

# BINDTI: A bi-directional Intention network for drug-target interaction identification based on attention mechanisms

Lihong Peng, Xin Liu, Long Yang, Longlong Liu, Zongzheng Bai, Min Chen\*, Xu Lu\*, Libo Nie\*

**Abstract**— The identification of drug-target interactions (DTIs) is an essential step in drug discovery. In vitro experimental methods are expensive, laborious, and time-consuming. Deep learning has witnessed promising progress in DTI prediction. However, how to precisely represent drug and protein features is a major challenge for DTI prediction. Here, we developed an end-to-end DTI identification framework called BINDTI based on bi-directional Intention network. First, drug features are encoded with graph convolutional networks based on its 2D molecular graph obtained by its SMILES string. Next, protein features are encoded based on its amino acid sequence through a mixed model called ACmix, which integrates self-attention mechanism and convolution. Third, drug and target features are fused through bi-directional Intention network, which combines Intention and multi-head attention. Finally, unknown drug-target (DT) pairs are classified through multilayer perceptron based on the fused DT features. The results demonstrate that BINDTI greatly outperformed four baseline methods (i.e., CPI-GNN, TransformerCPI, MolTrans, and IIFDTI) on the BindingDB, BioSNAP, DrugBank, and Human datasets. More importantly, it was more appropriate to predict new DTIs than the four baseline methods on imbalanced datasets. Ablation experimental results elucidated that both bi-directional Intention and ACmix could greatly advance DTI prediction. The fused feature visualization and case studies manifested that the predicted results by BINDTI were basically consistent with the true ones. We anticipate that the proposed BINDTI framework can find new low-cost drug candidates, improve drugs' virtual screening, and further facilitate drug repositioning as well as drug discovery. BINDTI is publicly available at <https://github.com/plhhnu/BINDTI>.

**Index Terms**— Drug-target interaction, bi-directional Intention network, attention, convolution, graph convolutional network.

## I. INTRODUCTION

Identification of potential linkages between drugs and targets plays an important role in drug discovery [1]–[7]. In vitro experiments screen many credible drug-target interactions (DTIs). However, laboratory techniques are time-consuming, labor-intensive, and expensive [8]–[10]. Consequently, accumulating computational tools have been developed to identify potential DTIs. These tools mainly contain docking-based methods, network-based methods, and machine learning-based methods [11].

Molecular docking-based DTI prediction methods mainly foresee binding ability for two small molecules [12]. Shaikh et al. [13] used molecular docking to compute chemical interaction affinity for DTI prediction. Yang et al. [14] proposed a ligand-protein docking model named FitDock to fit initial conformation through a hierarchical multi-feature alignment algorithm and then output precise docking poses between ligands and proteins. Molecular docking approaches require high-quality 3D structure information for computing the affinity strength of ligand-protein pairs. However, their co-crystal structures are very difficult to be determined, leading to that docking-based methods were subject to extremely computational resource consumption and were excessively dependent on the 3D structure information [15], [16].

Network-based methods first depict drugs and proteins as nodes and DTIs as edges, and then compute drug and target similarity matrices, finally use graph theoretic algorithms to predict DTI candidates [11]. These methods contain random walk [17], probabilistic soft logic model [18], attentive meta-path extraction [19], non-negative matrix factorization [20], heterogeneous information network [21], the combination of unbalanced bi-random walk and Laplacian regularized least squares, and network-based virtual screening [22]. Network-based methods fully adopt drug and protein similarities, but fail to infer DTIs for a new drug and/or target without interaction information.

Recently, machine learning techniques have attained widespread attention in bioinformatics because of their good performance [23]–[26]. In particular, deep learning can leverage diverse biomedical data and has obtained wide applications, such as lncRNA-protein interaction prediction [27], protein structure modeling [28]–[30], identification of linkages between various RNA data and diseases [31]–[38], and single cell data analysis [39]–[42]. Notably, deep learning has been

Manuscript received September 28, 2023; revised \*\*\*.

This manuscript was supported by National Natural Science Foundation of China under Grant No. 61803151 and 62172158.

Lihong Peng, Xin Liu, Long Yang, Longlong Liu, Zongzheng Bai, and Libo Nie are with College of Life Science and Chemistry, Hunan University of Technology, Zhuzhou, 412007, Hunan, China, e-mail: plhhnu@163.com (Lihong Peng); liuxinhut@163.com (Xin Liu); yl\_hut@163.com (Long Yang); soulorlong@163.com (Longlong Liu); zozeba@163.com (Zongzheng Bai); nielibo@hut.edu.cn (Libo Nie).

Min Chen is with School of Computer Science, Hunan Institute of Technology, Hengyang, 421002, China.

Xu Lu is with School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou, 510665, China.

Xin Liu is co-first author.

Corresponding authors: Min Chen (chenmin@hnit.edu.cn); Xu Lu (bruda@126.com); Libo Nie (nielibo@hut.edu.cn).

broadly applied to DTI inference. These methods mainly consist of two steps: drug and protein feature representation and DTI classification based on deep learning. In general, physicochemical properties as well as biological and topological features of drugs and/or targets have been used to generate Drug-Target (DT) feature vectors. These methods contain stacked autoencoder [43], the combination of multi-head attention and skip connection [44], meta-learning framework with subgraph matching [45]. More importantly, graph convolutional network (GCN) [46], graph attention network (GAT) [15], [47], and graph neural network (GNN) [48]–[50] have been broadly applied to encode local structures of drugs based on their 2D molecular graphs. And convolution neural network (CNN) [8], [46], [51]–[54] has been used to attain protein features based on their sequences. Particularly, one-side attention, i.e., neural attention mechanism [48], has acquired the joint representation of drug's molecular structures and protein's subsequences. The newest studies [15], [46], [54]–[56] fused drug and target features based on two-sided attention mechanism.

In this manuscript, we developed BINDTI, an end-to-end framework to identify DTI candidates. This study has four main contributions:

- An end-to-end framework called BINDTI is developed to predict potential DTIs based on attention mechanisms.
- A mixed model called ACmix is designed to encode protein features by incorporating convolution and self-attention.
- A bi-directional Intention network is proposed to fuse drug and protein features based on multi-head attention and Intention.
- Visualization of the fused features manifests that the predicted DTIs are basically consistent with the true ones.

## II. MATERIALS AND METHODS

### A. Datasets

The BINDTI performance is evaluated on four publicly available datasets: DrugBank [52], BindingDB [57], BioSNAP [58], and Human [59], with test sets composed of unobserved DT pairs. The DrugBank dataset is created by Zhao et al. [52] based on DTI data from the DrugBank database [60]. BindingDB [57] contains experimentally verified binding affinities, mainly including interactions between small molecule drugs and proteins. The BioSNAP dataset is a balanced dataset and was created by Huang et al. [61] and Marinka et al. [58] based on DTI data from the DrugBank database [60]. In this dataset, positive DTIs have been validated and negative DTIs were randomly selected from unseen pairs. Human [59] contains highly credible negative DTIs through a bioinformatics tool. Table I shows statistics of the four DTI datasets.

### B. BINDTI

In this manuscript, we developed BINDTI, an end-to-end framework to predict new DTIs. BINDTI mainly consists of four procedures: (i) encoding drug features with GCNs based on its 2D molecular graph obtained by its SMILES string; (ii) encoding protein features with ACmix based on its amino

**TABLE I:** The introduction of four DTI datasets

Dataset	Drug	Protein	Interaction	Positive	Negative
BindingDB	14643	2623	49199	20674	28525
BioSNAP	4505	2181	27464	13830	13634
DrugBank	6636	4254	34992	17501	17491
Human	2726	2001	5997	2633	3364

acid sequence; (iii) fusing drug and protein features through bi-directional Intention network and characterizing one DT pair as a feature vector; (iv) classifying unknown DT pairs via multilayer perceptron based on the fused DT features. The BINDTI pipeline is illustrated in Fig. 1.

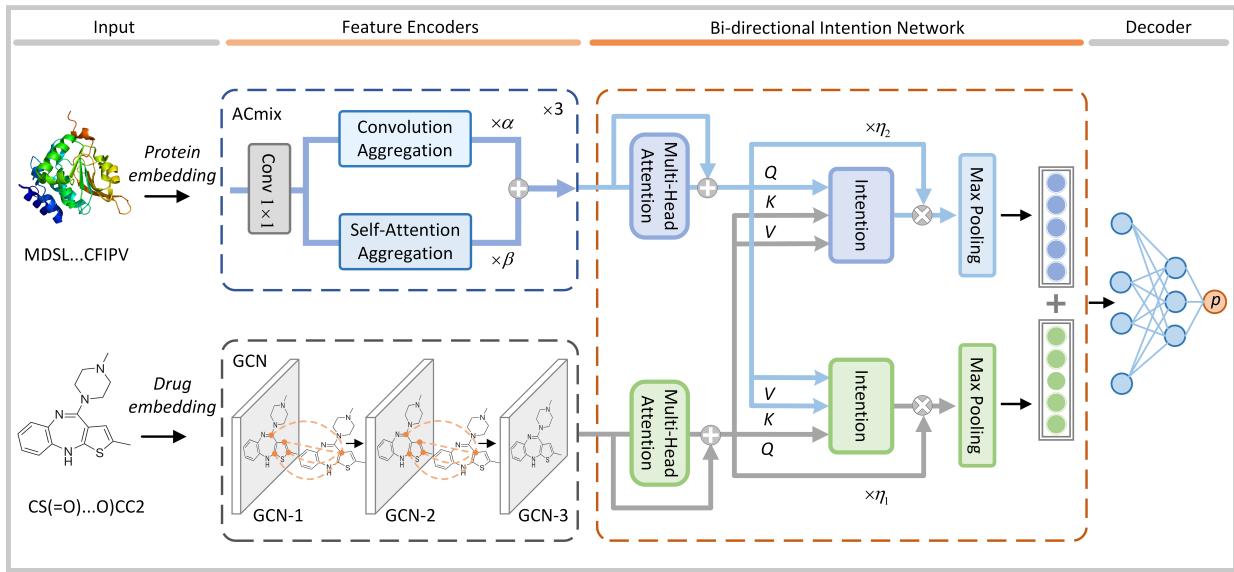
**1) Problem Formulation:** For a DTI dataset  $D = \{(\mathcal{G}_i, \mathcal{P}_i, y_i)\} (i = 1, 2, \dots, n)$  with  $n$  DT pairs, let  $(\mathcal{G}, \mathcal{P})$  denote a certain DT pair (i.e., one sample) and  $y$  denote its label.  $y = 1$  if the DT pair has an interaction, otherwise  $y = 0$ . Next, a DT pair is determined whether there is a linkage based on bi-directional Intention network with attention mechanisms and multilayer perceptron.

**2) Encoding drug features with GCN:** To obtain drug features, inspired by the drug feature extraction method proposed by Ref. [46], BINDTI first converts drug's SMILES string [62] into a 2-dimensional molecular graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  and  $\mathcal{E}$  denote the set of vertices (i.e., atoms) and set of edges (i.e., chemical bonds), respectively. To characterize each node in  $\mathcal{G}$ , first, each node is initialized based on their chemical properties through the DGL-LifeSci package [63]. Consequently, each node is denoted as a 74-dimensional vector with integer values. Next, since each type of drugs contain different atom numbers, a maximum atom number  $N_d$  is set to convert each drug as a tensor with the same dimension. Particularly, when a drug contains less atoms, we pad the obtained 74-dimensional vector with zero for each node. Consequently, each drug can be represented as  $\mathbf{M}_d \in \mathbb{R}^{N_d \times 74}$ . Finally, a dense matrix with real values  $\mathbf{Z}_d \in \mathbb{R}^{N_d \times D_d}$  is computed through a simple linear transformation model  $\mathbf{Z}_d = \mathbf{W}_0 \mathbf{M}_d^T$  and is further used as the input in GCN.

Subsequently, a three-layer GCN is constructed to learn drug's graph representation. The feature vector of each atom is updated by aggregating information of atoms neighboring to the atom. Both a atom and its neighbor atoms are linked by chemical bonds. As a result, we automatically acquire substructure information of a drug compound based on its neighborhoods. Thus, the drug encoder is represented as Eq. (1):

$$\mathbf{H}_G^{(l+1)} = \text{ReLU}(\tilde{\mathbf{A}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}_G^{(l)} \mathbf{W}^{(l)}) \quad (1)$$

where  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$  denotes an adjacency matrix corresponding to the undirected graph  $\mathcal{G}$  with additive self-connections, where  $\mathbf{A}$  and  $\mathbf{I}$  indicate two adjacency matrices corresponding to  $\mathcal{G}$  and identity matrix, respectively.  $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ ,  $\mathbf{W}^{(l)}$  denotes the trainable weights in the  $l$ -th layer of GCN with an activation function  $\text{ReLU}$ , and  $\mathbf{H}_G^{(l)}$  denotes outputs in the  $l$ -th layer with  $\mathbf{H}_G^{(0)} = \mathbf{Z}_d$ .



**Fig. 1:** An illustration of the proposed BINDTI model. It contains the following four steps: (i) encoding drug features with GCN based on drug's SMILES string; (ii) encoding protein features with ACmix based on its amino acid sequence; (iii) fusing drug and protein features with bi-directional Intention network; (iv) predicting interaction probability for each DT pair based on multilayer perceptron.

**3) Encoding target protein features with ACmix:** Convolution and self-attention have powerful representation learning performance. ACmix [64] is a mixed model with less computational overhead by enjoying the superiority of convolution and self-attention. To encode protein features, first, all 23 types of amino acids are initialized into a learnable feature matrix  $\mathbf{E}_p \in \mathbb{R}^{23 \times D_p}$ , where  $D_p$  denotes dimension in the latent space based on the principle of word embedding. Next, a protein sequence  $\mathcal{P}$  comprised of  $m$  amino acids can be initialized as  $\mathbf{X} \in \mathbb{R}^{D_p \times 1 \times N_p}$  based on the embedding matrix  $\mathbf{E}_p \in \mathbb{R}^{23 \times D_p}$ . Particularly,  $N_p$ , which represents the maximum amino acid number of a protein, is used to obtain all protein sequences with the same length for batch training. When the length of a protein sequence is less than  $N_p$ , corresponding columns in  $\mathbf{X}$  are padded with zeros. Consequently, we learn protein features  $\mathbf{Y}_{conv}$  based on the above obtained protein feature tensor  $\mathbf{X}$  and two stages of a standard convolution operation by Eq. (2):

$$\begin{aligned} \text{I: } \mathbf{Y}_{ij}^{p,q} &= \mathbf{K}_{p,q} \mathbf{X}_{ij} \\ \text{II: } \mathbf{Y}_{ij}^{p,q} &= Shift(\mathbf{Y}_{ij}^{p,q}, p - \lfloor k/2 \rfloor, q - \lfloor k/2 \rfloor) \quad (2) \\ \mathbf{Y}_{ij} &= \sum_{p,q} \mathbf{Y}_{ij}^{p,q} \end{aligned}$$

where  $\mathbf{K}_{p,q} \in \mathbb{R}^{C_{in} \times C_{out}}$  ( $p, q \in \{0, 1, \dots, k-1\}$ ) denotes the kernel weights corresponding to the indices of the kernel position  $(p, q)$ , and  $C_{in}$  and  $C_{out}$  represent the channel sizes in the input and output layers, respectively. Tensors  $\mathbf{X} \in \mathbb{R}^{C_{in} \times H \times W}$  and  $\mathbf{Y} \in \mathbb{R}^{C_{out} \times H \times W}$  with the height value of  $H$  and the width value of  $W$  indicate feature maps in the input and output layers, respectively.  $\mathbf{X}_{ij} \in \mathbb{R}^{C_{in}}$  and  $\mathbf{Y}_{ij} \in \mathbb{R}^{C_{out}}$  represent the  $i$ -th feature tensor of the  $j$ -th amino acid with regard to  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively. The Shift operation  $\tilde{f} \triangleq Shift(f, \Delta x, \Delta y)$  is defined as  $\tilde{f} = f_{i+\Delta x, j+\Delta y}, \forall i, j$ .

Moreover, we also obtain each protein's feature tensor  $\mathbf{Y}_{att}$  based on  $\mathbf{X}$  and two stages of multi-head self-attention by Eq. (3):

$$\begin{aligned} \text{I: } \mathbf{Q}_{ij}^{(l)} &= \mathbf{W}_q^{(l)} \mathbf{X}_{ij}, \quad \mathbf{K}_{ij}^{(l)} = \mathbf{W}_k^{(l)} \mathbf{X}_{ij}, \quad \mathbf{V}_{ij}^{(l)} = \mathbf{W}_v^{(l)} \mathbf{X}_{ij} \\ \text{II: } Att(\mathbf{Q}_{ij}^{(l)}, \mathbf{K}_{ab}^{(l)}) &= softmax_{\mathcal{N}_k(i,j)} \left( \frac{(\mathbf{Q}_{ij}^{(l)})^\top \mathbf{K}_{ij}^{(l)}}{\sqrt{d}} \right) \quad (3) \\ \mathbf{Y}_{ij} &= \left\| \sum_{l=1}^{N_h} \left( \sum_{a,b \in \mathcal{N}_k(i,j)} Att(\mathbf{Q}_{ij}^{(l)}, \mathbf{K}_{ab}^{(l)}) \mathbf{V}_{ab}^{(l)} \right) \right\| \end{aligned}$$

where  $\|\cdot\|$  represents the concatenation operation over the outputs on  $N_h$  attention heads,  $d$  denotes the feature dimension of  $\mathbf{Q}_{ij}^{(l)}$ ,  $\mathbf{W}_q^{(l)}$ ,  $\mathbf{W}_k^{(l)}$ , and  $\mathbf{W}_v^{(l)}$  are three projection matrices involving queries, keys and values in multi-head attention module.  $\mathcal{N}_k(i,j)$  denotes a region composed of  $k$  nodes centered around  $(i,j)$ , and  $Att(\mathbf{Q}_{ij}^{(l)}, \mathbf{K}_{ab}^{(l)})$  indicates the attention weight corresponding to features within  $\mathcal{N}_k(i,j)$ .

As shown in Fig. 1, the aforementioned operations elegantly integrate convolution and self-attention. Since they implement the same  $1 \times 1$  convolution operations, they can be only conducted one projection. We simultaneously perform convolution aggregation operation (i.e., Eq. (2) II) and self-attention aggregation operation (i.e., Eq. (3) II) after projection. Finally, the protein encoder is defined as Eq. (4) based on the outputs  $\mathbf{Y}_{conv}$  and  $\mathbf{Y}_{att}$  on convolution and self-attention:

$$\mathbf{H}_{\mathcal{P}} = \alpha \mathbf{Y}_{conv} + \beta \mathbf{Y}_{att} \quad (4)$$

where  $\alpha$  and  $\beta$  are two learnable parameters used to represent the importance of protein feature tensors obtained by convolution and self-attention on the final protein feature encoding, respectively. Their initial values are set to 0.5.

#### 4) Fusing drug and target features with bidirectional Intention

**network:** Multi-head attention [65] and Intention [66] are powerful deep learning techniques. To fuse drug and target features, we developed a bi-directional Intention network by integrating multi-head attention and Intention. Differed from previous works [44], [48], we fully utilize attention of atoms to protein and one of amino acid to drug and exploit a bi-directional attention to effectively integrate biological features of each DT pair through a neural network.

Multi-head attention [65] facilitates the model to jointly deal with information from different feature representation subspaces. The robustness of multi-head attention is obviously boosted when adding a residual module [16]. Inspired by transformer [65] and ResNet [67], we designed a multi-head self-attention residual module to enhance drugs' feature representation by Eq. (5):

$$\mathbf{H}_D = \text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) + \mathbf{H}_G \quad (5)$$

where  $\mathbf{Q}$ ,  $\mathbf{K}$  and  $\mathbf{V}$  are generated through a projection function  $f = \mathbf{w}^T \mathbf{x} + b$  based on the input  $\mathbf{H}_G$ , where  $\mathbf{w}$  and  $b$  denote weight and bias, respectively.

Similarly, we enhance target proteins' feature representation by Eq. (6):

$$\mathbf{H}_T = \text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) + \mathbf{H}_P \quad (6)$$

Garnet et al. [66] proposed the KVQ space and the Intention module. Intention, as a generalization version of attention, can serve as a unified linkage for many other methods. As shown in Fig.2, we build an Intention module for drug and protein feature fusion by Eq. (7):

$$H_{\text{int}}(\mathbf{K}, \mathbf{V}, \mathbf{Q}) = \mathbf{E}_Q \left[ (\mathbf{E}_K^T \mathbf{E}_K + \gamma \mathbf{I})^{-1} \mathbf{E}_K^T \mathbf{E}_V \right] \quad (7)$$

where  $\gamma$  is a learnable parameter.  $\mathbf{E}_K$ ,  $\mathbf{E}_V$ , and  $\mathbf{E}_Q$  represent learnable multilayer perceptron-based embeddings with regard to  $\mathbf{K}$ ,  $\mathbf{V}$ , and  $\mathbf{Q}$ , respectively.  $\mathbf{I}$  is an identity matrix.

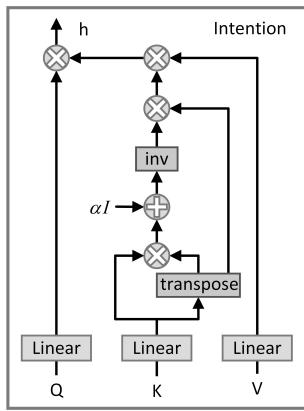


Fig. 2: The Intention block.

Based on the Intention module, the fused drug feature matrix  $\mathbf{f}_D$  and protein feature matrix  $\mathbf{f}_T$  can be represented as Eqs. (8) and (9), respectively:

$$\mathbf{f}_D = \eta_1 \mathbf{H}_D \sigma(H_{\text{int}}(\mathbf{H}_T, \mathbf{H}_T, \mathbf{H}_D)) \quad (8)$$

$$\mathbf{f}_T = \eta_2 \mathbf{H}_T \sigma(H_{\text{int}}(\mathbf{H}_D, \mathbf{H}_D, \mathbf{H}_T)) \quad (9)$$

where  $\eta_1$  and  $\eta_2$  are two learnable parameters, and  $\sigma$  denotes a nonlinear activation function *softmax*.

Finally, the fused feature matrices  $\mathbf{f}_D$  and  $\mathbf{f}_T$  with respect to drugs and proteins are concatenated by a max-pooling layer by Eq. (10):

$$\mathbf{f} = \text{Concat}(\text{maxpool}(\mathbf{f}_D), \text{maxpool}(\mathbf{f}_T)) \quad (10)$$

**5) Inference and Training:** We obtain the feature representation  $\mathbf{f}$  of each DT pair based on bi-directional Intention network. Subsequently, the fused DT features  $\mathbf{f}$  are fed into multilayer perceptron with four fully connected layers. Finally, the interaction probability  $p$  for each DT pair is computed by Eq. (11):

$$p = \text{Sigmoid}(\mathbf{W}\mathbf{f} + b) \quad (11)$$

Given an interaction probability threshold  $\delta$ , the DT pair is taken as interacting when  $p > \delta$ , otherwise, the DT pair is taken as not interacting.

During training, we minimize the following cross entropy loss to obtain the optimal solution of BINDTI by Eq. (12):

$$\arg \min_{\Theta} - \sum_i (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) + \lambda \|\Theta\|_2^2 \quad (12)$$

where  $\Theta$  denotes a set with respect to all learnable weight matrices and bias vectors,  $y_i$  and  $p_i$  represent the ground-truth and predicted labels of the  $i$ -th DT pair, respectively, and  $\lambda$  indicates a hyperparameter used to L2 regularization.

## III. RESULTS

### A. Experimental setup

AUC, AUPRC, accuracy, F1-score and Matthews' Correlation Coefficient (MCC) were taken as evaluation metrics to assess the BINDTI performance. BINDTI was implemented using Pytorch. We first trained the BINDTI model and obtained the optimal hyperparameter combination on the Human dataset. The other parameters in BINDTI were trained and their optimal values were determined when BINDTI computed the best performance on the validation set. The final performance of BINDTI was obtained on the test set. For BINDTI, we set epoch\_size=100, batch\_size=64, learning\_rate=1e-4, learning\_decay=0.5, weight\_decay=1e-5, atom\_dim=128, amino\_acid\_dim=128, ACmix\_layers=3, GCN\_layers=3, attention\_heads=8, and MLP\_hidden\_size=512.

To simulate real DTI prediction application scenarios and assess the generalization ability and reliability of the models, we compared BINDTI with other four baseline methods under two different experimental settings [46], [61]:

$E_1$ : Each dataset is randomly partitioned into three sets: training set, validation set, and test set at a ratio of 7:1:2. Under this setting, the distributions of the sizes of drugs (i.e., the number of atoms) and targets (i.e., the number of amino acids) in the validation/training/test sets were described in the form of categorical scatter plots. The results are available at: <https://github.com/plhhnu/BINDTI/tree/main/Distributions/distributions>. In addition, we still assessed the distribution of

similarity of drugs between the training and test sets as well as distribution of similarity of targets between the two sets. The similarity between drugs was calculated using the RDKit python toolkit, which is available at <https://www.rdkit.org/>. The similarity between proteins was calculated based on the method provided by Ref. [68]. Consequently, the distribution of similarity of drugs and the distribution of similarity of targets between the training and test sets are available at: <https://github.com/plhhnu/BINDTI/tree/main/Distributions/Similarity>.

$E_2$ : For each dataset, first, 10% of all DT pairs was randomly selected and was partitioned into validation set and test set at a ratio of 1:2; next, the remaining 90% of DT pairs were taken as training set; finally, DT pairs on the training set, where drugs and/or targets were appeared in the validation and test sets, were removed from the training set.

Under both  $E_1$  and  $E_2$ , we independently partitioned each dataset for five times through different random seeds, and the obtained partition at each time was independently run on BINDTI and four baseline methods. That is, the experiments for evaluating BINDTI and four baselines were repeated for five times. The average ( $\pm$ standard deviation) of the results from the five independent runs was regarded as the final performance.

Furthermore, we conducted 5-fold cross validation for evaluating the BINDTI performance. Under 5-fold cross validation, all DT pairs were randomly divided into roughly equal five parts, that is, in each round, 80% of DT pairs were randomly selected as the training set and the remaining 20% was taken as the test set.

### B. Baseline methods

The BINDTI performance was compared with the following four baselines: (1) CPI-GNN [48]: CPI-GNN utilizes GNN and CNN for encoding drug and protein features, respectively, and then concatenates the obtained two latent vectors into a neural network for compound-protein linkage inference; (2) TransformerCPI [69]: TransformerCPI takes drugs and proteins as two types of different sequences, and then separately generates drug atom representation and protein sequence representation through CNN and GCN, finally acquires DT features with decoder and calculates the interaction probability for each DT pair using linear layers; (3) MolTrans [61]: MolTrans first decomposes drugs and proteins into two sets with explicit substructure sequences through the frequent consecutive subsequence approach, and then acquires the augmented contextual embedding representations for drugs and proteins based on Transformer embedding modules, subsequently models the DTI map through dot product, and finally implements new DTI identification by incorporating CNN and fully connected network [52]; (4) IIFDTI [70]: IIFDTI first extracts the interactive features of drugs and proteins based on their substructures through the bi-directional encoder-decoder architecture, and then separately extracts their independent features through GAT and CNN, finally, all extracted features are fused and fed into a multi-layer fully connected network for inferring DTIs. The above baselines are classical deep

learning-based DTI prediction models and have manifested superior DTI classification performance over shallow models [61]. Their parameter settings were the same as ones provided by corresponding publications.

### C. Comparison of results

On the BindingDB, BioSNAP and DrugBank datasets, BINDTI was compared with four baseline methods under  $E_1$ : CPI-GNN, TransformerCPI, MolTrans, and IIFDTI. As shown in Table II, BINDTI computed the best AUC, AUPRC, accuracy, F1-score, and MCC compared to the above four baseline methods, followed by IIFDTI. In particular, on BindingDB, the AUC, AUPRC and MCC values for BINDTI were 2.53%, 3.62% and 8.92% better than MolTrans, respectively. GNN-CPI, TransformerCPI, MolTrans, and IIFDTI are end-to-end representation learning frameworks and significantly outperformed traditional DTI classification methods. BINDTI obtained the best performance, thereby elucidating its superior end-to-end learning ability.

The Human dataset is an imbalanced dataset. On the Human dataset, BINDTI was compared with the above four baseline methods under  $E_1$  and  $E_2$ . Under  $E_2$ , the ratio of positive DT pairs and negative DT pairs was set to 1:2. As shown in Table III, under  $E_1$ , BINDTI had the second-best performance. Under  $E_2$ , AUC, AUPRC, accuracy, F1-score, and MCC of the five DTI prediction methods greatly reduced, while BINDTI largely surpassed other four baselines. This suggests that BINDTI can infer potential interactions between unknown drugs and unknown proteins more precisely and effectively. In addition, it was more appropriate to deal with imbalanced datasets than other four methods.

Table IV shows the performance of BINDTI and all four baselines under 5-fold cross validation on the four DTI datasets. As shown in Table IV, under 5-fold cross validation, BINDTI obtained the best AUC, AUPRC, accuracy, F1-score, and MCC on the BindingDB, BioSNAP, and DrugBank datasets, significantly outperforming the four baselines. Although BINDTI computed slightly low performance on the Human dataset, the difference was very tiny. Thus, BINDTI could efficiently identify potential DTIs from unknown DT pairs.

In addition, we performed statistical tests on BINDTI and the four baselines under  $E_1$  and 5-fold cross validation. We carried out t-tests of the baseline models against BINDTI based on their AUC values, and the *difference* of average AUCs, *t*-value and *p*-value were calculated respectively. The larger the *difference* and *t*-value are, the better BINDTI outperforms the corresponding baseline method. The smaller *p*-value is, the greater the significant difference between BINDTI and the four baselines is. Generally, a *p*-value less than 0.05 indicates a significant difference. As shown in Table V, BINDTI significantly outperformed the four baselines on BindingDB, BioSNAP and DrugBank, and obtained competing performance on Human.

### D. Ablation experiments

In this section, we conducted multiple ablation experiments. First, as shown in Fig. 3, we evaluated the affect of different

**TABLE II:** Performance comparison on BindingDB, BioSNAP and DrugBank under  $E_1$ .

Method	AUC	AUPRC	Accuracy	F1-score	MCC
BindingDB					
CPI-GNN	0.6087±0.0089	0.5174±0.0147	0.5972±0.0080	0.3731±0.0567	0.1321±0.0120
TransformerCPI	0.8953±0.0023	0.8601±0.0030	0.8115±0.0037	0.7806±0.0036	0.6161±0.0067
MolTrans	0.9354±0.0033	0.9084±0.0044	0.8642±0.0080	0.8398±0.0095	0.7222±0.0163
IIFDTI	0.9242±0.0026	0.9014±0.0048	0.8329±0.0033	0.8179±0.0033	0.6743±0.0043
BINDTI	<b>0.9607±0.0019</b>	<b>0.9446±0.0024</b>	<b>0.9073±0.0026</b>	<b>0.9080±0.0018</b>	<b>0.8114±0.0045</b>
BioSNAP					
CPI-GNN	0.7071±0.0042	0.7237±0.0049	0.6656±0.0055	0.6457±0.0104	0.3349±0.0115
TransformerCPI	0.8436±0.0216	0.8590±0.0157	0.7709±0.0210	0.7668±0.0270	0.5431±0.0415
MolTrans	0.8796±0.0060	0.8829±0.0071	0.8005±0.0045	0.8018±0.0125	0.6034±0.0073
IIFDTI	0.8953±0.0038	0.8981±0.0056	0.7980±0.0037	0.8157±0.0020	0.6062±0.0050
BINDTI	<b>0.9027±0.0044</b>	<b>0.9020±0.0065</b>	<b>0.8333±0.0090</b>	<b>0.8355±0.0056</b>	<b>0.6675±0.0171</b>
DrugBank					
CPI-GNN	0.6951±0.0056	0.7344±0.0029	0.6599±0.0057	0.6097±0.0098	0.3331±0.0096
TransformerCPI	0.8557±0.0073	0.8584±0.0077	<u>0.7768±0.0078</u>	0.7801±0.0109	0.5540±0.0159
MolTrans	0.8600±0.0049	0.8661±0.0065	0.7726±0.0178	0.7790±0.0037	0.5497±0.0273
IIFDTI	0.8712±0.0025	0.8726±0.0029	0.7765±0.0081	0.7974±0.0035	0.5641±0.0128
BINDTI	<b>0.8819±0.0038</b>	<b>0.8830±0.0044</b>	<b>0.8093±0.0056</b>	<b>0.8132±0.0035</b>	<b>0.6195±0.0101</b>

Note: The best results are marked in bold and the second-best results are marked as underlined.

**TABLE III:** performance comparison on the Human dataset under  $E_1$  and  $E_2$ .

Method	AUC	AUPRC	Accuracy	F1-score	MCC
Setting $E_1$					
CPI-GNN	0.9582±0.0024	0.9554±0.0081	0.8946±0.0040	0.8909±0.0035	0.7893±0.0078
TransformerCPI	0.9654±0.0040	0.9599±0.0062	0.9120±0.0055	0.9030±0.0076	0.8228±0.0107
MolTrans	<b>0.9838±0.0044</b>	<b>0.9831±0.0044</b>	<b>0.9428±0.0126</b>	<u>0.9364±0.0127</u>	<b>0.8850±0.0246</b>
IIFDTI	0.9813±0.0031	0.9771±0.0039	0.9285±0.0125	<u>0.9216±0.0122</u>	0.8574±0.0236
BINDTI	<u>0.9813±0.0040</u>	0.9769±0.0045	0.9405±0.0087	<b>0.9385±0.0092</b>	0.8796±0.0175
Setting $E_2$					
CPI-GNN	0.7369±0.0343	0.7217±0.0668	0.6545±0.0518	0.5137±0.1105	0.3437±0.1134
TransformerCPI	0.7174±0.0457	0.6899±0.0587	0.6406±0.0530	0.3715±0.1840	0.2770±0.1135
MolTrans	0.7198±0.0276	0.6652±0.0284	0.6270±0.0354	0.3719±0.1079	0.2481±0.0639
IIFDTI	0.7232±0.0239	0.7163±0.0174	<b>0.6745±0.0197</b>	0.5054±0.0458	0.3541±0.0272
BINDTI	<b>0.7655±0.0402</b>	<b>0.7536±0.0330</b>	0.6424±0.0686	<b>0.7235±0.0269</b>	<b>0.3553±0.0990</b>

Note: The best performance and the second-best performance are highlighted in bold and underlined, respectively.

**TABLE IV:** Performance comparison on BindingDB, BioSNAP, DrugBank and Human under 5-fold cross-validation.

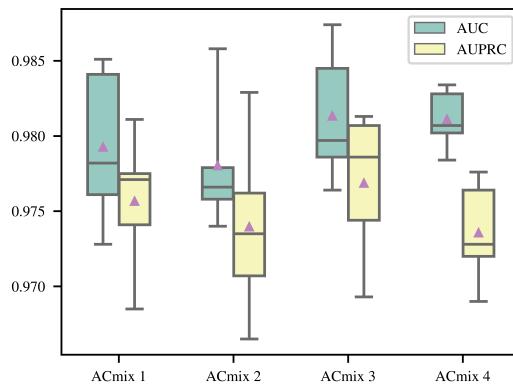
Method	AUC	AUPRC	Accuracy	F1-score	MCC
BindingDB					
CPI-GNN	0.6808±0.1498	0.6187±0.1756	0.6476±0.1299	0.5034±0.2150	0.2622±0.2746
TransformerCPI	0.9409±0.0033	<u>0.9248±0.0045</u>	0.8748±0.0066	0.8540±0.0081	0.7445±0.0133
MolTrans	0.9418±0.0049	0.9238±0.0102	0.8767±0.0111	0.8561±0.0147	0.7504±0.0222
IIFDTI	0.9272±0.0071	0.9057±0.0098	0.8363±0.0119	<u>0.8260±0.0112</u>	0.6835±0.0199
BINDTI	<b>0.9655±0.0036</b>	<b>0.9519±0.0045</b>	<b>0.9130±0.0085</b>	<b>0.9141±0.0061</b>	<b>0.8242±0.0162</b>
BioSNAP					
CPI-GNN	0.7109±0.0161	0.7295±0.0132	0.6730±0.0162	0.6522±0.0118	0.3506±0.0336
TransformerCPI	0.8966±0.0066	0.9030±0.0089	<u>0.8212±0.0074</u>	0.8235±0.0071	0.6430±0.0144
MolTrans	0.8895±0.0072	0.8949±0.0041	0.8105±0.0074	0.8114±0.0079	0.6239±0.0136
IIFDTI	0.8968±0.0082	0.8990±0.0132	0.7945±0.0110	0.8160±0.0082	0.6030±0.0163
BINDTI	<b>0.9046±0.0028</b>	<b>0.9061±0.0041</b>	<b>0.8341±0.0066</b>	<b>0.8373±0.0028</b>	<b>0.6694±0.0118</b>
DrugBank					
CPI-GNN	0.7244±0.0801	0.7674±0.0621	0.6780±0.0615	0.6504±0.0772	0.3665±0.1183
TransformerCPI	0.8718±0.0044	0.8773±0.0071	<u>0.7926±0.0070</u>	0.7991±0.0093	0.5856±0.0140
MolTrans	0.8709±0.0080	<u>0.8790±0.0092</u>	0.7892±0.0097	0.7983±0.0101	0.5803±0.0183
IIFDTI	0.8736±0.0030	0.8756±0.0065	0.7827±0.0121	<u>0.8057±0.0092</u>	0.5738±0.0183
BINDTI	<b>0.8881±0.0066</b>	<b>0.8884±0.0097</b>	<b>0.8223±0.0071</b>	<b>0.8249±0.0045</b>	<b>0.6451±0.0142</b>
Human					
CPI-GNN	0.9649±0.0069	0.9611±0.0081	0.8997±0.0180	0.8926±0.0237	0.8016±0.0341
TransformerCPI	<u>0.9820±0.0046</u>	0.9774±0.0046	0.9360±0.0084	0.9273±0.0083	0.8706±0.0162
MolTrans	<u>0.9809±0.0078</u>	<b>0.9799±0.0063</b>	<b>0.9500±0.0087</b>	0.9328±0.0173	<b>0.8989±0.0174</b>
IIFDTI	<b>0.9834±0.0036</b>	0.9798±0.0054	<u>0.9423±0.0079</u>	0.9353±0.0090	0.8840±0.0155
BINDTI	0.9794±0.0036	0.9753±0.0102	0.9407±0.0075	<b>0.9398±0.0055</b>	0.8800±0.0147

Note: The best performance and the second-best performance are highlighted in bold and underlined, respectively.

**TABLE V:** Statistical tests on BindingDB, BioSNAP, DrugBank and Human under  $E_1$  and 5-fold cross-validation.

Method	Setting $E_1$			5-fold cross-validation		
	<i>difference</i>	<i>t</i>	<i>p</i>	<i>difference</i>	<i>t</i>	<i>p</i>
BindingDB						
CPI-GNN	0.3520	77.5286	<0.001	0.2847	4.2480	<0.01
TransformerCPI	0.0654	43.6206	<0.001	0.0246	11.2497	<0.001
MolTrans	0.0253	13.2450	<0.001	0.0237	8.6735	<0.001
IIFDTI	0.0365	22.6115	<0.001	0.0383	10.7751	<0.001
BioSNAP						
CPI-GNN	0.1956	64.7855	<0.001	0.1937	26.4411	<0.001
TransformerCPI	0.0592	5.3759	<0.001	0.0080	2.4991	0.0370
MolTrans	0.0232	6.2478	<0.001	0.0151	4.3492	<0.01
IIFDTI	0.0074	2.5842	0.0324	0.0078	2.0308	0.0768
DrugBank						
CPI-GNN	0.1868	55.2741	<0.001	0.1637	4.5541	<0.001
TransformerCPI	0.0262	6.3739	<0.001	0.0163	4.6013	<0.01
MolTrans	0.0219	7.1005	<0.001	0.0172	3.7028	<0.01
IIFDTI	0.0108	4.7466	<0.01	0.0146	4.5204	<0.01
Human						
CPI-GNN	0.0231	9.9007	<0.001	0.0145	4.1605	<0.01
TransformerCPI	0.0159	5.5660	<0.001	-0.0026	-0.9944	0.3491
MolTrans	-0.0025	-0.8413	0.4246	-0.0015	-0.4007	0.6991
IIFDTI	0.0000	0.0033	0.9974	-0.0040	-1.7540	0.1175

ACmix layer numbers on the BINDTI performance on the Human dataset. The results demonstrated that the best performance was obtained when BINDTI used a 3-layer ACmix.



**Fig. 3:** Performance of BINDTI with different ACmix layer numbers on the Human dataset under  $E_1$ .

Next, we carried out ablation experiments to assess the effectiveness of bi-directional Intention network in the BINDTI framework on the BioSNAP dataset. We considered three variants of BINDTI: one-side drug Intention, one-side protein Intention, and their linear concatenation [46]. The two one-side Intentions are similar to the neural attention mechanism proposed by CPI-GNN [48], which can learn the joint representation corresponding to drug feature vector and protein feature matrix. We replaced bi-directional Intention in BINDTI using one-side attention to acquire the above two variants. The linear concatenation is a simple concatenation operation over drug and protein feature vectors after a max-pooling layer [46]. As shown in Table VI, bi-directional Intention (i.e., BINDTI) can more effectively fuse drug and protein features than linear concatenation and one-side drug or target attention.

Finally, GCN demonstrated the powerful drug feature rep-

**TABLE VI:** Ablation experiments about bi-directional mechanism on BioSNAP under  $E_1$ .

Method	AUC	AUPRC	MCC
Linear concatenation	<b><math>0.8979 \pm 0.0031</math></b>	$0.9007 \pm 0.0039$	$0.6567 \pm 0.0071$
One side protein Intention	$0.8842 \pm 0.0062$	$0.8824 \pm 0.0081$	$0.6236 \pm 0.0187$
One side drug Intention	$0.8934 \pm 0.0016$	$0.8976 \pm 0.0022$	$0.6482 \pm 0.0091$
Bi-directional Intention	<b><math>0.9027 \pm 0.0044</math></b>	<b><math>0.9020 \pm 0.0065</math></b>	<b><math>0.6675 \pm 0.0171</math></b>

Note: The best performance and the second-best performance are highlighted in bold and underlined, respectively.

resentation ability. To evaluate the effectiveness of GCN in the BINDTI framework, we replaced it with GAT and evaluated the performance of BINDTI under three different combinations, that is, GAT+ACmix, GAT+CNN, and GCN+ACmix (i.e., BINDTI), on the four datasets. As shown in Table VII, BINDTI computed the best AUC, AUPRC, and MCC when BINDTI combined GCN and ACmix. Therefore, we chose GCN as the drug feature encoder in our model.

### E. Visualization of the fused features

To investigate the distribution of the fused features and more intuitively demonstrate the superior performance of BINDTI, we visualized the fused features by BINDTI on the Human and BioSNAP datasets. We first randomly selected 1,200 DT pairs from the test set, and then reduced dimensions of the fused features with UMAP [71]. The results after dimension reduction were visualized in Fig. 4. In particular, each dataset were visualized through three types of data: (1) drug features which fused protein features through Intention and protein features which fused drug features through Intention; (2) joint features of DT pair with real labels; and (3) joint features of DT pair with the predicted labels. The results demonstrated that BINDTI can map the fused drug and protein features to different regions, thereby forming denser drug clusters and protein clusters. The two DT joint features were relatively

**TABLE VII:** Ablation experiments about GCN on the four datasets under  $E_1$ .

Method	AUC	AUPRC	MCC
BindingDB			
GAT+ACmix	<u>0.9575±0.0021</u>	<u>0.9407±0.0023</u>	0.7967±0.0074
GAT+CNN	<u>0.9576±0.0023</u>	<u>0.9404±0.0050</u>	<u>0.8028±0.0072</u>
GCN+ACmix	<b>0.9607±0.0019</b>	<b>0.9446±0.0024</b>	<b>0.8114±0.0045</b>
BioSNAP			
GAT+ACmix	0.8968±0.0026	0.8929±0.0034	<u>0.6658±0.0048</u>
GAT+CNN	<u>0.8993±0.0033</u>	<u>0.8966±0.0079</u>	0.6646±0.0052
GCN+ACmix	<b>0.9027±0.0044</b>	<b>0.9020±0.0065</b>	<b>0.6675±0.0171</b>
DrugBank			
GAT+ACmix	0.8725±0.0022	0.8698±0.0048	0.6026±0.0069
GAT+CNN	<u>0.8764±0.0016</u>	<u>0.8757±0.0041</u>	0.6158±0.0053
GCN+ACmix	<b>0.8819±0.0038</b>	<b>0.8830±0.0044</b>	<b>0.6195±0.0101</b>
Human			
GAT+ACmix	0.9765±0.0014	<u>0.9717±0.0042</u>	<u>0.8724±0.0060</u>
GAT+CNN	<u>0.9771±0.0030</u>	0.9700±0.0043	0.8694±0.0125
GCN+ACmix	<b>0.9813±0.0040</b>	<b>0.9769±0.0045</b>	<b>0.8796±0.0175</b>

Note: The best performance and the second-best performance are highlighted in bold and underlined, respectively.

sparserly distributed in the space. Furthermore, the visualized results on DT joint features with the predicted labels were basically consistent with those with true labels, demonstrating that it learned useful DT features for downstream prediction.

#### F. Case study

Although BINDTI manifested a good DTI classification ability on each benchmark dataset, it's not enough to illustrate whether it can effectively identify novel DTIs. To determine its credibility, we implemented two case studies from the perspectives of drugs and targets, respectively. First, BINDTI was trained on the DrugBank dataset, where all DTIs involving a query drug/target in the case study were removed. Subsequently, the interaction probabilities between the query drug/target and the other all targets/drugs were computed, and DT pairs were ranked based on their linkage probabilities. Finally, the predicted DTIs were validated through the DrugBank database [60].

We selected two drugs, NADH and Beta-D-Glucose, for conducting drug-based case studies. Oral NADH (DB00157) supplementation could decrease simple fatigue and energy-sapping disorders, for instance, chronic fatigue syndrome and fibromyalgia. NADH supplements could also strengthen concentration and memory capacity as well as enhance athletic endurance [60], [72], [73]. Beta-D-Glucose (DB02379) is a main energy source for living organisms [60]. The top 10 discovered targets interacting with the two drugs are shown in Table VIII. The results demonstrated that all top 10 predicted targets have been validated to have true linkages with the two drugs on the DrugBank database.

Similarly, we selected two targets, H1R and ADRA1C, for conducting target-based case studies. H1R (P35367) mediates neurotransmission in the central nervous system [60], [74]. ADRA1C (P35348) implements its effect through association with G proteins [75]. The top 10 predicted drugs interacting with the two targets are shown in Table IX. The results demonstrated that the other 9 predicted drugs except DB00943

**TABLE VIII:** The top 10 targets which were predicted to interact with NADH and Beta-D-Glucose by BINDTI.

Rank	NADH (DB00157)		Beta-D-Glucose (DB02379)	
	UniProt ID	Result	UniProt ID	Result
1	P48163	True	Q60053	True
2	P49821	True	P14090	True
3	Q16798	True	O07653	True
4	P11586	True	P26827	True
5	Q13630	True	Q9BW91	True
6	P07327	True	P17802	True
7	Q92781	True	Q54331	True
8	P31937	True	P23904	True
9	P17516	True	P05162	True
10	Q13423	True	Q701N4	True

have been verified to have true interactions with the two targets on the DrugBank database. The results further demonstrated the reliability of BINDTI in discovering DTIs for new drugs or targets.

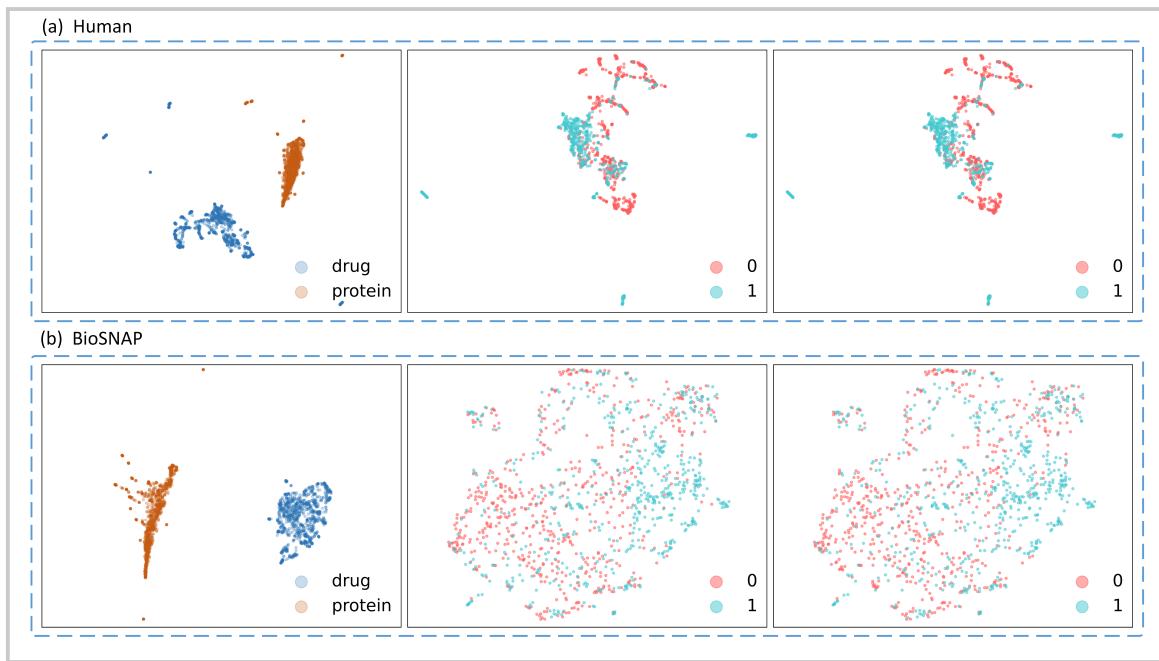
**TABLE IX:** The top 10 drug candidates which were predicted to interact with H1R and ADRA1C by BINDTI.

Rank	H1R (P35367)		ADRA1C (P35348)	
	DrugBank ID	Result	DrugBank ID	Result
1	DB06698	True	DB00943	Unconfirmed
2	DB04905	True	DB00540	True
3	DB01237	True	DB01151	True
4	DB09167	True	DB00661	True
5	DB04837	True	DB00458	True
6	DB00719	True	DB00211	True
7	DB00726	True	DB04948	True
8	DB00366	True	DB06207	True
9	DB00283	True	DB06207	True
10	DB04946	True	DB00575	True

## IV. DISCUSSION

DTI identification is very important to promote drug discovery. Current deep learning-based DTI prediction methods can be roughly divided into non-end-to-end methods and end-to-end methods. Non-end-to-end methods extracted feature descriptors of drugs and target proteins based on their chemical structures and biological information. However, end-to-end DTI prediction methods can extract high-quality features from raw biological data without human intervention. Thus, end-to-end-based deep learning framework supervised with biochemical data provides a promising insight into DTI prediction.

In this manuscript, we developed BINDTI, an end-to-end framework to identify new DTIs. BINDTI first used GCN to encode drug features based on its 2D molecular graph obtained by its SMILES string. Next, BINDTI utilized ACmix to encode protein features based on its amino acid sequence. Third, BINDTI fused drug and protein features through bi-directional Intention network and depicted one DT pair as a vector. Finally, BINDTI classified unknown DT pairs through multilayer perceptron based on the fused DT features.



**Fig. 4:** Visualization of the fused features. (a) Visualization on the Human dataset. (b) Visualization on the BioSNAP dataset. The three subfigures from left to right on the two datasets correspond to: (1) drug features which fused protein features through Intention and protein features which fused drug features through Intention; (2) joint features of DT pair with real labels; and (3) joint features of DT pair with the predicted labels. 0 represents DT pairs without interaction, and 1 represents DT pairs with interaction.

We conducted multiple experiments to assess the BINDTI performance. On the BindingDB, BioSNAP and DrugBank datasets, BINDTI greatly outperformed four baseline methods (CPI-GNN, TransformerCPI, MolTrans, and IIFDTI) under  $E_1$ . On the Human dataset, BINDTI computed good performance compared to the above four baseline methods under  $E_1$  and  $E_2$ . The experimental results under 5-fold cross validation further validated its good DTI classification performance and generalization ability. In addition, the results from ablation experiments demonstrated that bi-directional Intention, ACmix, and GCN can obviously advance DTI prediction performance. Finally, the fused feature visualization and case studies manifested that the predicted results by BINDTI were basically consistent with the true ones.

The proposed BINDTI framework obtained good DTI identification ability, it could attribute to the following features: (i) GCN can extract rich feature representation of drugs by comprehensively considering their chemical structural characteristics and neighboring information; (ii) ACmix, which perfectly aggregates CNN and self-attention mechanism, effectively extracted protein features; (iii) Bi-directional Intention network, which incorporates multi-head attention and Intention, significantly enhanced drug and protein's feature representation, better fused drug and target features, and boosted the model' interpretability and generalization ability.

In our proposed BINDTI framework, we used drug's 2D molecular graph to extract its features, which resulted in sparse features of many drug molecules. In the future, we will further improve the feature representation of drugs by combining their SMILES strings based on the concept of word embedding. Furthermore, BINDTI learned protein representations through their sequences and did not consider their 3D structures. Deep-

Mind's AlphaFold can generate abundant 3D structures for proteins and opened doors for protein 3D structure-based DTI prediction [76], [77]. Thus, we will also extend our ideas on 3D structure-based protein feature encoding to further improve DTI prediction performance and increase its interpretability.

Furthermore, in recent years, more and more non-coding RNAs (ncRNAs) have been found and increasing evidences have shown that they may affect gene expression and disease progression, making them a new class of drug targets. It thus becomes important to understand the relationship between ncRNAs and drug targets [78]. Consequently, Wang et al. [79] designed a robust PCA model for predicting potential associations between small molecules and microRNAs by combining  $\gamma$ -norm regularization. Liu et al. [80] proposed a multimodal graph neural network for inferring circRNA-drug resistance associations. Similar to the feature representation of target proteins, non-coding RNAs (such as microRNAs or circRNAs) can be characterized based on their nucleotide sequences, therefore, BINDTI may be suitable for predicting associations between small molecule drugs and non-coding RNAs. In the future, we will predict the relationship between small molecule drugs and non-coding RNAs by combining drug molecular structure, non-coding RNA sequences, and the bi-directional Intention network.

## V. CONCLUSION

In this manuscript, we developed an end-to-end DTI prediction framework BINDTI by integrating GCN, ACmix and bi-directional Intention network. The comparative experimental results demonstrated that BINDTI provided a powerful and robust tool for identifying potential DTIs. We anticipate that

our proposed BINDTI model can find new low-cost drug candidates, promote drugs' virtual screening, and further facilitate drug repositioning as well as drug discovery.

## REFERENCES

- [1] L. Peng, B. Liao, W. Zhu, Z. Li, and K. Li, "Predicting drug-target interactions with multi-information fusion," *IEEE journal of biomedical and health informatics*, vol. 21, no. 2, pp. 561–572, 2015.
- [2] X. Chen, C. C. Yan, X. Zhang, X. Zhang, F. Dai, J. Yin, and Y. Zhang, "Drug-target interaction prediction: databases, web servers and computational models," *Briefings in bioinformatics*, vol. 17, no. 4, pp. 696–712, 2016.
- [3] C.-C. Wang, Y. Zhao, and X. Chen, "Drug-pathway association prediction: from experimental results to computational models," *Briefings in bioinformatics*, vol. 22, no. 3, p. bbaa061, 2021.
- [4] J. Hu, W. Yu, C. Pang, J. Jin, N. T. Pham, B. Manavalan, and L. Wei, "Drugomerdti: Drug graphomer for drug-target interaction prediction," *Computers in Biology and Medicine*, vol. 161, p. 106946, 2023.
- [5] T.-H. Li, C.-C. Wang, L. Zhang, and X. Chen, "Snrmpacdc: computational model focused on siamese network and random matrix projection for anticancer synergistic drug combination prediction," *Briefings in Bioinformatics*, vol. 24, no. 1, p. bbac503, 2023.
- [6] H. Hu, Z. Feng, H. Lin, J. Cheng, J. Lyu, Y. Zhang, J. Zhao, F. Xu, T. Lin, Q. Zhao, et al., "Gene function and cell surface protein association analysis based on single-cell multiomics data," *Computers in Biology and Medicine*, vol. 157, p. 106733, 2023.
- [7] Y. Zhang, Y. Hu, N. Han, A. Yang, X. Liu, and H. Cai, "A survey of drug-target interaction and affinity prediction methods via graph neural networks," *Computers in Biology and Medicine*, p. 107136, 2023.
- [8] Q. Zhao, G. Duan, H. Zhao, K. Zheng, Y. Li, and J. Wang, "Gifdti: Prediction of drug-target interactions based on global molecular and intermolecular interaction representation learning," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 3, pp. 1943–1952, 2023.
- [9] L. Peng, F. Wang, Z. Wang, J. Tan, L. Huang, X. Tian, G. Liu, and L. Zhou, "Cell-cell communication inference and analysis in the tumour microenvironments from single-cell transcriptomics: data resources and computational strategies," *Briefings in Bioinformatics*, vol. 23, no. 4, p. bbac234, 2022.
- [10] S.-H. Wang, C.-C. Wang, L. Huang, L.-Y. Miao, and X. Chen, "Dual-network collaborative matrix factorization for predicting small molecule-mirna associations," *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab500, 2022.
- [11] Y.-S. Jung, Y. Kim, and Y.-R. Cho, "Comparative analysis of network-based approaches and machine learning algorithms for predicting drug-target interactions," *Methods*, vol. 198, pp. 19–31, 2022. Network-Based Approaches in Bioinformatics and Biomedicine.
- [12] S. Singh, Q. Bani Baker, and D. B. Singh, "Chapter 18 - molecular docking and molecular dynamics simulation," in *Bioinformatics* (D. B. Singh and R. K. Pathak, eds.), pp. 291–304, Academic Press, 2022.
- [13] Shaikh, Naeem and Sharma, Mahesh and Garg, Prabha, "An improved approach for predicting drug-target interaction: proteochemometrics to molecular docking," *Mol. BioSyst.*, vol. 12, pp. 1006–1014, 2016.
- [14] X. Yang, Y. Liu, J. Gan, Z.-X. Xiao, and Y. Cao, "FitDock: protein-ligand docking by template fitting," *Briefings in Bioinformatics*, vol. 23, p. bbac087, 03 2022.
- [15] M. Li, Z. Lu, Y. Wu, and Y. Li, "BACPI: a bi-directional attention neural network for compound-protein interaction and binding affinity prediction," *Bioinformatics*, vol. 38, pp. 1995–2002, 01 2022.
- [16] L. Huang, J. Lin, R. Liu, Z. Zheng, L. Meng, X. Chen, X. Li, and K.-C. Wong, "CoaDTI: multi-modal co-attention based framework for drug-target interaction annotation," *Briefings in Bioinformatics*, vol. 23, p. bbac446, 10 2022.
- [17] X. Chen, M.-X. Liu, and G.-Y. Yan, "Drug-target interaction prediction by random walk on the heterogeneous network," *Molecular BioSystems*, vol. 8, no. 7, pp. 1970–1978, 2012.
- [18] S. Fakhraei, B. Huang, L. Raschid, and L. Getoor, "Network-based drug-target interaction prediction with probabilistic soft logic," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 11, no. 5, pp. 775–787, 2014.
- [19] H. Wang, F. Huang, Z. Xiong, and W. Zhang, "A heterogeneous network-based method with attentive meta-path extraction for predicting drug-target interactions," *Briefings in Bioinformatics*, vol. 23, no. 4, p. bbac184, 2022.
- [20] S. Sadeghi, J. Lu, and A. Ngom, "A network-based drug repurposing method via non-negative matrix factorization," *Bioinformatics*, vol. 38, no. 5, pp. 1369–1377, 2022.
- [21] X. Su, P. Hu, H. Yi, Z. You, and L. Hu, "Predicting drug-target interactions over heterogeneous information network," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 562–572, 2022.
- [22] Z. Wu, H. Ma, Z. Liu, L. Zheng, Z. Yu, S. Cao, W. Fang, L. Wu, W. Li, G. Liu, et al., "wsdtmbi: a novel network-based inference method for virtual screening," *Chemical Science*, vol. 13, no. 4, pp. 1060–1079, 2022.
- [23] M. Bagherian, E. Sabeti, K. Wang, M. A. Sartor, Z. Nikolovska-Coleska, and K. Najarian, "Machine learning approaches and databases for prediction of drug-target interaction: a survey paper," *Briefings in bioinformatics*, vol. 22, no. 1, pp. 247–269, 2021.
- [24] X. Chen, Y. Gong, D.-H. Zhang, Z.-H. You, and Z.-W. Li, "Drmda: deep representations-based mirna-disease association prediction," *Journal of cellular and molecular medicine*, vol. 22, no. 1, pp. 472–485, 2018.
- [25] L. Peng, X. He, X. Peng, Z. Li, and L. Zhang, "Stgnnks: Identifying cell types in spatial transcriptomics data based on graph neural network, denoising auto-encoder, and k-sums clustering," *Computers in Biology and Medicine*, vol. 166, p. 107440, 2023.
- [26] L. Peng, W. Xiong, C. Han, Z. Li, and X. Chen, "Celldialog: A computational framework for ligand-receptor-mediated cell-cell communication analysis iii," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 2, 2024.
- [27] L. Peng, J. Tan, X. Tian, and L. Zhou, "Enanndeep: An ensemble-based lncrna-protein interaction prediction framework with adaptive k-nearest neighbor classifier and deep models," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 1, pp. 209–232, 2022.
- [28] S.-S. Guo, J. Liu, X.-G. Zhou, and G.-J. Zhang, "Deepumqa: ultrafast shape recognition-based protein model quality assessment using deep learning," *Bioinformatics*, vol. 38, no. 7, pp. 1895–1903, 2022.
- [29] C.-X. Peng, X.-G. Zhou, Y.-H. Xia, J. Liu, M.-H. Hou, and G.-J. Zhang, "Structural analogue-based protein structure domain assembly assisted by deep learning," *Bioinformatics*, vol. 38, no. 19, pp. 4513–4521, 2022.
- [30] J. Liu, K. Zhao, and G. Zhang, "Improved model quality assessment using sequence and structural information by enhanced deep neural networks," *Briefings in bioinformatics*, vol. 24, no. 1, p. bbac507, 2023.
- [31] X. Chen, T.-H. Li, Y. Zhao, C.-C. Wang, and C.-C. Zhu, "Deep-belief network for predicting potential mirna-disease associations," *Briefings in bioinformatics*, vol. 22, no. 3, p. bbaa186, 2021.
- [32] L. Peng, L. Huang, Q. Su, G. Tian, M. Chen, and G. Han, "LDA-VGHB: identifying potential lncRNADisease associations with singular value decomposition, variational graph auto-encoder and heterogeneous Newton boosting machine," *Briefings in Bioinformatics*, vol. 25, p. bbad466, 1 2024.
- [33] L. Zhang, X. Chen, and J. Yin, "Prediction of potential mirna-disease associations through a novel unsupervised deep learning framework with variational autoencoder," *Cells*, vol. 8, no. 9, p. 1040, 2019.
- [34] Y. Wang, G. Yu, J. Wang, G. Fu, M. Guo, and C. Domeniconi, "Weighted matrix factorization on multi-relational data for lncrna-disease association prediction," *Methods*, vol. 173, pp. 32–43, 2020.
- [35] F. Sun, J. Sun, and Q. Zhao, "A deep learning method for predicting metabolite-disease associations via graph neural network," *Briefings in Bioinformatics*, vol. 23, no. 4, p. bbac266, 2022.
- [36] W. Wang, L. Zhang, J. Sun, Q. Zhao, and J. Shuai, "Predicting the potential human lncrna-mirna interactions based on graph convolution network with conditional random field," *Briefings in Bioinformatics*, vol. 23, no. 6, p. bbac463, 2022.
- [37] G. Fu, J. Wang, C. Domeniconi, and G. Yu, "Matrix factorization-based data fusion for the prediction of lncrna-disease associations," *Bioinformatics*, vol. 34, no. 9, pp. 1529–1537, 2018.
- [38] L. Peng, M. Ren, L. Huang, and M. Chen, "Genddn: A deep learning framework for lncrna-disease association identification based on graph attention auto-encoder, dual-net neural architecture, and deep neural network," *Interdisciplinary Sciences: Computational Life Sciences*, 2024.
- [39] T. Wang, J. Sun, and Q. Zhao, "Investigating cardiotoxicity related with herg channel blockers using molecular fingerprints and graph attention mechanism," *Computers in Biology and Medicine*, vol. 153, p. 106464, 2023.
- [40] P. Zhang, Y. Wu, H. Zhou, B. Zhou, H. Zhang, and H. Wu, "Clnn-loop: a deep learning model to predict ctcf-mediated chromatin loops in the different cell lines and ctcf-binding sites (cbs) pair types," *Bioinformatics*, vol. 38, no. 19, pp. 4497–4504, 2022.
- [41] H. Wu, Y. Wu, Y. Jiang, B. Zhou, H. Zhou, Z. Chen, Y. Xiong, Q. Liu, and H. Zhang, "schicstackl: a stacking ensemble learning-based method

- for single-cell hi-c classification using cell embedding,” *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab396, 2022.
- [42] L. Peng, P. Gao, W. Xiong, Z. Li, and X. Chen, “Identifying potential ligandreceptor interactions based on gradient boosted neural network and interpretable boosting machine for intercellular communication analysis,” *Computers in Biology and Medicine*, p. 108110, 2024.
- [43] L. Wang, Z.-H. You, X. Chen, S.-X. Xia, F. Liu, X. Yan, Y. Zhou, and K.-J. Song, “A computational-based method for predicting drug–target interactions by using stacked autoencoder deep neural network,” *Journal of Computational Biology*, vol. 25, no. 3, pp. 361–373, 2018.
- [44] L. Zhang, C.-C. Wang, and X. Chen, “Predicting drug–target binding affinity through molecule representation block based on multi-head attention and skip connection,” *Briefings in Bioinformatics*, vol. 23, no. 6, p. bbac468, 2022.
- [45] Y. Wang, Y. Xia, J. Yan, Y. Yuan, H.-B. Shen, and X. Pan, “Zerobind: a protein-specific zero-shot predictor with subgraph matching for drug–target interactions,” *Nature Communications*, vol. 14, no. 1, p. 7861, 2023.
- [46] P. Bai, F. Miljković, B. John, and H. Lu, “Interpretable bilinear attention network with domain adaptation improves drug–target prediction,” *Nature Machine Intelligence*, vol. 5, pp. 126–136, Feb 2023.
- [47] M. Li, X. Cai, S. Xu, and H. Ji, “Metapath-aggregated heterogeneous graph neural network for drug–target interaction prediction,” *Briefings in Bioinformatics*, vol. 24, p. bbac578, 01 2023.
- [48] M. Tsubaki, K. Tomii, and J. Sese, “Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences,” *Bioinformatics*, vol. 35, pp. 309–318, 07 2018.
- [49] X. Lin, Z. Quan, Z.-J. Wang, Y. Guo, X. Zeng, and P. S. Yu, “Effectively identifying compound-protein interaction using graph neural representation,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, pp. 932–943, March 2023.
- [50] Y. Luo, Y. Liu, and J. Peng, “Calibrated geometric deep learning improves kinase–drug binding predictions,” *Nature Machine Intelligence*, vol. 5, no. 12, pp. 1390–1401, 2023.
- [51] I. Lee, J. Keum, and H. Nam, “Deepconv-dti: Prediction of drug–target interactions via deep learning with convolution on protein sequences,” *PLOS Computational Biology*, vol. 15, pp. 1–21, 06 2019.
- [52] Q. Zhao, H. Zhao, K. Zheng, and J. Wang, “HyperAttentionDTI: improving drug–protein interaction prediction by sequence-based deep learning with attention mechanism,” *Bioinformatics*, vol. 38, pp. 655–662, 10 2021.
- [53] S. Zheng, Y. Li, S. Chen, J. Xu, and Y. Yang, “Predicting drug–protein interaction using quasi-visual questionanswering system,” *Nature Machine Intelligence*, vol. 2, pp. 134–140, Feb 2020.
- [54] Q. Zhao, G. Duan, M. Yang, Z. Cheng, Y. Li, and J. Wang, “Attentiondta: Drug–target binding affinity prediction by sequence-based deep learning with attention mechanism,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, pp. 852–863, March 2023.
- [55] M. Zhao, M. Yuan, Y. Yang, and S. X. Xu, “Cpgl: Prediction of compound–protein interaction by integrating graph attention network with long short-term memory neural network,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, pp. 1935–1942, May 2023.
- [56] J. Bian, X. Zhang, X. Zhang, D. Xu, and G. Wang, “MCANet: shared-weight-based MultiheadCrossAttention network for drug–target interaction prediction,” *Briefings in Bioinformatics*, vol. 24, p. bbad082, 03 2023.
- [57] M. K. Gilson, T. Liu, M. Baitaluk, G. Nicola, L. Hwang, and J. Chong, “BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology,” *Nucleic Acids Research*, vol. 44, pp. D1045–D1053, 10 2015.
- [58] M. Zitnik, R. Sosić, S. Maheshwari, and J. Leskovec, “Biosnap datasets: Stanford biomedical network dataset collection. 2018,” URL <http://snap.stanford.edu/biodata>.
- [59] H. Liu, J. Sun, J. Guan, J. Zheng, and S. Zhou, “Improving compound–protein interaction prediction by building up highly credible negative samples,” *Bioinformatics*, vol. 31, pp. i221–i229, 06 2015.
- [60] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, and J. Woolsey, “DrugBank: a comprehensive resource for in silico drug discovery and exploration,” *Nucleic Acids Research*, vol. 34, pp. D668–D672, 01 2006.
- [61] K. Huang, C. Xiao, L. M. Glass, and J. Sun, “MolTrans: Molecular Interaction Transformer for drug–target interaction prediction,” *Bioinformatics*, vol. 37, pp. 830–836, 10 2020.
- [62] D. Weininger, “Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules,” *Journal of Chemical Information and Computer Sciences*, vol. 28, pp. 31–36, Feb 1988.
- [63] M. Li, J. Zhou, J. Hu, W. Fan, Y. Zhang, Y. Gu, and G. Karypis, “Dgl-lifesci: An open-source toolkit for deep learning on graphs in life science,” *ACS Omega*, vol. 6, no. 41, pp. 27233–27238, 2021.
- [64] X. Pan, C. Ge, R. Lu, S. Song, G. Chen, Z. Huang, and G. Huang, “On the integration of self-attention and convolution,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–825, June 2022.
- [65] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, Curran Associates, Inc., 2017.
- [66] M. Garnelo and W. M. Czarnecki, “Exploring the space of key-value-query models with intention,” 2023.
- [67] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [68] K. Bleakley and Y. Yamanishi, “Supervised prediction of drug–target interactions using bipartite local models,” *Bioinformatics*, vol. 25, no. 18, pp. 2397–2403, 2009.
- [69] L. Chen, X. Tan, D. Wang, F. Zhong, X. Liu, T. Yang, X. Luo, K. Chen, H. Jiang, and M. Zheng, “TransformerCPI: improving compound–protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments,” *Bioinformatics*, vol. 36, pp. 4406–4414, 05 2020.
- [70] Z. Cheng, Q. Zhao, Y. Li, and J. Wang, “IIFDTI: predicting drug–target interactions through interactive and independent features based on attention mechanism,” *Bioinformatics*, vol. 38, pp. 4153–4161, 07 2022.
- [71] L. McInnes, J. Healy, and J. Melville, “Umap: Uniform manifold approximation and projection for dimension reduction,” 2020.
- [72] P. Belenky, K. L. Bogan, and C. Brenner, “Nad+ metabolism in health and disease,” *Trends in biochemical sciences*, vol. 32, no. 1, pp. 12–19, 2007.
- [73] S.-J. Lin and L. Guarente, “Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease,” *Current opinion in cell biology*, vol. 15, no. 2, pp. 241–246, 2003.
- [74] M. Debacker, W. Gommeren, H. Moereels, G. Nobels, P. Vangompel, J. E. Leyten, and W. H. Luyten, “Genomic cloning, heterologous expression and pharmacological characterization of a human histamine h1 receptor,” *Biochemical and biophysical research communications*, vol. 197, no. 3, pp. 1601–1608, 1993.
- [75] C. D. Wright, S. C. Wu, E. F. Dahl, A. J. Sazama, and T. D. O’Connell, “Nuclear localization drives  $\alpha 1$ -adrenergic receptor oligomerization and signaling in cardiac myocytes,” *Cellular signalling*, vol. 24, no. 3, pp. 794–802, 2012.
- [76] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., “Highly accurate protein structure prediction with alphafold,” *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.
- [77] K. Zhao, Y. Xia, F. Zhang, X. Zhou, S. Z. Li, and G. Zhang, “Protein structure and folding pathway prediction based on remote homologs recognition using pathreader,” *Communications Biology*, vol. 6, no. 1, p. 243, 2023.
- [78] X. Chen, N.-N. Guan, Y.-Z. Sun, J.-Q. Li, and J. Qu, “Microna-small molecule association identification: from experimental results to computational models,” *Briefings in bioinformatics*, vol. 21, no. 1, pp. 47–61, 2020.
- [79] S. Wang, C. Ren, Y. Zhang, Y. Li, S. Pang, and T. Song, “Identifying potential small molecule–miRNA associations via robust pca based on  $\gamma$ -norm regularization,” *Briefings in Bioinformatics*, vol. 24, no. 5, p. bbad312, 2023.
- [80] Z. Liu, Q. Dai, X. Yu, X. Duan, and C. Wang, “Predicting circrna–drug resistance associations based on a multimodal graph representation learning framework,” *IEEE Journal of Biomedical and Health Informatics*, 2023.