graph representations of ligands' SMILE and proteins' binding sites, and we create a graph convolutional neural network armed with an attention pooling mechanism to extract learnable embeddings from graphs, as well as a self-attention mechanism to learn relationship between ligands and proteins' binding sites; (3) Prediction module to predict unknown interaction in a drug-target pair, which can address both classification and regression tasks; (4) Interpretation module to provide a deeper understanding of which binding sites of a target protein are more probable to bind with a given ligand. (5) In-lab validations, where we compare our computationally-predicted results with experimentally-observed (measured) drug-target interactions in laboratory to test and validate the practical potential of our proposed model.

tures capable of learning from long sequences of proteins, such as RNNs or Long-Short Term Memory (LSTMs). Despite the remarkable success of CNNs and RNNs in DTI prediction, they are incapable of capturing structural information of the molecules, which in turn, leads to degraded predictive power of these models. This motivates the use of a more natural representation of the molecules and the convention of second class of DL models, namely Graph Neural Networks (GNNs). The basic strategy of these models is to use graph descriptions of the molecules, where atoms and chemical bonds correspond to nodes and edges, respectively. Graph Convolutional Neural Network (GCNN) and graph Attention Network (GAT) are the two widely used GNN-based models in computer-aided drug design and discovery [27, 16, 25]. Although all these graph-based models use graph representation for drug-like molecules, most of them use amino acid sequence representations for proteins. Given the fact that, in reality, protein sequences fold in 3 dimensional space, the drawback of using 1D protein sequence is that it can not capture the 3D structural features that are key factors in the prediction of drug-target interactions. On the other hand, obtaining the high-resolution 3D structure of the proteins is a challenging task, beside the fact that proteins contain a large number of atoms requiring a large scale 3D (sparse) matrix to capture the whole structure. This not only makes the training computationally expensive, but also leads to models with relatively low accuracy with no practical benefits. Perhaps, it is because of these reasons that there has been only a limited number of works taking the 3D structure of proteins as the direct input to their models. To alleviate this issue, an alternative strategy has been adopted in the literature wherein the proteins are represented by a 2D contact (or distance) map that shows the interaction of proteins' residue pairs in the form of a matrix. DGraphDTA [10] introduced these contact maps and uses them as the output of protein structure prediction. These contact maps are consistent with the adjacency matrix in GNNs, allowing for a combination of two data sources to perform an affinity prediction task. Another study [37] utilized these 2D distance maps along with their corresponding SMILES and fed these inputs into a dynamic attentive convolutional neural network (DynCNN) and a self-attentional sequential model. It is worth mentioning that a contact (or distance) map is typically the output of protein structure prediction, which is based on heuristics and provides only an approximation abstraction of the real structure of protein, generally, different from the one determined experimentally via X-ray crystallography or by nucleic magnetic resonance spectroscopy (NMR) [28]. Taken all together, and considering the fact that the binding of a protein to many molecules occurs at different binding pockets (of the protein's surface) rather than the whole protein, in this paper, we represent protein pockets as graphs where the key protein residues correspond to the nodes that are connected based on residue proximity. Furthermore, the features associated with each node are encoded as a vector describing the local amino acid environment. Also, to further improve prediction performance for DTIs, we propose a computational method that utilizes topological information of protein binding pockets as well as the drug ligands in the form of graphs. Our model is inspired by the ones developed for text classification in the field of Natural Language Processing (NLP), and is highly explainable due to its self attention mechanism.

In terms of methodology, our contribution can be summarized in three parts. First, we use graph representation of protein pockets as the input for target protein. Given the fact that intermolecular interactions between protein and ligand occur in pocket-like regions of the pro-

tein, prediction models that utilize the binding sites (pockets) of the proteins are expected to have better generalizability, compared to those relying on certain patterns present in drug molecules or protein sequences. Second, we devise a self-attention mechanism to make the model learn which parts of the protein interact with the ligand, thus complement the black-box nature of deep learning-based methods and enables interpretability, while achieving better DTI prediction performance [44], [52], [73]. Third, we build an end-to-end Graph Convolutional Neural Network (GCNN)-based model, which (1) automatically learns useful embeddings from the graphs of raw molecules and protein pockets, that is, the embeddings are not fixed, but they change according to the context (i.e., sentence) in which they appear and (2) use the learned embeddings similar to the word embeddings, by treating the drug-target complex as a sentence with relational meaning between its biochemical entities a.k.a. protein pockets and drug molecule. This consideration is motivated by the fact that the structure of drug-target complex can be very similar to the structure of a natural language sentence in that the structural and relational information of the entities are keys in understanding the most important information of the sentence. In this regard, each protein pocket or drug is analogous to a word, and each drug-target pair is analogous to a sentence. More specifically, we hypothesize that self-attention bidirectional Long Short-Term Memory (LSTM) mechanism can be used to capture any relationship between binding sites of a given protein and the drug in a sequence, and thus provide a better understanding of their binding relationships. Finally, we conduct in-lab experimental investigations to test the practical potential of our model in prediction and evaluation of compound-target binding interactions in a real world application. Visualization of the aforementioned method can be found in Figure 1. To the best of our knowledge, we are the first to use attention-based bidirectional LSTM networks to perform a relation classification to capture the most important contextual semantic or relational information in a biochemical sequence (i.e. sentence). Each part of our contribution will be described in Section 3.

2 Related works

Deep learning based approaches have been successfully deployed to address drug target interaction prediction. These approaches show better performance compared to machine learning algorithms and have lower computational cost compared to the docking methods. The main difference between these deep learning approaches are in their architecture and the structure of the neural network and representation of the input data. Using linear representation in these methods for drugs and proteins is very common. Small molecules of the drugs can be easily and effectively represented in one-dimensional space, but proteins are much bigger molecules with complex interaction and 1D representations can be insufficient. Although the datasets for 3D structure of the protein are limited, some recent deep learning based literature have benefited from them which we will introduce them:

AtomNet [33] is the first study that has used the 3D structure of the protein and the structure-based, deep convolutional neural network to predict the binding of drug- target pairs. This study used the voxelized complex of the ligand and protein into a cube then 3D convolutional neural networks (CNNs) were deployed to build a binary classifier.

[23] proposed a CNN scoring function that took the 3D representation of the protein-ligand and learned the features critical in binding prediction. This model outperformed AutoDock Vina score in terms of discriminat-

ing and ranking the binding poses.

Pafnucy [26] has proposed 3D convolutional neural networks in this field that uses regression instead of binary classification and predicted the binding affinity values for the drug-target pairs. This study represented the input with 3D grid and considers both proteins and ligands atoms similar. Using a regularization technique, their designed network focused on capturing the general properties of interactions between proteins and ligands.

There are some limitations associated with all these studies. For example, it is a highly challenging task to experimentally obtain high-quality 3D structure of proteins, which explains why the number of datasets with 3D structure information is very limited [37]. Most studies that use 3D structural information utilize convolutional neural networks, which are sensitive to different orientations of the 3D structure, beside the fact that these approaches are computationally expensive. To use all orientations of the same structure these approaches become even more expensive and time consuming. To overcome these limitations, recent studies have proposed graph convolutional network approaches. Previous studies such as [8, 11, 20] have used Graph Convolutional Neural Networks (GCNNs) in the field of ligand protein interaction. There are other studies that have utilized the GCNN architectures to the 3D structure of the protein-ligand complex. Among these studies, GraphBAR [25], is the first 3D graph convolutional neural network that used a regression approach to predicts drug-target binding affinities. They used graphs to represent the complex of protein-ligand instead of 3D voxelized grid cube. These graphs were in the form of multiple adjacency matrices in which the entries were calculated based on distance and feature matrices of molecular properties of the atoms. Also, to overcome the limitation on 3D structure data, they used a docking simulation method to augment additional data to their model.

Lim et. al. [16] proposed a graph convolutional network model along with a distance-aware graph attention mechanism to extract features of the interactions binding pose, directly from 3D structure of drug-target complexes from docking softwares. Their model improved over docking and several deep learning-based models, in terms of virtual screening and pose prediction task. However, their approach had limitations such as less explainability as well as addition docking errors added to the deep learning model.

Pocket Feature is an unsupervised autoencoder model, which was proposed by Torng et. al. [27], to learn representations from binding sites of the target proteins. The model uses 3D graph representations for protein pockets along with 2D graph representations for drugs. They train a GCNN model to extract features from the graphs of protein pockets and drugs' SMILEs. This model performed better than 3DCNN[23] and docking simlutation models such as AutoDock Vina[29], RF-Score[18], and NNScore[18].

Zheng et al [37] pointed out the low efficiency of using direct input of three-dimensional structure and utilized a 2D distance map to represent the proteins. They further converted the problem of drug-target interaction prediction into a classical visual question and answering (VQA) problem, wherein, given a distance map of a protein, the question was whether or not a given drug interacts with the target protein. Although their model outperformed several state-of-the-art models, their VQA system is able to solve a classification task only, where it predicts if there is an interaction between drug-target pairs.

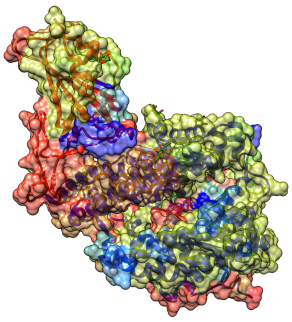


Figure 2: Depiction of COVID spike protein and ACE2 complex with PDB-ID of 6M0J; Color of the surface represents the binding sites computed through Saberi Fathi et. al algorithm which yields the binding site of the proteins. All protein visualization was produced with UCSF Chimera software [22]

3 Materials and Methods

3.1 Preprocessing

3.1.1 Proteins

We use 3D structure of the proteins in this study. These 3D structures are extracted from Protein Data Bank (PDB) files of proteins. PDB data are collections of submitted experimental values (e.g. from NMR, x-ray diffraction, cryo-electron microscopy) for proteins. We use the algorithm proposed by Saberi Fathi et. al [24] to find binding pockets of proteins. Figure 2 provides a visualization of a protein's binding sites. This algorithm computes bounding box coordination for each binding site of a protein. These coordinations were then used for reducing complete protein structures to a subset of peptide fragments. These fragments can be represented as a graph wherein each atom is a node and the connection between atoms are edges in the graph. For each atom, a vector was constructed to represent the atom's features. Also, one-hot encoding of atom type, atom degree, total number of hydrogen atoms and implicit valence of the atom were used to compute feature vector of each atom. This approach yields vector with a size of 1×31 for each node.

3.1.2 Ligands

A bidirectional graph is constructed for each ligand, which is represented in Simplified molecular-input line-entry system(SMILE) format in drug-target interaction data sets. Each atom in the ligand molecule is represented as a node and the connections between atoms are represented with edges in graph. In this study, hydrogen atoms are not explicitly represented as nodes in the graph. Also, a vector was constructed to represent atom's features in the graph. Similarly, one-hot encoding of atom type, atom degree, formal charge of the atom, number of radical electrons of the atom, the atom's hybridization, atom's aromaticity, and the number of total hydrogens of the atom were used to construct the features of the atoms in a ligand. This approach yields vector with a size of 1×74 for each node.

3.2 Embedding

Generated graphs for proteins and ligands are then fed into a graph convolutional neural network to learn embeddings.

3.2.1 Topology Adaptive Graph Convolutional Networks

Topology Adaptive Graph CNN (TAGCN) is a variant of graph convolutional neural network that bases its convolutional layers on graph signal processing techniques in order to learn nonlinear representations of graph-structured data [3]. It works by simultaneously sliding a set of fixed-size learnable filters on a given graph.

This produces a weighted sum of the filter's outputs, representing both strength correlation between graph vertices and the vertex features themselves [3]. The graph convolutional layer for TAGCN is defined as

$$H^k = \sum_{k=0}^{K} \left((D^{-1/2} A D^{-1/2})^k X \Theta_k + b_k \right), \quad (1)$$

where A denotes the adjacency matrix, $D_{ii} = \sum_{j=0} A_{ij}$ is its corresponding diagonal degree matrix, Θ_k is the linear weights that accumulates the results of different hops together, with K being the number of hops, indicating the length of a path from a given node.

3.2.2 Pooling Mechanism

Once the constructed graphs for proteins and drugs are fed into a series of different graph convolutional layers, we then utilize the method proposed by Li et. al. [14] to extract embedding from the corresponding graphs. For the graph level representation, they define a vector as

$$h_{\mathcal{G}} = \tanh\left(\sum_{\nu \in \mathcal{V}} \sigma\big(i(h_\nu^{(T)}, x_\nu)\big) \odot \tanh\big(j(h_\nu^{(T)}, x_\nu)\big) \right), \quad (2)$$

where, i and j are neural networks, $h_\nu^{(T)}$ and x_ν are input and outputs real-valued vectors, and $\sigma(i(h_\nu^{(T)}, x_\nu))$ is a soft attention mechanism [14].

3.3 Sequence Handling

Following the extraction of embeddings, we then treat the problem as a text classification problem, which can be defined as follows:
Let $d \in \mathbb{X}$ denote a protein-ligand complex, where $\mathbb{X}$ is space of embeddings for protein pockets and ligands. Also, define the fixed set of classification labels as $\mathbb{C} = \{0, 1\}$, with 0 for non-active and 1 for active interactions for a given drug-target pair. Let $\mathbb{D}$ denote the labeled training set of protein-ligand complexes $\langle d, c \rangle$, where $\langle d, c \rangle \in \mathbb{X} \times \mathbb{C}$, and it is defined as Eq. (3).

$$\langle d, c \rangle = \langle \text{sequence(protein pockets embeddings, ligand embedding)}, c \rangle. \quad (3)$$

Following the approach proposed by Zhou et. al. [38], the goal is to learn a classifier γ that maps created sequences to $\{0, 1\}$.

$$\gamma : \mathbb{X} \to \{0, 1\}. \quad (4)$$

3.3.1 Self-Attention

Attention mechanism is a method for selectively concentrating on most relevant part of the input vector. It accomplishes this task by mapping a query and a set of key-value pairs to a weighted sum of the values, computed by the relationship of the query and the corresponding key. More specifically, given the query and associated key, the linear weights are computed based on some compatibility function [31].

Vaswani et al. [31] describes a particular attention called "Scaled Dot-Product Attention", where the input is composed of queries, keys and values. Instead of computing a dot product between the inputs and the query, this attention mechanism contains learnable parameters through adopting three trainable weight matrices. More specifically, as shown in Eq. 5, the queries, keys, and values are packed into matrices Q, K, and V, respectively Also, the output matrix is computed by calculating the dot product of the query with all the keys, divided by the square root of the dimension of the keys. The division of the square root of the dimension of the keys serves as a scaling factor to avoid pushing the softmax function into small gradient regions [31].

$$\text{Attention}(Q, K, V) = \text{Softmax}(\frac{QK^T}{\sqrt{d_k}})V. \quad (5)$$

This self-attention mechanism uses sequence of embeddings as input, and extracts query, key and value from each embedding. The attention output is then computed using Eq. 5.

3.3.2 BiLSTM

Long Short Term Memory (LSTM) is a variation of recurrent neural networks with three gates in its architecture: the input gate, forget gate, and output gate. The cells in an LSTM remember the information in the sequence for an arbitrary index, and the gates regulate the flow of information to and out of each cell. The forget gate, then, decides which information should be forgotten and which information should persist through the next cell. Zhou et al. [38] developed a BiLSTM network with two subnetworks for the forward and backward sequence context, respectively. Using an element-wise sum, the outputs of the forward and backward passes are then combined, as shown below:

$$\vec{h_i} = \text{LSTM}(x_i, \vec{h}_{i-1}) \tag{6}$$

$$\overleftarrow{h_i} = \text{LSTM}(x_i, \overleftarrow{h}_{i+1}) \tag{7}$$

$$h_i = [\vec{h}_i \oplus \overleftarrow{h}_i] \tag{8}$$

3.4 Classifier

The new values previously computed are then concatenated to a 1D vector I, and are passed to the classification layers. In this study, 2 fully connected layers are used to classify concatenated vector to either an active or inactive interaction. Eqs. 9 and 10 represent the two input and output layers of the classifier network, respectively:

$$O_h = \text{RELU}(W_{\text{input}} I + b_{\text{input}}), \tag{9}$$

$$O = \sigma(W_{\text{output}} O_h + b_{\text{output}}). \tag{10}$$

Sigmoid function is then used in the final layer 10 to predict the output in the form of a probability. Moreover, the following cross entropy loss function is used to train the model:

$$\mathcal{L}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]. \tag{11}$$

4 Experiments

4.1 Datasets

In our experiments, we compare our AttentionSiteDTI with several state-of-the-art methods, using three publicly available benchmark datasets: DUD-E dataset, Human dataset and the customized BindingDB dataset. We use the simplest docking-based method to find the binding sites of proteins [24]. We expect a boost in the performance of our model with incorporation of more complex (ML-based) binding site prediction algorithms, and/or higher-level computational physics approaches (e.g. molecular dynamics, density functional theory).

DUD-E This dataset [19] consists of 102 targets from 8 protein families. Each target has around 224 active compounds and more than 10,000 decoys, which were computationally generated in a way that their physical attributes are similar to active compounds but topologically dissimilar. We used three fold cross validation for our experiment, each fold was splitted based on the target, similar targets were kept in the same fold. We used random under sampling on the decoys to make the dataset balanced for training and used unbalanced dataset for evaluation.

Human This dataset [17] was built using a systematic screening framework to create credible and reliable negative sample pairs. the dataset consists of 5,423 interactions. We used the same split used in the DrugVQA [37] (80%,10%,10% random split for training,validation and test set) for a head-to-head comparison.

BindingDB This dataset [7] contains experimentally based assays of the interactions between small molecules and proteins. Following

Table 1: Dataset's Training Hyperparameters (FC is representing number of fully connected layers, L-GCN and P-GCN are number of Graph convolutional layers for extracting ligands and protein binding sites embedding respectively)

Datasets	FC	L-GCN	P-GCN
DUD-E	2	18	16
Human	6	4	3
Bindingdb	6	4	3

the work in DrugVQA [37], we used a small subset of the dataset, which consists of 39,747 positive and 31,218 negative samples. Further, in order to validate the generalization ability of the proposed model, the testing set was splited into two groups of the proteins; those that are seen in the time of training vs those that are not being seen by the model.

4.2 Implementation and evaluation strategy

Experimentation strategies. We used Pytorch 1.8.2 (long time support version) for implementing aforementioned model. We train the models for 30 epochs and used Adam optimizer for training the network with learning rate of 0.001. We used batch size of 100 for better generalization of the network along with a dropout with probability 0.3 after each fully connected layer. The GPU used for the experimentation was (Nvidia RTX 3090) with 24 GB of memory. We used 4 as number of hops in TAGCN for proteins and 2 for ligands. Size of the hidden state for BiLSTM layer in our model was set to 31, which was the output of the graph convolution layer (TAGCN). We used padding of zero to reshape each matrix to the maximum number of binding pockets in the datasets. Also, in order to prevent the attention layer to focus on relationship between different pockets of the protein, the corresponding values for inner protein relationships were set to zero. All other hyperparameters were tuned to yield the best result for each data set, which can be seen in Table 1.

Evaluation metrics. We evaluated our models in terms of several metrics, including AUC, which is the area under the receiver operating characteristic curve. In order to compare our performance against [37], we additionally report precision and recall for human dataset, ROC enrichment metric(RE) which is information on early enrichment of the precision recall curve which AUC does not contain such information. This metric is used on different thresholds and it can be defined as Equation 12. for DUD-E dataset and accuracy for BindingDB. Note that the RE score is defined as below:

$$RE = \frac{TPR}{FPR} \text{at a given FPR threshold.} \quad (12)$$

Ablation study Our ablation study was performed on three different benchmark datasets, which illusterate the effectiveness of several text classification methods in AttentionSiteDTI framework. We report AUC of all experiments, which is widely used in the literature. The results of this study can be found in Table 2 that shows that attention mechanism is the most effective method in text classification, and it is particularly advantageous because of the additional information on model explainability. Although, the attention mechanism is showing superior performance compared to Bi-LSTM, it is noteworthy that, for more challenging datasets, attention mechanism cannot capture the relationship between binding sites and ligands. Therefore, for instance in DUD-E dataset, which is intentionally generated in the way that the negative interactions are extremely close to positive ones, attention mechanism with Bi-LSTM architecture gives better results compared to only self-attention mechanism. The Bi-LSTM architecture cannot focus solely on interactions between ligand and

binding sites; therefore, it has inferior results compared to other proposed architectures.
Finally, it is worth mentioning that in our experiments, the TAGCN architecture for calculating graph embeddings has a better performance than GAT [32] and GCN [12] architecture, but the performance on other graph convolutional layers is yet to be explored.

Comparison on the DUD-E dataset On DUD-E dataset, we compared our proposed model with several state-of-the-art approaches. These models can be divided into 4 categories: (1) machine learning-based methods such as NN-score[4], and Random Forest-score (RF-score)[1]; (2) open source molecular docking programs including AutoDock Vina[29] and Smina[13]; (3) deep learning-based models such as AtomNet[33], 3D-CNN[23], which use neural networks to extract features from 3D structural information; and (4) graph-based models like PocketGCN[27], GNN[30], DrugVQA[37], which are all based on graph representations. PocketGCN utilizes two Graph-CNNs that automatically extracts features from the graph of protein pockets and ligands, in order to capture protein-ligand binding interactions. CPI-GNN[34] is a prediction model that combines a graph neural network (GNN) for compounds and a convolutional neural network (CNN) for targets. DrugVQA utilizes a 2D distance map to represent proteins in a Visual Question Answering system, where the images are the distance maps of the proteins, the questions are the SMILES of the drugs, and the answers are whether the drug-target pair will interact. Note that the scores of these models are derived form Zheng et. al.[37]. Also, following Zheng et al.'s work, we perform 3-fold cross-validation on this dataset, and report the average evaluation metrics. Also, we employ the F1 score and ROC enrichment (RE) metrics, where the RE score is defined as the ratio of the true positive rate (TPR) to the false positive rate (FPR) at a given FPR threshold. In Table 3, we report the RE scores at 0.5%, 1%, 2%, and 5% FPR thresholds. The results indicate that our model achieves state-of-the-art performance in DTI prediction on all metrics. As the results show, our AttentionSiteDTI has significant improvement at 0.5% RE. Also, we hypothesize that the poor performance of AtomNet and 3D-CNN may be due to the sparsity of 3D space, as they use the whole 3D structure of the proteins.

Comparison on the human dataset On Human dataset, we compared our model against several traditional machine learning models such as K nearest neighbors (KNN), random forest (RF), L2-logistic (L2)(these results were gathered from [18]); and some recently developed graph-based approaches including graph CNNs(GCNs)[12] , CPI–GNN[34], DrugVQA[37], TransformerCPI[2] as well as GraphDTA[20] that was originally designed for regression task, and was tailored to binary classification task by [36]. For a head-to-head comparison with other models, we followed the same experimental setting as in [15, 30]. Also, we repeated our experiments with three different random seeds, similar to DrugVQA[37]. The performances of the aformentioned models were obtained from [36], and are summarized in Table 4. It can be observed that the prediction accuracy of our proposed model is superior than all ML- and GNN-based models; and it achieves competitive performance with DrugVQA in terms of precision and recall. The relatively low performance of ML-based models is indeed in line with our expectation, and is due to their use of low-quality features, unable of capturing complex non-linear relationships in protein-drug interaction. The deep learning

Table 2: Ablation Study Results

Ablation tests	**BindingDB**	**Human**	**DUD-E**
Bi-LSTM	0.863	0.976	0.954
Self-Attention	**0.940**	**0.991**	0.961
Bi-LSTM + Self-Attention	0.925	0.984	**0.971**

Table 3: DUD-E Dataset Comparision

	AUC	0.5% RE	1.0% RE	2.0% RE	5.0% RE
NN Score	0.584	4.166	2.980	2.460	1.891
RF-score	0.622	5.628	4.274	3.499	2.678
Vina	0.716	9.139	7.321	5.811	4.444
Smina	0.696	-	-	-	-
3D-CNN	0.868	42.559	26.655	19.363	10.710
AtomNet	0.895	-	-	-	-
PocketGCN	0.886	44.406	29.748	19.408	10.735
GNN	0.940	-	-	-	-
DrugVQA	0.972	88.17	58.71	35.06	17.39
AttentionSiteDTI	**0.971**	**101.74**	**59.92**	**35.07**	**16.74**

models, on the other hand, are very powerful in extracting important features governing the complex interactions in a drug-target pair. On this basis, our model further improves on the accuracy, indicating that the quality of learned information in drug-target interactions is guaranteed by the back propagation of the end-to-end learning of our AttentionSiteDTI.

Comparison on the BindingDB dataset As the final experiment, on BindingDB dataset, we further compared our model against Tiresias[5], DBN[35], CPI-GNN [34], E2E[6], DrugVQA[37] and Bridge-DPI[36] as baselines. Tiresias uses similarity measures of drug and target pairs. DBN uses stacked restricted Boltzmann machines with the inputs in the form of extended connectivity fingerprints. As mentioned earlier, CPI-GNN combines a graph neural network (GNN) for compounds and a convolutional neural network (CNN) for targets to capture drug-target interactions. E2E is a GNN-based model that uses LSTM to learn drug-target pair information with Gene Ontology annotations. DrugVQA, as previously mentioned, is a Visual Question Answering system, where the images are the distance maps of the proteins, the questions are the SMILES of the drugs, and the answers are whether the drug-target pair will interact. Finally, BridgeDPI uses convolutional neural networks to obtain embeddings for drugs and proteins, as well as a GNN to learn the associations between proteins/drugs using some hyper-nodes that are connections between proteins/drugs. Note that the scores for all these models are derived from [36]. Also, following suggestions from previous works, we report the prediction results in terms of AUC and Accuracy (ACC) on the test set, which is divided into a set of unseen protein (the proteins that are not observed in training set) and a set of seen protein (the proteins that are observed in training set). This, indeed, makes the cus-

Table 4: Human Dataset Comparision

	AUROC	Precision	Recall	F1 Score
K-NN	0.86	0.798	0.927	0.858
RF	0.940	0.861	0.897	0.879
L2	0.911	0.861	0.913	0.902
SVM	0.910	0.966	0.950	0.958
GraphDTA	0.960	0.882	0.912	-
GCN	0.956	0.862	0.928	-
CPI–GNN	0.970	0.923	0.918	0.920
E2E/GO	0.970	0.893	0.914	0.903
DrugVQA	0.979	0.954	0.961	0.957
BridgeDPI	0.990	0.963	0.949	0.956
AttentionSiteDTI	**0.991**	**0.951**	**0.975**	**0.963**

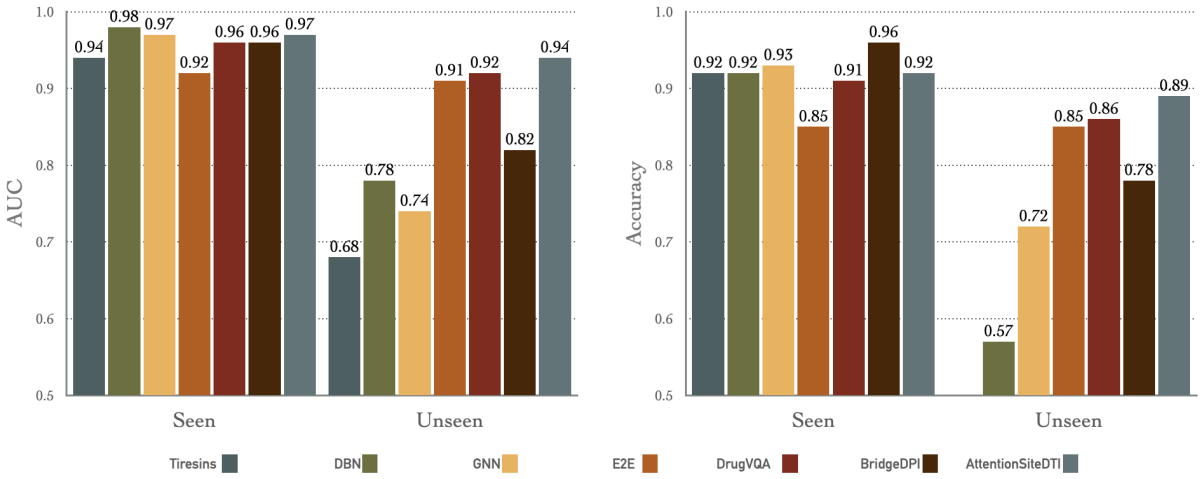


Figure 3: Comparison of AttentionSiteDTI with six baselines: (left) shows Area Under the Curve (AUC) for seen proteins and unseen proteins in the test; (right) shows Accuracy for seen proteins and unseen proteins in the test. Note that the accuracy scores of Tiresias do not show in unseen case, because it is lower than the lower bound of the y-axis (0.5). Note that for a head-to-head comparison with all models including ours, we implemented the BridgeDPI model with our experimental setting.

tomized BindingDB dataset suitable to assess models' generalization ability to unknown proteins, which should be the focus in prediction problems (i.e., cold-start problem), as there are a large number of unknown proteins in nature.

As experimental results indicate in Figure 3, all models generally perform well on seen proteins with AUC above 0.9, and ACC exceeding 0.85. However, these models show different and much worse performance on unseen proteins, which reflects the complexity of this more realistic learning scenario. Tiresias is a similarity-based model that uses a set of expert designed similarity measures as the features for proteins and drugs. The poor performance of Tiresias on the unseen proteins is perhaps due to the fact that these handcrafted features are not sufficient in capturing interactions between drug-target pairs, thus resulting in the accuracy even less than 0.5 on unseen proteins. On the other hand, the good performance of deep learning-based models including DBN, CPI-GNN, E2E, DrugVQA, BridgeDPI as well as our AttentionSiteDTI shows the effectiveness of these models in capturing relevant features that are critical in DTI prediction problem. As the results show, our model achieves the best performance with AUC of 0.97 and 0.94 on seen and unseen proteins, respectively. Also, in terms of accuracy, our AttentionSiteDTI outperforms all other models with accuracy reaching 0.89 in unseen proteins. This is an indication that our attention-based bidirectional LSTM network is, indeed, effective in relation classification of drug-target (protein pocket) pairs by learning the deeper interaction rules, governing the relationship between proteins' binding sites (pockets) and drugs. Although, the baselines seem to be more effective when the tested proteins are observed in the training, this indeed can be an indication of over-fitting.

Model Explainability Ligands bind to certain parts (active sites) of proteins either blocking the binding of other ligands or inducing a change in the protein structure, which produces a therapeutic effect. Binding at other sites that provide no therapeutic value are "non-active," and generally do not cause a direct biological effect.

In the case of drugs (ligands), which interact with proteins to prevent some biological process, binding to a specific protein site blocks a natural binding event (e.g. spike protein binding ACE2). Ligands binding to active sites and inducing a change in protein structure (conformation) are less likely in our system of study, and are probably not as helpful for building models (usually these ligands/therapeutic agents are employed/considered when a patient has an ailment, which causes natural biochemicals to be produced in insufficient quantities).

In this work the attention mechanism enables the model to predict which protein binding sites are more probable to bind with a given ligand. This probability is the attention matrix computed in the model. The attention visualization can be found in Figure 4 as the heat map of the protein for the complex of SARS-CoV2 Spike protein and human, host cell-expressing ACE2 (Angiotensin Converting Enzyme2; surface membrane protein; agonist binding partner of SARS-CoV2 spike protein) in the interaction with the drug named Darunavir. The projection of the heat map on the protein is depicted in this figure, as well.

5 Covid Case study and In Lab Validation

To further evaluate the practical potential of our proposed model, we experimentally tested and validated the binding interactions between seven candidate compounds and spike (or ACE2) protein. In-silico

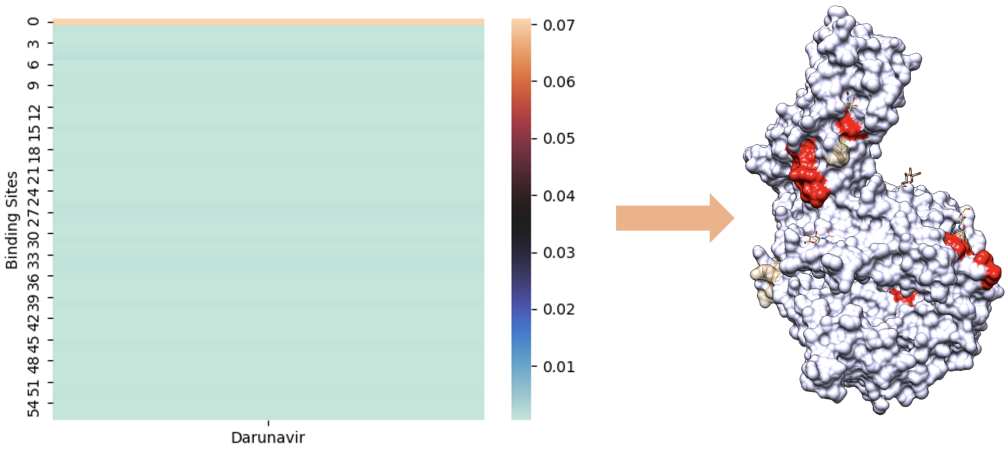


Figure 4: (Left) shows Heatmap of self-attention mechanism weights in the proposed method with input of Darunavir as ligand and complex of COVID spike protein and ACE2 as protein, which translates to probability of each calculated binding sites of the protein being active for that specific ligand.(Right) shows projected heatmap of self-attention weights on the complex of COVID spike protein and ACE2

predictions were performed on the binding interaction of seven candidate compounds including N-acetyl-neuraminic acid, 3α,6αMannopentaose, N-glycolylneuraminic acid, 2-Keto-3-deoxyoctonate ammonium salt, cytidine5-monophospho-N-acetylneuraminic acid sodium salt and Darunavir as inhibitor molecules to bind to the spike protein, or the ACE2 receptor protein, which has been shown to be the primary host factor recognized and targeted by SARS-CoV-2 Spike protein. To be more specific, the primary goal in our experimental investigations is to determine the ability of those seven compounds to disrupt the important interaction between spike protein and ACE2, which, in turn, leads to inability of SARS-CoV-2 virions to infect host cells (i.e. disrupting the host-virus complex which mediates infection). As the results show in Table 5, we observe high agreement (five out of seven matched results) between the predicted and experimentally-measured drug-target interactions, which illustrates the potential of our AttentionSiteDTI as an effective complementary pre-screening tool to accelerate the exploration, and recommendation of lead compounds with desired interaction properties toward their targets. Note that, in our experiment, we set the activity threshold to 15 nm to only capture highly active compounds; thereby, limiting the influence of interactions at neighboring sites and weak interactions with poor coordination to the binding site center.

6 Conclusion

In this work, we proposed an end-to-end Graph Convolutional Neural Network(GCNN)-based model, built on self-attention bidirectional Long Short-Term Memory mechanism, which captures any relationship between binding sites of a given protein and the drug in a sequence analogous to a sentence with relational meaning between its biochemical entities a.k.a. protein pockets and drug molecule. Our proposed framework enables learning which binding sites

Table 5: In-lab Validation of AttentionSiteDTI in the case study of Covid-19

Compound	**AttentionSiteDTI**	**Lab Results**
2-keto-3-deoxynononic acid	Non-interacting	Interacting
N-Glycolylneuraminic acid	Non-interacting	Interacting
Cytidine-5-monophospho-N-acetylneuraminic acid sodium salt	**Interacting**	**Interacting**
Darunavir	**Interacting**	**Interacting**
N-acetyl-neuraminic acid	**Non-interacting**	**Non-interacting**
N-Acetyllactosamine	**Non-interacting**	**Non-interacting**
3,6-Mannopentaose	**Non-interacting**	**Non-interacting**

of a protein interact with a given ligand, thus allows interpretability and better generalizability, while outperforms state-of-the-art methods in prediction of drug-target interaction. Moreover, we implemented a computational-experimental framework to experimentally test and validate the predicted binding interactions between seven candidate compounds and spike (or ACE2) protein. The results of our in-lab investigations also showed high agreement between the computationally-predicted and experimentally-observed binding interactions, which in turn, illustrated the potential of our AttentionSiteDTI as an effective pre-screening tool to accelerate the exploration and recommendation of lead compounds with desired interaction properties toward their targets.

7 Data and Code availability

All datasets, used in this paper, are publicly available. DUD-E dataset is available at `http://dude.docking.org`, Human dataset is available at `https://github.com/IBMInterpretableDTIP` and finally the customized BindingDB-IBM dataset can be found at `https://github.com/masashitsubaki/CPI_prediction/tree/master/`. We used 3D structures of proteins in Human dataset from `https://github.com/prokia/drugVQA`. Also, all instructions and codes for our experiments are available at `https://github.com/yazdanimehdi/AttentionSiteDTI`

8 Acknowledgement

We thank Ms.Katalina Biondi for discussions and her valuable feedback and comments on earlier versions of the manuscript.

References

[1] Pedro J. Ballester and John B. O. Mitchell. A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. *Bioinformatics*, 26(9):1169–1175, 2010.

[2] Lifan Chen, Xiaoqin Tan, Dingyan Wang, Feisheng Zhong, Xiaohong Liu, Tianbiao Yang, Xiaomin Luo, Kaixian Chen, Hualiang Jiang, and Mingyue Zheng. Transformercpi: improving compound–protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments. *Bioinformatics*, 36(16):4406–4414, 2020.

[3] Jian Du, Shanghang Zhang, Guanhang Wu, Jose M. F. Moura, and Soummya

Kar. Topology adaptive graph convolutional networks, 2018.

[4] Jacob D. Durrant and J. Andrew McCammon. NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *Journal of Chemical Information and Modeling*, 51(11):2897–2903, 2011.

[5] Achille Fokoue, Mohammad Sadoghi, Oktie Hassanzadeh, and Ping Zhang. Predicting drug-drug interactions through large-scale similarity-based link prediction. In *European Semantic Web Conference*, pages 774–789. Springer, 2016.

[6] Kyle Yingkai Gao, Achille Fokoue, Heng Luo, Arun Iyengar, Sanjoy Dey, and Ping Zhang. Interpretable drug target prediction using deep neural representation. In *IJCAI*, volume 2018, pages 3371–3377, 2018.

[7] Michael K Gilson, Tiqing Liu, Michael Baitaluk, George Nicola, Linda Hwang, and Jenny Chong. Bindingdb in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic acids research*, 44(D1):D1045–D1053, 2016.

[8] Joseph Gomes, Bharath Ramsundar, Evan N Feinberg, and Vijay S Pande. Atomic convolutional networks for predicting protein-ligand binding affinity. *arXiv preprint arXiv:1703.10603*, 2017.

[9] Tong He, Marten Heidemeyer, Fuqiang Ban, Artem Cherkasov, and Martin Ester. Simboost: a read-across approach for predicting drug–target binding affinities using gradient boosting machines. *Journal of cheminformatics*, 9(1):1–14, 2017.

[10] Mingjian Jiang, Zhen Li, Shugang Zhang, Shuang Wang, Xiaofeng Wang, Qing Yuan, and Zhiqiang Wei. Drug–target affinity prediction using graph neural network and contact maps. *RSC Advances*, 10(35):20701–20712, 2020.

[11] Mostafa Karimi, Di Wu, Zhangyang Wang, and Yang Shen. Deepaffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics*, 35(18):3329–3338, 2019.

[12] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017.

[13] David Ryan Koes, Matthew P. Baumgartner, and Carlos J. Camacho. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling*, 53(8):1893–1904, 2013.

[14] Yujia Li, Daniel Tarlow, Marc Brockschmidt, and Richard Zemel. Gated graph sequence neural networks, 2017.

[15] Hansaim Lim, Paul Gray, Lei Xie, and Aleksandar Poleksic. Improved genome-scale multi-target virtual screening via a novel collaborative filtering approach to cold-start problem. *Scientific reports*, 6(1):1–11, 2016.

[16] Jaechang Lim, Seongok Ryu, Kyubyong Park, Yo Joong Choe, Jiyeon Ham, and Woo Youn Kim. Predicting drug–target interaction using a novel graph neural network with 3d structure-embedded graph representation. *Journal of chemical information and modeling*, 59(9):3981–3988, 2019.

[17] Hui Liu, Jianjiang Sun, Jihong Guan, Jie Zheng, and Shuigeng Zhou. Improv-

ing compound–protein interaction prediction by building up highly credible negative samples. *Bioinformatics*, 31(12):i221–i229, 2015.

[18] Hui Liu, Jianjiang Sun, Jihong Guan, Jie Zheng, and Shuigeng Zhou. Improving compound–protein interaction prediction by building up highly credible negative samples. *Bioinformatics*, 31(12):i221–i229, 2015.

[19] Michael M Mysinger, Michael Carchia, John J Irwin, and Brian K Shoichet. Directory of useful decoys, enhanced (dud-e): better ligands and decoys for better benchmarking. *Journal of medicinal chemistry*, 55(14):6582–6594, 2012.

[20] Thin Nguyen, Hang Le, Thomas P Quinn, Tri Nguyen, Thuc Duy Le, and Svetha Venkatesh. Graphdta: Predicting drug–target binding affinity with graph neural networks. *Bioinformatics*, 37(8):1140–1147, 2021.

[21] Tapio Pahikkala, Antti Airola, Sami Pietilä, Sushil Shakyawar, Agnieszka Szwajda, Jing Tang, and Tero Aittokallio. Toward more realistic drug–target interaction predictions. *Briefings in bioinformatics*, 16(2):325–337, 2015.

[22] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

[23] Matthew Ragoza, Joshua Hochuli, Elisa Idrobo, Jocelyn Sunseri, and David Ryan Koes. Protein–ligand scoring with convolutional neural networks. *Journal of chemical information and modeling*, 57(4):942–957, 2017.

[24] Seyed Majid Saberi Fathi and Jack A. Tuszynski. A simple method for finding a protein's ligand-binding pockets. *BMC Structural Biology*, 14(1):18, 2014.

[25] Jeongtae Son and Dongsup Kim. Development of a graph convolutional neural network model for efficient prediction of protein-ligand binding affinities. *PloS one*, 16(4):e0249404, 2021.

[26] Marta M Stepniewska-Dziubinska, Piotr Zielenkiewicz, and Pawel Siedlecki. Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics*, 34(21):3666–3674, 2018.

[27] Wen Torng and Russ B. Altman. Graph Convolutional Neural Networks for Predicting Drug-Target Interactions. *Journal of Chemical Information and Modeling*, 59(10):4131–4149, 2019.

[28] Giuseppe Tradigo. *Protein Contact Maps*, pages 1771–1773. Springer New York, New York, NY, 2013.

[29] Oleg Trott and Arthur J Olson. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

[30] Masashi Tsubaki, Kentaro Tomii, and Jun Sese. Compound–protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. *Bioinformatics*, 35(2):309–318, 2018.

[31] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones,

Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.

[32] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018.

[33] Izhar Wallach, Michael Dzamba, and Abraham Heifets. Atomnet: a deep convolutional neural network for bioactivity prediction in structure-based drug discovery. *arXiv preprint arXiv:1510.02855*, 2015.

[34] Erniu Wang, Fan Wang, Zhihao Yang, Lei Wang, Yin Zhang, Hongfei Lin, and Jian Wang. A Graph Convolutional Network–Based Method for Chemical-Protein Interaction Extraction: Algorithm Development. *JMIR Medical Informatics*, 8(5):e17643, 2020.

[35] Ming Wen, Zhimin Zhang, Shaoyu Niu, Haozhi Sha, Ruihan Yang, Yonghuan Yun, and Hongmei Lu. Deep-learning-based drug–target interaction prediction. *Journal of proteome research*, 16(4):1401–1409, 2017.

[36] Yifan Wu, Min Gao, Min Zeng, Feiyang Chen, Min Li, and Jie Zhang. BridgeDPI: A Novel Graph Neural Network for Predicting Drug-Protein Interactions. *arXiv*, 2021.

[37] Shuangjia Zheng, Yongjian Li, Sheng Chen, Jun Xu, and Yuedong Yang. Predicting drug–protein interaction using quasi-visual question answering system. *Nature Machine Intelligence*, 2(2):134–140, 2020.

[38] Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu. Attention-based bidirectional long short-term memory networks for relation classification. In *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)*, pages 207–212, 2016.

Supplementary Material

To assess the validity of determined models, for the current study, predictions were tested via standard laboratory binding assay. The assay chosen relies on the inhibition of complex formation between SARS-CoV2 spike protein and human ACE2 binding partner. Addition of candidate inhibitor molecule causes binding to a complex component, preventing the native binding, and reducing the intensity of a luminescence signal. Candidate inhibitor molecules were chosen for their (dis-)similarity to known inhibiting compounds from literature, as well as to probe the sensitivity of model predictions; specifically: N-acetyl-neuraminic acid, cytidine-5-monophospho-N-acetylneuraminic acid sodium salt, Darunavir (from Sigma Aldrich), 3α,6α-Mannopentaose, N-glycolylneuraminic acid, and 2-Keto-3-deoxyoctonate ammonium salt (from Fisher Scientific) were investigated.
ELISA-type ACE2: SARS-CoV-2 Spike Inhibitor Screening assay (Spike S1 RBD; BPS Bioscience Cat : 79936) was used per manufacturer protocol/procedure. All candidate inhibitor compounds were tested over a concentration range from 0.01 μM to 30 μM. To conduct the assay, stock ACE2 protein was thawed on ice and diluted to 1 μg/mL with phosphate buffered saline (PBS). From here, 50 μL of ACE2 diluted sample was added to a nickel-coated 96 well plate and incubated for 1 hr. at room temperature on a shake-table at low speed. The plate was then washed 3x with PBS and incubated for 10 minutes in a blocking buffer. 10 μL of the candidate inhibitor molecule, dispersed to desired concentration in PBS, was subsequently added and incubated at room temperature for 1 hr. under slow shaking. An inhibitor buffer containing 5% DMSO was used as positive control and blank measurements. SARS-CoV-2 Spike (RBD)-Fc was thawed under similar conditions to ACE2 and diluted to 0.25 ng/mL in Assay Buffer 1. Spiked protein was added to each well (except blank wells) and shaken slowly on the shake-table for an additional 1 hr. at room temperature. Plates were then washed 3x with PSB and incubated in Blocking Buffer for 10 minutes. Anti-mouse Fc-HRP (horseradish peroxidase) was then added and incubated for 1 hr. under slow shaking. Lastly, HRP substrate was added and chemiluminescence was measured using a FluoroStar Omega microplate reader. Reduction in chemiluminescence, relative to the negative control, is interpreted as disruption of the Spike-ACE2 complex formation and the consequential washing away of the HRP chemiluminescence producing substrate. Spike protein standard curves were produced via titration from 0.1 to 100 nM to determine/compare concentration dependent luminescence signal.

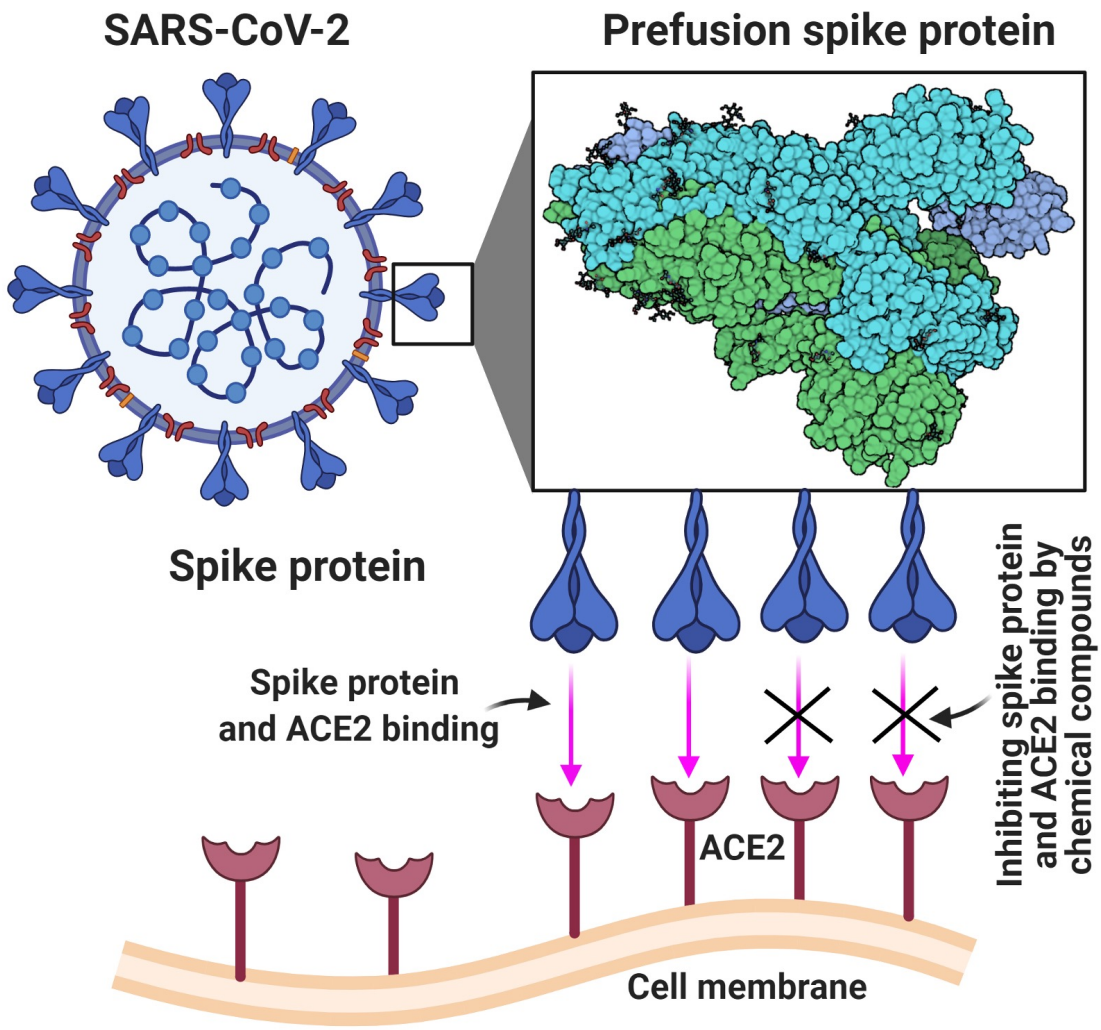
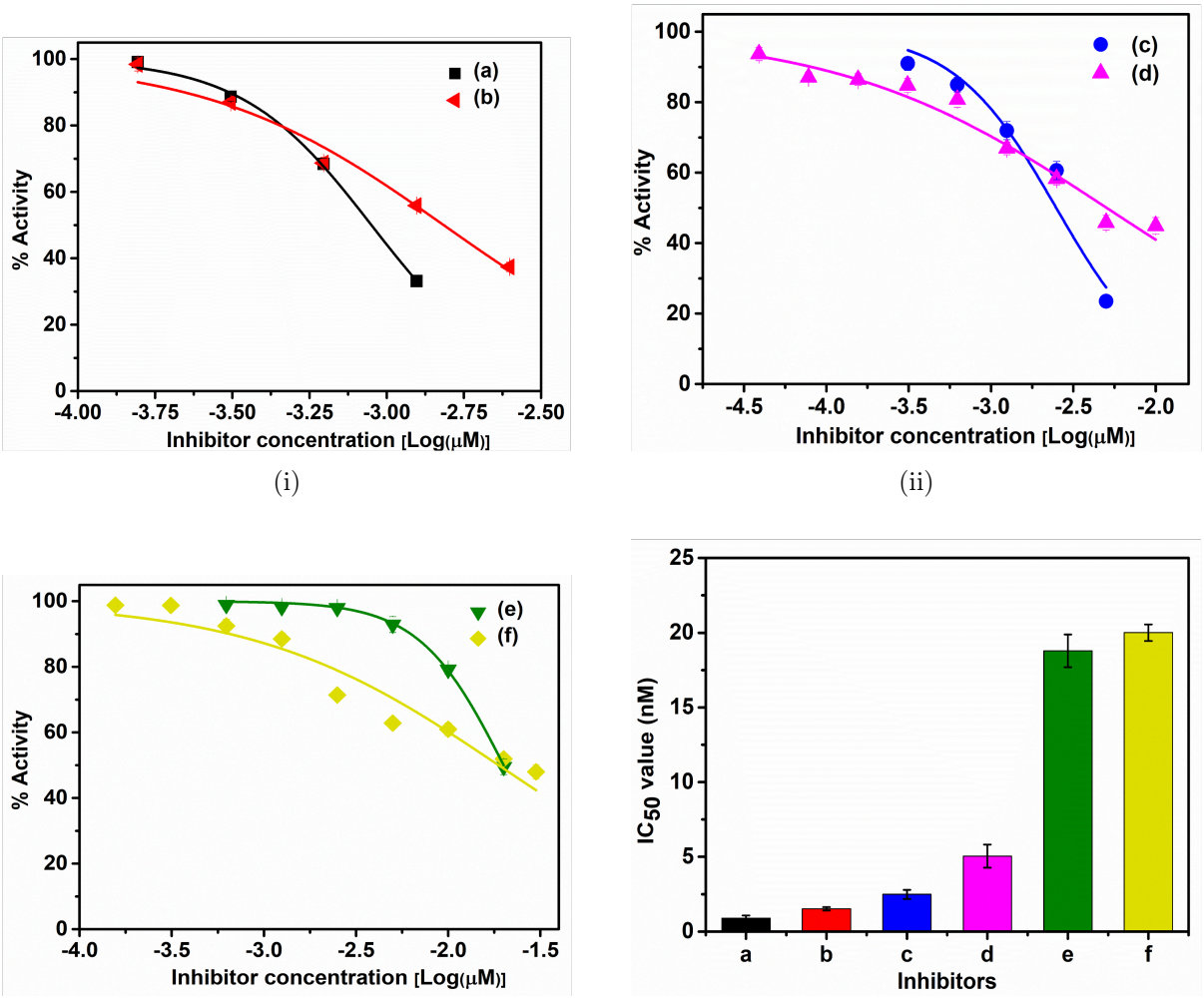


Figure 5: The schematic diagram represents binding between ACE2 protein on cell membrane and SAR-CoV-2 spik protein that causes coronavirus infection. Currently, many pharmacological and non-pharmacological compounds have been tested as the inhibitory compounds to protect against current pandemic infection disease. We have chosen a set of inhibitor's such as 2-Keto-3-deoxyoctonate ammonium salt, cytidine-5-monophospho-N-acetylneuraminic acid sodium salt, Darunavir, (N-acetyl-neuraminic acid, 3α,6α-Mannopentaose, and N-glycolylneuraminic acid to test the inhibition concentration under in vitro condition using the BPS bioscience ACE-2:SARS-CoV-2 spike inhibitor screening ELISA assay kit.



(i) (ii) (iii) (iv)

Figure 6: A family of silicic acid molecules and darunavir drug were used as a inhibitors to test for inhibition of the binding between spike and ACE2 protein using ACE-2:SARS-CoV-2 spike inhibitor screening ELISA assay kit. To assess the concentration of inhibition, the different concentration (0.1 nM to 30 nM) of silicic acid and drug molecules were added to screening assay kit and chemiluminescence was measured using microplate reader (FluoStar Omega luminesces fluorescent mode) according to the manufacture instructions. Fig's. 2(I), (II) and (III) shows graphs of inhibitor log concentration vs percentage activity graphs of 2-Keto-3-deoxyoctonate ammonium salt (a), N-glycolylneuraminic acid (b), cytidine-5'-monophospho-N-acetylneuraminic acid (c), darunavir (d), N-acetyl-neuraminic acid (e), and N-acetyllactosamine (f). The non-linear equation function associated with dose-response curve fitting method was used to fit the all dates and obtain the IC50 value. Fig. 2 (IV) shows the IC50 concentration of all tested compounds results. The experiment was performed in duplicates for all compounds.