JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

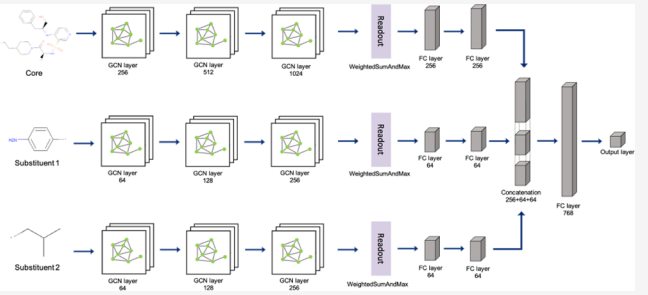
ACGCN: Graph Convolutional Networks for Activity Cliff Prediction between Matched Molecular Pairs

Junhui Park, Gaeun Sung, SeungHyun Lee, SeungHo Kang, and ChunKyun Park*

Cite This: J. Chem. Inf. Model. 2022, 62, 2341−2351 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: One of the interesting issues in drug−target interaction studies is the activity cliff (AC), which is usually defined as structurally similar compounds with large differences in activity toward a common target. The AC is of great interest in medicinal chemistry as it may provide clues to understanding the complex properties of the target proteins, paving the way for practical applications aimed at the discovery of more potent drugs. In this paper, we propose graph convolutional networks for the prediction of AC and designate the proposed models as Activity Cliff prediction using Graph Convolutional Networks (ACGCNs). The results show that ACGCNs outperform several off-the-shelf methods when predicting ACs of three popular target data sets for thrombin, Mu opioid receptor, and melanocortin receptor. Finally, we utilize gradient-weighted class activation mapping to visualize activation weights at nodes in the molecular graphs, demonstrating its potential to contribute to the ability to identify important substructures for molecular docking.



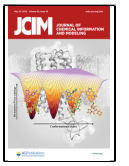
INTRODUCTION

In drug discovery, identifying drug−target interactions (DTIs) is an important initial task, in which the drugs are treated as chemical compounds and the targets are considered as proteins. Because an experimental assay is a time-consuming and expensive procedure by which to investigate DTIs, computer-aided DTI studies have become increasingly used.[1] The *in silico* methods have several significant parts that involve, for example, the molecular docking[2] and chemical similarity method (e.g., similarity ensemble approach,[3] SwissTargetPrediction[4]). Along with the pharmacophore-based method,[5,6] these are important in the drug design and chemical information and modeling community. Molecular docking studies, although done in a rigid body manner (especially for proteins) when simulating intermolecular interactions, provide reasonable accuracy and opportunities for visual interpretation. However, the method has some limitations in its use when it is difficult to obtain a three-dimensional structure of the target protein (e.g., a membrane protein) or when large-scale simulation is required. Chemical similarity methods play an important role in drug design studies by identifying similar compounds in large screening databases. The method is based on the "chemical similarity principle", which states that when two compounds are structurally similar, they tend to have similar properties.[7] Pharmacophore-based methods have received much attention, especially at a time when the structure of protein targets is not available.[8] A pharmacophore is defined by the International Union of Pure and Applied Chemistry[9] as "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response". A pharmacophore query can be done utilizing database molecules to find new putative hits or, conversely, to compare a ligand to several pharmacophore models (parallel screening) to look for potential targets.[10]

In this paper, we focus on the machine learning-based *in silico* methods for finding DTIs, which have been in the spotlight recently.[11] Machine learning, an alternative that has arisen over the past few decades, has become one of the most prominent approaches. It comes with an ever-increasing number of extensive biological data sets, including both chemical structure graphs and genomic sequences. It is an attractive method for accelerating the long candidate-testing process. In particular, the application of deep learning[12] in DTI studies has grown significantly.[13−17]

One issue of interest in DTI studies is the activity cliff (AC). ACs are generally defined as structurally similar compounds with large differences in activity toward a common target.[18] Because it is assumed that similar compounds exhibit similar potency in the structure−activity relationship (SAR) analysis, AC is an interesting outlier, often referred to as the SAR

Received: March 20, 2022
Published: May 6, 2022



https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

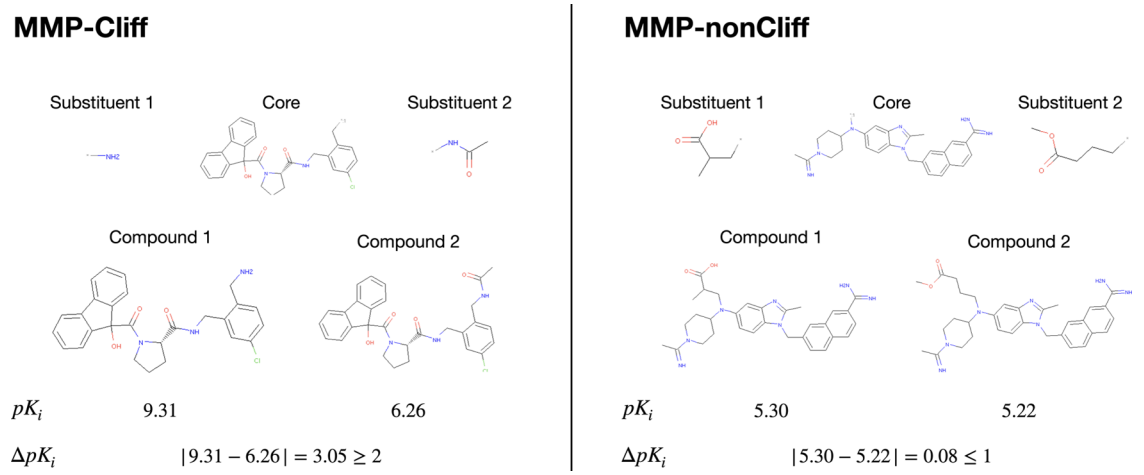


Figure 1. Example of MMP-Cliff and MMP-nonCliff in the MMP data set generated from compounds for thrombin as the target protein. One MMP consists of five elements: compound 1, compound 2, core, substituent 1, and substituent 2. If $\Delta pK_i \geq 2$, the MMP is labeled as MMP-Cliff (left), and if $\Delta pK_i \leq 1$, the MMP is labeled as MMP-nonCliff (right).

discontinuity.[18−20] Understanding ACs may provide clues to improve comprehension of the complex properties of target proteins, thereby paving the way for practical applications aimed at the discovery of newer, more potent drugs. In this context, ACs have been an active research area.[18,19,21−25] To predict an AC relationship between two similar compounds, a consistent definition for the two compounds must be established. A matched molecular pair (MMP) indicates a pair of compounds sharing a common core structure distinguished only by a single-site chemical modification. ACs, based on MMPs, are called MMP-Cliffs.[21]

There have been several strategies for using machine learning methods in AC prediction.[26−28] Most importantly, molecules can be represented graphically, and hence graph-based neural networks[29,30] can be utilized in SAR studies. Graph neural networks (GNNs) have been actively researched with significant achievements in graph representation learning.[31−34] Moreover, many studies have been done to improve GNN architectures for some well-known graph benchmarks, in particular, the expressive power of various graph neural network models.[35−37]

In this paper, we propose deep-learning models that predict the activity cliff relationship in MMPs based on a graph-based approach. Among a variety of GNNs, we use graph convolutional networks (GCNs) and designate the proposed models as Activity Cliff prediction using Graph Convolutional Networks (ACGCNs).

DATA MATERIALS

MMP Data Set. In general, deep-learning methods require a sufficient number of data to get good performance. For successful training of our model, we selected three targets (thrombin, Mu opioid receptor, and melanocortin receptor 4) with known binding affinities for a relatively large number of compounds. The generation of an MMP is conducted by means of exchanging a pair of distinguished substructures, such as terminal groups or central fragments. The algorithm proposed by Hussain and Rea[38] has often facilitated the performance of recent studies. The focus of the algorithm is the fragmenting and indexing of the compounds in the MMP analysis. Taking the single-cut example (i.e., MMPs that differ with respect to a single terminal fragment) to briefly illustrate the algorithm, in the first stage, the algorithm conducts fragmentation of all of the compounds in the input data set and enumerates all of the possible single cuts at the marked bonds in the compounds. It then performs indexing of the resulting fragments (see ref 38 for more details on the algorithm such as double-cut or triple-cut examples). A user-specified parameter can be applied to the algorithm to control the amount of difference between the two compounds in an MMP. For example, if the specified difference is 20%, then the MMP generated by the algorithm is a pair of molecules that differs at a single substructural site that is no larger than 20%. We empirically found the use of 20% as this parameter achieved a reasonable setting in our prediction study.

Based on pairwise potency differences, each MMP is labeled either MMP-Cliffs or MMP-nonCliffs. An MMP, a pair of ACs, is typically regarded as an MMP-cliff when the potency difference of the pair is at least 100-fold or 2 orders of magnitude (i.e., $\Delta pK_i \geq 2$). To avoid potency boundary effects on AC prediction, the potency difference of compounds forming an MMP-nonCliff is limited to a maximum of 10-fold or 1 order of magnitude (i.e., $\Delta pK_i \leq 1$). Other MMPs with potency differences between 1 and 2 orders of magnitude are not considered. All MMPs can be configured with five elements: compound 1, compound 2, common core of two compounds, substituent 1 except for a shared core in compound 1, and substituent 2 except for a shared core in compound 2. Figure 1 depicts examples of MMP-cliff and MMP-nonCliff.

To construct the data set, we obtained lists of compounds corresponding to each target from the ChEMBL database[39] (version 28) and found the MMPs by utilizing the Hussain and Rea algorithm.[38] The number of total compounds, total MMPs, MMP-Cliffs, and MMP-nonCliffs for the three targets is given in Table 1.

Expressing Molecular Structure in Graph Form. The graph consists of nodes and edges containing relational information between its nodes. We consider the atoms of a molecule as nodes and the bonds between atoms as edges. To graphically represent a molecule, the feature extraction is needed and an adjacency matrix should be defined for the connection information between atoms. Denoting some notations here will be allowed for a more specific explanation.

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

Table 1. Data Set Figures for Three Targets[a]

target name	total compounds	total MMPs	MMP-Cliffs	MMP-nonCliffs
thrombin	3171	5751	317	4408
Mu opioid receptor	3625	8725	219	7097
melanocortin receptor 4	1858	7169	111	5750

[a]Thrombin, Mu opioid receptor, melanocortin receptor 4.

A molecular graph $\mathcal{G}$ has the sets of atoms and bonds of these atoms denoted by $\mathcal{V}$ and $\mathcal{E}$, respectively. The ith node is denoted by $v_i \in \mathcal{V}$, and an edge is denoted by $e_{ij} = (v_i v_j) \in \varepsilon$ for $i,j \in [1,N_a]$. The adjacency matrix $A = (l_{ij}) \in \mathbb{R}^{N_a \times N_a}$ denotes the node connection information of a molecule.

The atom features are extracted, referring to the usage of Duvenaud et al.[29]

(1) Atom types: Br, C, Cl, F, H, I, N, O, P, S, B, and unknown.
(2) Atom degrees that indicate the number of bonded neighboring atoms.
(3) Implicit valence for atoms.
(4) Hybridization types.
(5) Whether an atom is part of an aromatic ring or not.

The atom features are one-hot-encoded to generate a 32-dimensional vector for each node, v_i.

An edge e_{ij} has its one-hot-encoded feature vector with a length of 6 from bond features. There are four bond types (single, double, triple, and aromatic); the other two bond features relate to whether a bond is conjugated and whether it is in a ring. A total of six adjacency matrices, $A^{(1)},\ldots,A^{(6)}$, are generated from each of six elements in the bond feature vector, $e_{ij} = \{e_{ij}^{(1)},\ldots,e_{ij}^{(6)}\}$, for all $i,j \in [1,N_a]$. When nodes v_i and v_j are connected, at least one element in their edge feature vector should be "1". Then, these six adjacency matrices are combined to produce the final adjacency matrix A. If at least one element of $l_{ij}^{(1)},\ldots,l_{ij}^{(6)}$ of the six adjacency matrices is 1, the element l_{ij} of the final version of the adjacency matrix A is 1, and otherwise, it is zero.

Therefore, a molecule is represented as the node feature matrix of size $N_a \times 32$ and an adjacency matrix of size $N_a \times N_a$, where N_a is the number of atoms in each molecule. The open-source cheminformatics software RDKit[40] was used to convert a molecule written in the simplified molecular-input line-entry system (SMILES)[41] notation to a graphic format.

PROPOSED MODELS

Graph Convolutional Layers for Graph Embedding. In the proposed models, the graph convolutional layer (GC layer) plays a role in embedding the molecular structure represented by the graph into a one-dimensional vector. For this, GC layers perform two operations: a graph convolution and a readout function. The graph convolution updates the node feature matrix through convolution operations of the atom and adjacent atoms. The readout function serves to reduce a two-dimensional node feature matrix to a one-dimensional node matrix.

Graph convolution is written as mathematical equations, e.g., for the feature vector of ith node v_i in the $(l+1)$th hidden layer

$$h_i^{(l+1)} = \sigma\left(b^{(l)} + \sum_{j \in \mathcal{N}(i)} \frac{1}{c_{ji}} h_j^{(l)} W^{(l)}\right) \quad (1)$$

where $b^{(l)}$ is the bias of the lth layer, $W^{(l)}$ is the specific weight matrix of the lth layer, $\mathcal{N}(i)$ is the set of neighbors of node v_i, the term c_{ji} is the product of the square root of node degrees (i.e., $c_{ji} = \sqrt{|\mathcal{N}(j)|}\sqrt{|\mathcal{N}(i)|}$), and σ is an activation function.[30]

First, each of the two input compounds passes through each corresponding GC layer for feature extraction. From these layers, the node feature matrix $h^{(k)} \in \mathbb{R}^{N_a \times N_f}$ is updated by eq 1, where N_f is the number of features from the last kth GC layer. Then, a one-dimensional vector $R_{out} \in \mathbb{R}^{1 \times 2N_f}$ is obtained from the updated feature $h^{(k)}$ by applying the readout function WeightedSumAndMax. This is a concatenation of $H_{\text{weightedsum}}$ and H_{maxpool}, as shown in the following equation, which is a function of a column-wise weighted sum and max pooling of a final updated feature, respectively

$$R_{\text{out}} = H_{\text{weightedsum}} \| H_{\text{maxpool}}$$

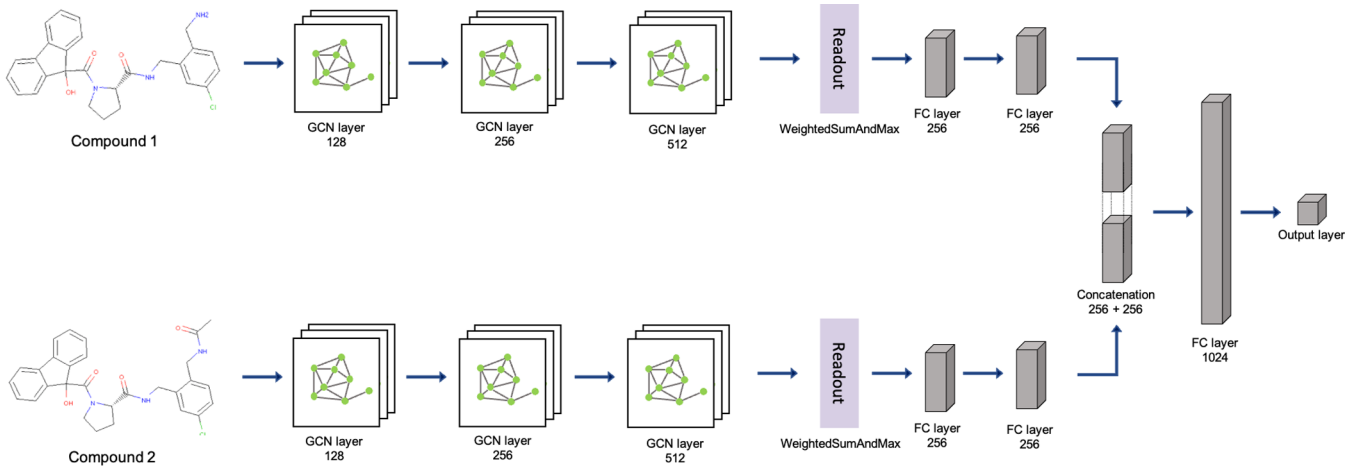


Figure 2. ACGCN-mmp architecture. This model uses two compounds as input, and each compound passes through three graph convolution layers and is expressed as a single vector through a readout function. Then, after combining the features through the two FC layers, the relationship of the MMP is predicted by one FC layer and the output layer.

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

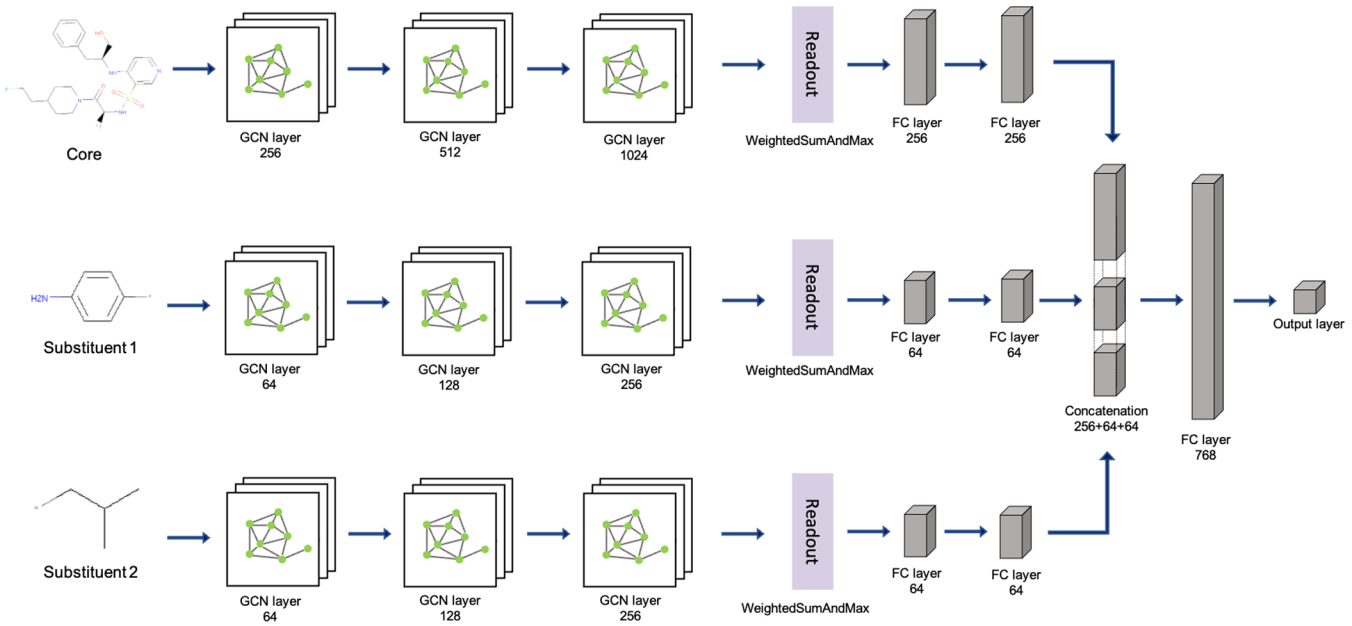


Figure 3. ACGCN-sub architecture. The input values for this model are core, substituent 1, and substituent 2. These inputs pass through three graph convolution layers and are expressed as vectors through readout functions. Then, after combining the features through the three FC layers, the relationship of the MMP is predicted by one FC layer and the output layer.

where $H_{\text{weightedsum}} = \text{sum}_{1\le i\le Na}(h_{ij}^{(R)})$ and $H_{\text{maxpool}} = \max_{1\le i\le Na}(h_{ij}^{(k)})$. Here, $h_i^{(R)} = b_R + h_i^{(k)}W_R$, where W_R and b_R are the weight matrix and the bias for computing the weights used in $H_{\text{weightedsum}}$, respectively.

Concatenation of Embedded Vectors. Because an MMP consists of two molecules, it would be appropriate to consider a model that could have multiple inputs. For this, we obtained a vector by passing each input to the GC layers and their own readout functions and, then, vectors generated from all inputs were concatenated into a single vector. For example, for two input compounds, two vectors could be obtained from two input compounds with $p_1 \times 1$- and $p_2 \times 1$-dimensions, respectively, by the GC layers and the readout functions. After these two vectors passed through n ($n \in \mathbb{N}_0$) fully connected (FC) layers, we integrated them into a one-dimensional vector by concatenation. Finally, the concatenated vector is passed through n FC layers and a sigmoid function to be classified as an MMP-Cliff or an MMP-nonCliff.

Overview of the Proposed Models. In this section, we introduce two specific models created by combining the embedding methods discussed above. The first model employs whole molecular structures in an MMP as an input, and we call this model the ACGCN-mmp. The second model, termed the ACGCN-sub, has its inputs generated from the common core of the MMP and the substituents of each molecule. Each input is converted from GC layers and readout function into a one-dimensional vector. After converting all inputs, they are concatenated into one single vector. The vector then passes through several FC layers and connects to the output layers to classify the molecule as an MMP-Cliff or an MMP-nonCliff.

The specific architecture of the ACGCN-mmp is as follows: the initial molecular features (with 32 dimensions) are successively updated to 128, 256, and 512 dimensions through each GCN block. Then, the molecule with the dimension of nodes × 512 is compressed by the readout function to a 256-dimensional single vector. The two vectors pass through two 256-dimensional FC layers, and they are concatenated into a single vector. After it passes through a 1024-dimensional FC layer, the final results are formed by an output layer. Figure 2 shows a graphic representation of the ACGCN-mmp structure.

In addition, the ACGCN-sub has an architecture similar to that of the ACGCN-mmp, except for the number of inputs: instead of two, three inputs are needed for the ACGCN-sub. The number of dimensions for GC layers is not necessarily large with a relatively simple molecular structure of the substituent. The model structures were explored through preliminary experiments. The optimal model with the best performance is presented in Figure 3.

EXPERIMENTS

Baseline Models. A deep neural network (DNN) with seven FC layers, light gradient boosting machine[42] (LightGBM), eXtreme gradient boosting[43] (XGBoost), and random forest[44] (RF) models were selected as the baseline models. An MMP should be transformed into a single vector to be applied as inputs for these models. Among the kernels introduced by Heikamp et al.[28] to generate a single-feature vector, we used two kernels for these experiments. Using fingerprint vectors of the common core and two substituents, the first kernel (named pair) generates the single vector by concatenating two vectors: one is calculated through an XOR computation of two substitute vectors and the other is obtained by comparing two core vectors. Furthermore, the second kernel, which is designated by the name diff, creates the vector simply by concatenating fingerprint vectors of the two substructures and the core. One may refer to Heikamp et al.[28] for a more detailed explanation of the kernels. In addition, Molecular ACCess System (MACCS) fingerprints[45] and the Extended Connectivity FingerPrints (ECFPs)[46] were applied to extract the MMP features.

Experimental Settings. In the MMP data set, a single strongly or weakly potent compound is likely to be included in several MMP-Cliffs.[47] Therefore, a naive splitting of the MMP data set into training and test sets would most likely result in

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

Table 2. Comparison of Model Performance for the Target Thrombin under the Item-Out Cross-Validation Setting[a]

			evaluation metrics					
model	feature	kernel	BA	TPR	TNR	F1-score	MCC	AUC
ACGCN	Duvenaud et al.[29]	mmp	0.858	0.878	0.839	0.463	0.446	0.899
		sub	**0.862**	**0.845**	0.879	**0.490**	**0.487**	**0.906**
DNN	MACCS	pair	0.830	0.758	**0.901**	0.487	0.477	0.885
		diff	0.794	0.732	0.857	0.395	0.383	0.881
	ECFP	pair	0.827	0.820	0.834	0.398	0.403	0.898
		diff	0.800	0.706	0.893	0.444	0.425	0.872
LightGBM	MACCS	pair	0.816	0.736	0.897	0.464	0.449	0.898
		diff	0.813	0.731	0.894	0.458	0.443	0.891
	ECFP	pair	0.817	0.736	0.899	0.470	0.454	0.892
		diff	0.790	0.717	0.863	0.397	0.382	0.865
XGBoost	MACCS	pair	0.827	0.794	0.861	0.429	0.426	0.902
		diff	0.822	0.773	0.872	0.440	0.433	0.897
	ECFP	pair	0.819	0.765	0.874	0.440	0.431	0.893
		diff	0.791	0.717	0.864	0.402	0.387	0.863
RF	MACCS	pair	0.839	0.835	0.843	0.422	0.427	0.905
		diff	0.835	0.805	0.865	0.440	0.439	0.899
	ECFP	pair	0.823	0.773	0.872	0.439	0.432	0.892
		diff	0.798	0.722	0.873	0.419	0.404	0.869

[a]ACGCN: Activity Cliff prediction using Graph Convolutional Networks, BA: balanced accuracy, DNN: deep neural networks, LightGBM: light gradient boosting machine, XGBoost: eXtreme gradient boosting, RF: random forests, BA: balanced accuracy, TPR: true positive rate, TNR: true negative rate, MACCS: Molecular ACCess System, and ECFP: Extended Connectivity FingerPrints.

many compounds simultaneously occurring in both sets. This can artificially increase the performance of machine learning methods, which might not have learnt any MMP-cliff characteristics but might only be able to recognize the weakly or strongly active compounds of an MMP. To address this problem, referring to the study of Horvath et al.,[26] we consider an advanced cross-validation (AXV) scheme in addition to conventional cross validation.

First, at each training stage, conventional p-fold "item-out" cross validation divides the MMP data set into two parts. The $(1 - p)/p$ part is for the training set, and the remaining $1/p$ part is for the test set. Throughout this paper, the value of p is set to "5".

Next, in the case of an AXV scheme of Horvath et al.,[26] a special concept called a "kept-out" pool is utilized to provide a partitioning criterion for the MMP data set. After generating a list of compounds in the MMP data set, $1/p$ of compounds is assigned to the kept-out pool. Depending on whether one or both of the compounds in an MMP belong to the kept-out pool, the MMP is assigned to either the "compound-out" or the "both-out" test set. If neither compound of an MMP is in the kept-out pool, then the MMP is included in the training set. With this advanced scheme, the compounds in all MMPs in the test and training sets do not overlap. Because the compounds in both test sets are not used to train the model, there is assurance that the model can properly distinguish the features of MMP-Cliffs and MMP-nonCliffs when the model achieves good performance.

Finally, model evaluation is repeated 10 times each in both conventional and advanced cross validation. For more detailed information on cross-validation methods including the AXV, please refer to Horvath et al.[26]

Model Settings. The hyperparameters for the learning models were a maximum of 1000 epochs and a minibatch size of 64. Moreover, an Adam optimizer with a learning rate of 0.001 was used. All hidden layers included a rectified linear unit (ReLU) activation function, a dropout rate of 0.2, and batch normalization. To prevent overfitting of model learning, the early stopping method was applied. Finally, we used the Deep Graph Library[48] for implementation of the graph convolutional layer and PyTorch[49] to construct the overall models.

Data Augmentation. When the input values of both ACGCN-mmp and ACGCN-sub pass the individual GC layers, each of the GC layers can be learned by only one molecule in the MMP. To solve this problem, we constructed the training set for ACGCN-mmp by swapping compound 1 and compound 2 of every MMP and adding swapped MMPs to the data set. Similarly, for ACGCN-sub, the data augmentation was implemented by swapping substituent 1 and substituent 2 of the MMPs in the training set.

RESULTS

Comparison of Performance between ACGCN and Baseline Models. We proposed ACGCN to predict an AC among the MMPs based on molecular graphs constructed from three separate targets. These included thrombin, Mu opioid receptor, and melanocortin receptor 4. In this paper, we compare the performance of different classifiers using several evaluation metrics such as true positive rate (TPR), true negative rate (TNR), balanced accuracy (BA), F1-score, Matthews correlation coefficient[50] (MCC), and area under the receiver operating characteristic curve (AUC). Details on the evaluation metrics are provided in the Supporting Information. When measuring the performance of a classifier, it is common to observe several metrics simultaneously rather than to rely on only one metric. For example, in terms of TNR, because the data are imbalanced due to a much larger number of MMP-nonCliffs, as shown in Table 1, a classifier favoring such MMP-nonCliffs could be considered high-performing. As such, having a high TNR alone does not mean that the classifier achieves high performance, so it is reasonable to consider several metrics simultaneously. Therefore, we decided to designate a classifier with the highest score on the most

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

Table 3. Comparison of Model Performance for the Target Mu Opioid Receptor under the Item-Out Cross-Validation Setting[a]

			evaluation metrics					
model	feature	kernel	BA	TPR	TNR	F1-score	MCC	AUC
ACGCN	Duvenaud et al.[29]	mmp	**0.889**	**0.907**	0.871	0.334	0.373	**0.929**
		sub	0.887	0.877	0.896	0.341	**0.400**	0.927
DNN	MACCS	pair	0.806	0.668	**0.945**	**0.386**	0.385	0.882
		diff	0.796	0.667	0.926	0.376	0.396	0.867
	ECFP	pair	0.820	0.711	0.929	0.262	0.315	0.891
		diff	0.811	0.693	0.929	0.258	0.307	0.888
LightGBM	MACCS	pair	0.836	0.763	0.910	0.325	0.366	0.919
		diff	0.839	0.803	0.874	0.274	0.327	0.909
	ECFP	pair	0.844	0.822	0.867	0.269	0.327	0.911
		diff	0.825	0.817	0.833	0.226	0.285	0.910
XGBoost	MACCS	pair	0.847	0.794	0.899	0.318	0.365	0.922
		diff	0.832	0.797	0.867	0.266	0.318	0.915
	ECFP	pair	0.838	0.764	0.912	0.332	0.372	0.912
		diff	0.810	0.752	0.867	0.249	0.295	0.888
RF	MACCS	pair	0.859	0.850	0.868	0.288	0.347	0.911
		diff	0.851	0.816	0.886	0.304	0.356	0.902
	ECFP	pair	0.837	0.783	0.892	0.313	0.357	0.897
		diff	0.813	0.766	0.860	0.251	0.298	0.871

[a]The abbreviations are the same as those in Table 2.

Table 4. Comparison of Model Performance for the Target Melanocortin Receptor 4 under the Item-Out Cross-Validation Setting[a]

			evaluation metrics					
model	feature	kernel	BA	TPR	TNR	F1-score	MCC	AUC
ACGCN	Duvenaud et al.[29]	mmp	**0.899**	**0.916**	0.882	**0.365**	**0.451**	**0.936**
		sub	0.896	0.874	0.919	0.354	0.385	0.922
DNN	MACCS	pair	0.838	0.803	0.873	0.272	0.326	0.913
		diff	0.844	0.763	0.925	0.363	0.398	0.917
	ECFP	pair	0.858	0.767	0.948	0.343	0.394	0.935
		diff	0.829	0.730	0.928	0.269	0.323	0.873
LightGBM	MACCS	pair	0.858	0.765	0.951	0.358	0.405	0.927
		diff	0.875	0.802	0.948	0.355	0.410	0.934
	ECFP	pair	0.846	0.593	0.943	0.253	0.287	0.828
		diff	0.848	0.775	0.922	0.268	0.331	0.917
XGBoost	MACCS	pair	0.851	0.739	**0.964**	0.352	0.444	0.920
		diff	0.873	0.793	0.953	0.358	0.427	0.915
	ECFP	pair	0.828	0.719	0.937	0.294	0.343	0.881
		diff	0.826	0.715	0.936	0.290	0.338	0.890
RF	MACCS	pair	0.877	0.829	0.926	0.304	0.371	0.915
		diff	0.871	0.805	0.936	0.320	0.381	0.909
	ECFP	pair	0.867	0.806	0.928	0.308	0.370	0.918
		diff	0.843	0.767	0.920	0.274	0.333	0.888

[a]The abbreviations are the same as those in Table 2.

metrics as the best-performing classifier for a comprehensive comparison.

First, in Table 2, we compared the proposed ACGCN and baseline models (DNN, LightGBM, XGBoost, and RF) for the target thrombin with the item-out cross-validation method. Thrombin, as is well known, is an enzyme that plays an important role in the blood coagulation process.[51] Blood clotting issues could lead to life-threatening conditions, such as bleeding, bruising, or thrombosis. As shown in Table 2, for thrombin, ACGCN-sub produced the best overall results compared to the other models. Moreover, ACGCN-sub showed the highest BA value of 0.862, which is frequently used when analyzing imbalanced data.

The second target for AC prediction between MMPs is the Mu opioid receptor, which mediates analgesic and brain reward pathways. Understanding the Mu opioid receptor is important because it is central to the development of addiction therapies.[52] Table 3 shows that for the Mu opioid receptor, ACGCN-mmp outperformed other models under the item-out cross-validation setting. Moreover, the BA value of ACGCN-mmp was the highest at 0.889.

The last target for AC prediction was melanocortin receptor 4, which is a crucial regulator of energy homeostasis. Because it regulates both appetite and energy expenditure in the hypothalamus, it is one of the most important receptors for the treatment of obesity.[53] The results for the target

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

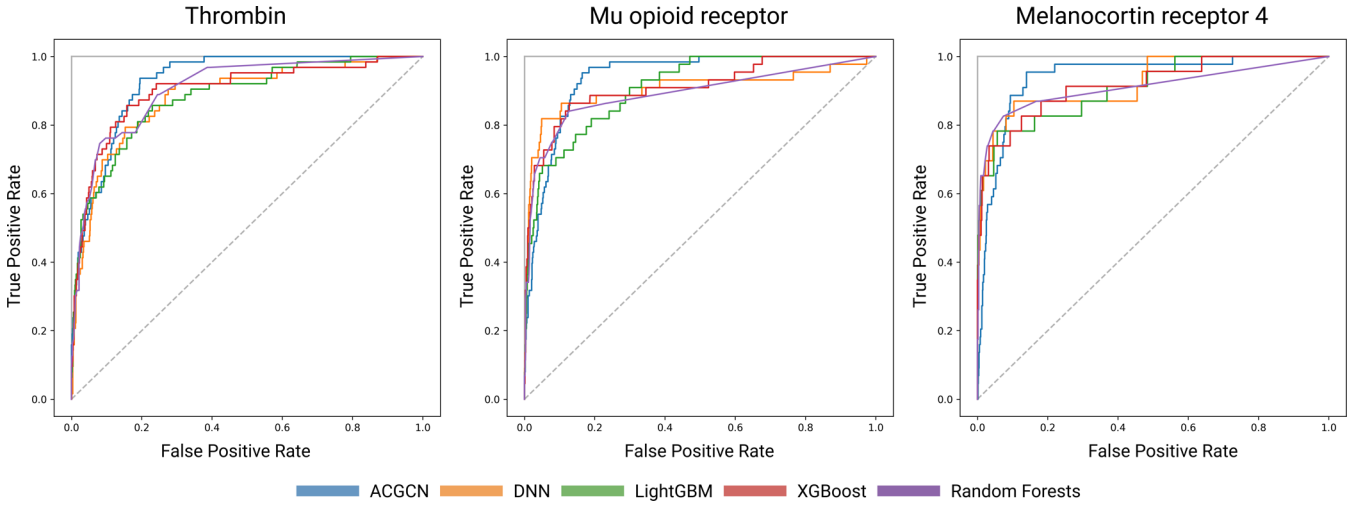


Figure 4. ROC curves for thrombin (left), Mu opioid receptor (center), and melanocortin receptor 4 (right) targets under the item-out cross-validation setting. At each target, the model's results came from the feature extraction method and kernel combination with the highest AUC values.

Table 5. Comparison of Model Performance for the Three Targets under the Advanced Cross-Validation Setting[a]

			thrombin		Mu opioid receptor		melanocortin receptor 4	
model	feature	kernel	compound-out	both-out	compound-out	both-out	compound-out	both-out
ACGCN	Duvenaud et al.[29]	mmp	0.785	**0.767**	**0.842**	**0.779**	0.761	0.711
		sub	**0.798**	0.751	0.814	0.755	**0.796**	**0.727**
DNN	MACCS	pair	0.724	0.684	0.745	0.728	0.772	0.710
		diff	0.733	0.704	0.711	0.687	0.758	0.665
	ECFP	pair	0.728	0.724	0.712	0.648	0.746	0.717
		diff	0.728	0.691	0.759	0.669	0.768	0.632
LightGBM	MACCS	pair	0.750	0.676	0.791	0.706	0.766	0.645
		diff	0.765	0.709	0.786	0.776	0.768	0.669
	ECFP	pair	0.748	0.685	0.757	0.707	0.680	0.621
		diff	0.728	0.657	0.750	0.756	0.736	0.645
XGBoost	MACCS	pair	0.763	0.745	0.787	0.762	0.744	0.638
		diff	0.770	0.729	0.774	0.734	0.758	0.671
	ECFP	pair	0.760	0.678	0.777	0.700	0.734	0.637
		diff	0.736	0.651	0.758	0.712	0.735	0.702
RF	MACCS	pair	0.774	0.726	0.800	0.748	0.771	0.660
		diff	0.779	0.732	0.788	0.752	0.764	0.664
	ECFP	pair	0.754	0.705	0.779	0.744	0.750	0.691
		diff	0.744	0.676	0.761	0.747	0.734	0.687

[a]Evaluation metric: balanced accuracy. The abbreviations are the same as those in Table 2.

melanocortin receptor 4 are shown in Table 4 in the case of the item-out cross-validation. ACGCN-mmp showed the best performance compared to other models with the highest BA value of 0.899.

We visualized the results for these three targets through the receiver operating characteristic (ROC) curves in Figure 4 under the item-out cross-validation setting. When comparing classifiers, we considered the combination of feature extraction methods and kernels with the highest AUC values in each model. For example, in the leftmost figure for thrombin, ACGCN represents the results for ACGCN-sub, DNN denotes the results for DNN-ECFP-pair, LightGBM indicates the results for LightGBM-MACCS-pair, and so on. Likewise, when the results were compared between models for the remaining targets (the Mu opioid receptor and melanocortin receptor 4), the feature extraction method and kernel combination with the highest AUC were considered. From the results in Figure 4, it was confirmed that the ACGCN model generally provided the best performance, as shown in Tables 2−4.

As mentioned, using only the conventional item-out cross-validation would most likely result in artificially increasing the performance of the machine learning methods that may not have learnt any MMP-cliff characteristics but can only recognize the weakly or highly active compound of an MMP. Thus, we additionally consider the AXV scheme of Horvath et al.[26] to address this problem. Similar to the case in Tables 2−4, we compared the model performance for the three targets in the AXV scheme. As expected, in the AXV scheme, the overall performance of all classifiers was lower than in the item-out cross-validation setting, with the both-out cross-validation being the lowest. Because aspects of the relative performance differences between classifiers are similar to those in the case of item-out cross-validation setting, only BA values are reported in Table 5 as representative due to space constraints. The

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

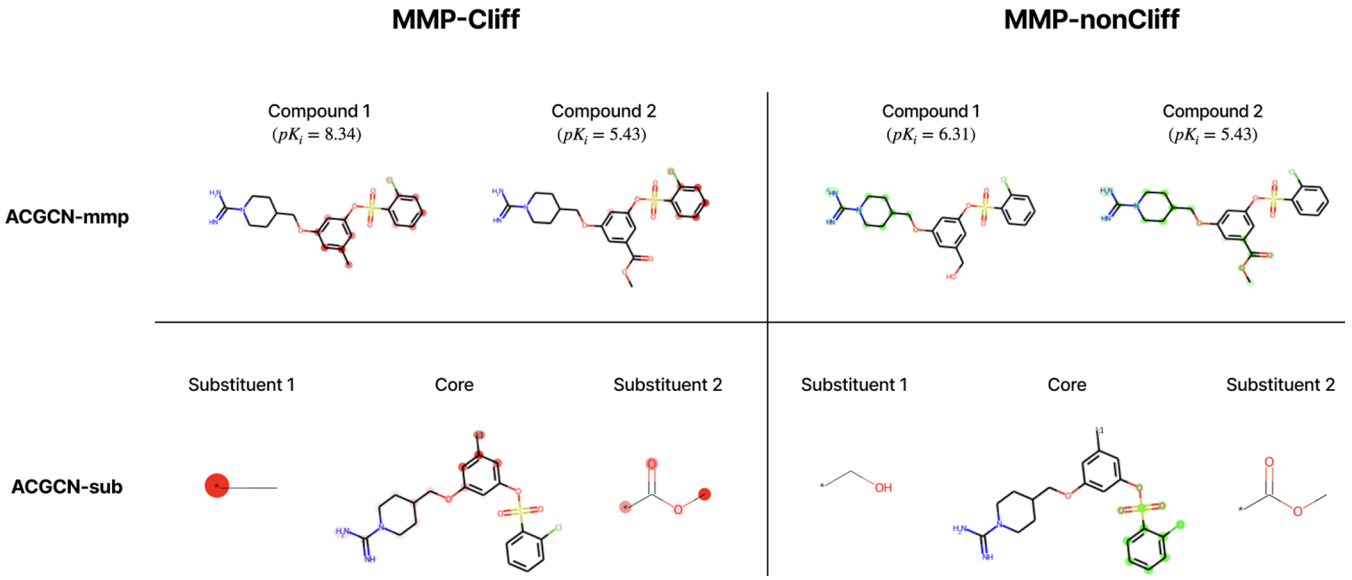


Figure 5. Visualization of activation weights using Grad-CAM. The mean gradient weights of each node are displayed in the molecular graph for two exemplary MMPs from thrombin. Red and green colors are for MMP-Cliff and MMP-nonCliff, respectively, and the thicker colors indicate higher values.

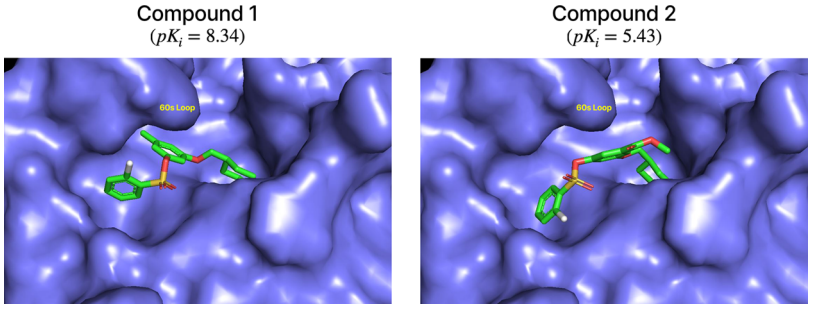


Figure 6. Exemplar contrast between two compounds in the case of MMP-Cliff.

classifier with the highest BA value also had the highest value for most of the remaining evaluation metrics. From the results in Table 5, we found that the ACGCN model generally performed superior or at least comparable to other baseline models.

Interpretability Using Grad-CAM. Gradient-weighted class activation mapping[54] (Grad-CAM) is a type of explainable artificial intelligence (XAI) method that can provide information concerning which node is accountable for the results (i.e., MMP-Cliff or MMP-nonCliff) by visualizing activation weights. We utilize the explanatory power of Grad-CAM to understand which node deserves to be noticed. Figure 5 depicts an exemplar image of two MMPs for the target thrombin. Here, the superimposed color shows the activation weights for the corresponding nodes. The red and green colors indicate the mean activation weights for MMP-Cliff and MMP-nonCliff, respectively, and the thicker colors indicate higher values. Note that, in Figure 5, in the case of the MMP-Cliff, the value of the activation weight increased at the substituent site of compound 1 but not in the case of compound 2.

To visualize how the changed substituent interacts with the target, we performed a molecular docking study with reference to the binding mode of thrombin and compound 1 (in the MMP-Cliff case in Figure 5), which was identified through X-ray structural analysis.[55] From the online Protein Data Bank (PDB) database, a three-dimensional structure of human thrombin (PDB ID: 1D3D) was obtained,[56] and molecular docking was performed with AutoDock Vina (version 1.2)[57] via AMDock (version 1.5.2).[58] The docking results were visualized using PyMOL (version 2.5).[59] Figure 6 shows that the bulkiness of the substituent for compound 2 (in the MMP-Cliff case in Figure 5) interferes near the 60s loop of the target thrombin, resulting in suboptimal binding. We believe that this demonstrates the potential of Grad-CAM, that is, the ability to visualize the activation weights on graphs may provide clues that will be of great help in understanding the actual mechanism of molecular docking. Generalizing the activation weight patterns with respect to the observed cliffs will be the subject of an interesting future study.

CONCLUSIONS

In this paper, we presented ACGCNs for the prediction of ACs from MMPs and compared their performance with other baseline models. As far as we know, these are the first comparisons of AC predictions using models based on molecular graphs at the compound pair level. Our method achieves reasonably satisfactory performance in predicting ACs in most evaluation metrics scores, including the BA. Moreover, we presented Grad-CAM approaches to visualize activation weights, which could be used for understanding which node deserves to be noticed. These improvements could contribute to the ability to identify important substructures for molecular

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351

docking. In future works, we would like to explore the following:

1. We propose using a modern graph convolutional network to explore a powerful approach for predicting activity cliffs between MMPs based on the given values of pK_i differences. We plan to construct a model that directly predicts the pK_i difference of two compounds with an MMP relationship,[27] thereby yielding a more integrated and comprehensive method along these lines.
2. Recently, there has been some progress in the related graph network approaches such as the message passing neural networks[60] framework and graph attention networks.[61] In particular, the method of Pattanaik et al.[60] deals with tetrahedral chirality, that is, a type of stereoisomerism, and this is a significant advance given the weakness of graph-based methodologies in which stereoisomer discrimination is inherently impossible.[60] Further study could be valuable for making use of such advances to predict activity cliffs between MMPs.
3. Finally, modeling studies for individual targets may naturally suffer from a lack of predictability for untrained target proteins. Future studies could improve on this by enlarging the scope of input so that it could better incorporate target protein sequencing information. We expect that the resulting approach could predict the activity cliff for general target proteins with a single model.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.2c00327.

Detailed description is provided in the Supporting Information regarding the metrics used to evaluate the model performance (PDF)

■ AUTHOR INFORMATION

Corresponding Author

ChunKyun Park − *Department of Statistics and Data Science, Yonsei University, Seoul 03722, South Korea;* orcid.org/0000-0002-3280-9416; Phone: +82 010-3583-1028; Email: chunkyun@yonsei.ac.kr

Authors

Junhui Park − *Department of Statistics and Data Science, Yonsei University, Seoul 03722, South Korea*

Gaeun Sung − *TESSER Inc., Seoul 06147, South Korea;* orcid.org/0000-0002-2887-5816

SeungHyun Lee − *TESSER Inc., Seoul 06147, South Korea;* orcid.org/0000-0002-5358-325X

SeungHo Kang − *Department of Statistics and Data Science, Yonsei University, Seoul 03722, South Korea;* orcid.org/0000-0002-7593-7899

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.2c00327

Author Contributions

Conceptualizing the study: J.P. and C.P.; implementing the experiments: J.P., G.S., and S.L.; and writing the paper: all authors. All of the authors read and approved the final manuscript.

Notes

The authors declare no competing financial interest.

The data and software that support the findings of this study are available at https://github.com/chunkyun/ACGCN.

■ ACKNOWLEDGMENTS

S.K. and J.P. were supported by the National Research Foundation (NRF), Korea, under project BK21 FOUR (2021-11-0077). S.K. was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2020R1F1A1A01048240).

■ ABBREVIATIONS

AC, activity cliff; ACGCN, activity cliff prediction using graph convolutional network; AUC, area under the ROC curve; AXV, advanced cross validation; DNN, deep neural network; DTI, drug−target interaction; ECFPs, Extended Connectivity FingerPrints; FC layer, fully connected layer; FPR, false positive rate; MCC, Matthews correlation coefficient; MMP, matched molecular pair; GC layer, graph convolutional layer; GCN, graph convolutional network; GNNs, graph neural networks; Grad-CAM, gradient-weighted class activation mapping; Light-GBM, light gradient boosting machine; MACCS, Molecular ACCess System; ReLU, rectified linear unit; RFs, random forests; ROC, receiver operating characteristic; SAR, structure−activity relationship; TNR, true negative rate; TPR, true positive rate; XAI, explainable artificial intelligence; XGBoost, eXtreme gradient boosting

■ REFERENCES

(1) Sachdev, K.; Gupta, M. K. A comprehensive review of feature based methods for drug target interaction prediction. *J. Biomed. Inf.* **2019**, *93*, No. 103159.

(2) Tao, X.; Huang, Y.; Wang, C.; Chen, F.; Yang, L.; Ling, L.; Che, Z.; Chen, X. Recent developments in molecular docking technology applied in food science: a review. *Int. J. Food Sci. Technol.* **2020**, *55*, 33−45.

(3) Keiser, M. J.; Roth, B. L.; Armbruster, B. N.; Ernsberger, P.; Irwin, J. J.; Shoichet, B. K. Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* **2007**, *25*, 197−206.

(4) Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* **2019**, *47*, W357−W364.

(5) Horvath, D. Pharmacophore-Based Virtual Screening. In *Chemoinformatics and Computational Chemical Biology*; Springer, 2011; pp 261−298.

(6) Schaller, D.; Šribar, D.; Noonan, T.; Deng, L.; Nguyen, T. N.; Pach, S.; Machalz, D.; Bermudez, M.; Wolber, G. Next generation 3D pharmacophore modeling. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, No. e1468.

(7) Johnson, M. A.; Maggiora, G. M. *Concepts and Applications of Molecular Similarity*; Wiley, 1990.

(8) Guner, O. History and evolution of the pharmacophore concept in computer-aided drug design. *Curr. Top. Med. Chem* **2002**, *2*, 1321−1332.

(9) Wermuth, C. G.; Ganellin, C.; Lindberg, P.; Mitscher, L. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* **1998**, *70*, 1129−1143.

(10) Steindl, T. M.; Schuster, D.; Laggner, C.; Langer, T. Parallel screening: a novel concept in pharmacophore modeling and virtual screening. *J. Chem. Inf. Model.* **2006**, *46*, 2146−2157.

(11) Bagherian, M.; Sabeti, E.; Wang, K.; Sartor, M. A.; Nikolovska-Coleska, Z.; Najarian, K. Machine learning approaches and databases for prediction of drug-target interaction: a survey paper. *Briefings Bioinf.* **2020**, *22*, 247−269.

(12) Goh, G. B.; Hodas, N. O.; Vishnu, A. Deep learning for computational chemistry. *J. Comput. Chem.* **2017**, *38*, 1291−1307.
(13) Jiang, D.; Wu, Z.; Hsieh, C.-Y.; Chen, G.; Liao, B.; Wang, Z.; Shen, C.; Cao, D.; Wu, J.; Hou, T. Could graph neural networks learn better molecular representation for drug discovery? A comparison study of descriptor-based and graph-based models. *J. Cheminf.* **2021**, *13*, 1−23.
(14) Wu, Y.; Gao, M.; Zeng, M.; Chen, F.; Li, M.; Zhang, J. BridgeDPI: A Novel Graph Neural Network for Predicting Drug-Protein Interactions. 2021, arXiv:2101.12547. arXiv.org e-Print archive. https://arxiv.org/abs/2101.12547.
(15) Rifaioglu, A. S.; Nalbat, E.; Atalay, V.; Martin, M. J.; Cetin-Atalay, R.; Doğan, T. DEEPScreen: high performance drug-target interaction prediction with convolutional neural networks using 2-D structural compound representations. *Chem. Sci.* **2020**, *11*, 2531−2557.
(16) Gao, K. Y.; Fokoue, A.; Luo, H.; Iyengar, A.; Dey, S.; Zhang, P. In *Interpretable Drug Target Prediction Using Deep Neural Representation*, Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI), 2018; pp 3371−3377.
(17) Jiang, M.; Li, Z.; Zhang, S.; Wang, S.; Wang, X.; Yuan, Q.; Wei, Z. Drug-target affinity prediction using graph neural network and contact maps. *RSC Adv.* **2020**, *10*, 20701−20712.
(18) Stumpfe, D.; Hu, H.; Bajorath, J. Evolving concept of activity cliffs. *ACS Omega* **2019**, *4*, 14360−14368.
(19) Maggiora, G. M. On outliers and activity cliffs why QSAR often disappoints. *J. Chem. Inf. Model.* **2006**, *46*, No. 1535.
(20) Iyer, P.; Stumpfe, D.; Vogt, M.; Bajorath, J.; Maggiora, G. Activity landscapes, information theory, and structure-activity relationships. *Mol. Inf.* **2013**, *32*, 421−430.
(21) Hu, X.; Hu, Y.; Vogt, M.; Stumpfe, D.; Bajorath, J. MMP-cliffs: systematic identification of activity cliffs on the basis of matched molecular pairs. *J. Chem. Inf. Model.* **2012**, *52*, 1138−1145.
(22) Stumpfe, D.; Hu, Y.; Dimova, D.; Bajorath, J. Recent progress in understanding activity cliffs and their utility in medicinal chemistry: miniperspective. *J. Med. Chem.* **2014**, *57*, 18−28.
(23) Medina-Franco, J. L.; Mendez-Lucio, O.; Yoo, J. Rationalization of activity cliffs of a sulfonamide inhibitor of DNA methyltransferases with induced-fit docking. *Int. J. Mol. Sci.* **2014**, *15*, 3253−3261.
(24) Husby, J.; Bottegoni, G.; Kufareva, I.; Abagyan, R.; Cavalli, A. Structure-based predictions of activity cliffs. *J. Chem. Inf. Model.* **2015**, *55*, 1062−1076.
(25) Hu, H.; Bajorath, J. Systematic exploration of activity cliffs containing privileged substructures. *Mol. Pharmaceutics* **2020**, *17*, 979−989.
(26) Horvath, D.; Marcou, G.; Varnek, A.; Kayastha, S.; de la Vega de Leon, A.; Bajorath, J. Prediction of activity cliffs using condensed graphs of reaction representations, descriptor recombination, support vector machine classification, and support vector regression. *J. Chem. Inf. Model.* **2016**, *56*, 1631−1640.
(27) de la Vega de León, A.; Bajorath, J. Prediction of compound potency changes in matched molecular pairs using support vector regression. *J. Chem. Inf. Model.* **2014**, *54*, 2654−2663.
(28) Heikamp, K.; Hu, X.; Yan, A.; Bajorath, J. Prediction of activity cliffs using support vector machines. *J. Chem. Inf. Model.* **2012**, *52*, 2354−2365.
(29) Duvenaud, D.; Maclaurin, D.; Aguilera-Iparraguirre, J.; Gomez-Bombarelli, R.; Hirzel, T.; Aspuru-Guzik, A.; Adams, R. P. Convolutional Networks on Graphs for Learning Molecular Fingerprints. 2015, arXiv:1509.09292. arXiv.org e-Print archive. https://arxiv.org/abs/1509.09292.
(30) Kipf, T. N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. 2016, arXiv:1609.02907. arXiv.org e-Print archive. https://arxiv.org/abs/1609.02907.
(31) Scarselli, F.; Gori, M.; Tsoi, A. C.; Hagenbuchner, M.; Monfardini, G. The graph neural network model. *IEEE Trans. Neural Netw.* **2009**, *20*, 61−80.
(32) Battaglia, P. W.; Hamrick, J. B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; Gulcehre, C.; Song, F.; Ballard, A.; Gilmer, J.; Dahl, G.; Vaswani, A.; Allen, K.; Nash, C.; Langston, V.; Dyer, C.; Heess, N.; Wierstra, D.; Kohli, P.; Botvinick, M.; Vinyals, O.; Li, Y.; Pascanu, R. Relational Inductive Biases, Deep Learning, and Graph Networks. 2018, arXiv:1806.01261. arXiv.org e-Print archive. https://arxiv.org/abs/1806.01261.
(33) Korolev, V.; Mitrofanov, A.; Korotcov, A.; Tkachenko, V. Graph convolutional neural networks as general-purpose property predictors: the universality and limits of applicability. *J. Chem. Inf. Model.* **2020**, *60*, 22−28.
(34) Liu, C.-H.; Korablyov, M.; Jastrzębski, S.; Włodarczyk-Pruszyński, P.; Bengio, Y.; Segler, M. H. RetroGNN: Approximating Retrosynthesis by Graph Neural Networks for De Novo Drug Design. 2020, arXiv:2011.13042. arXiv.org e-Print archive. https://arxiv.org/abs/2011.13042.
(35) Xu, K.; Hu, W.; Leskovec, J.; Jegelka, S. How Powerful are Graph Neural Networks?. 2018, arXiv:1810.00826. arXiv.org e-Print archive. https://arxiv.org/abs/1810.00826.
(36) Murphy, R.; Srinivasan, B.; Rao, V.; Ribeiro, B. In *Relational Pooling for Graph Representations*, International Conference on Machine Learning; PMLR, 2019; pp 4663−4673.
(37) Sato, R. A Survey on the Expressive Power of Graph Neural Networks. 2020, arXiv:2003.04078. arXiv.org e-Print archive. https://arxiv.org/abs/2003.04078.
(38) Hussain, J.; Rea, C. Computationally efficient algorithm to identify matched molecular pairs (MMPs) in large data sets. *J. Chem. Inf. Model.* **2010**, *50*, 339−348.
(39) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Felix, E.; Magarinos, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Maranon, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **2019**, *47*, D930−D940.
(40) Landrum, G. *Rdkit Documentation*, Release, 2013.
(41) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* **1988**, *28*, 31−36.
(42) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146−3154.
(43) Chen, T.; Guestrin, C. In *XGBoost: A Scalable Tree Boosting System*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 2016; p 785794.
(44) Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5−32.
(45) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273−1280.
(46) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742−754.
(47) Dimova, D.; Stumpfe, D.; Bajorath, J. Method for the evaluation of structure-activity relationship information associated with coordinated activity cliffs. *J. Med. Chem.* **2014**, *57*, 6553−6563.
(48) Wang, M.; Zheng, D.; Ye, Z.; Gan, Q.; Li, M.; Song, X.; Zhou, J.; Ma, C.; Yu, L.; Gai, Y.; Xiao, T.; He, T.; Karypis, G.; Li, J.; Zhang, Z. Deep Graph Library: A Graph-Centric, Highly-Performant Package for Graph Neural Networks. 2019, arXiv:1909.01315. arXiv.org e-Print archive. https://arxiv.org/abs/1909.01315.
(49) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 8026−8037.
(50) Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **2020**, *21*, 1−13.
(51) Di Cera, E. Thrombin. *Mol. Aspects Med.* **2008**, *29*, 203−254.

(52) Contet, C.; Kieffer, B. L.; Befort, K. Mu opioid receptor: a gateway to drug addiction. *Curr. Opin. Neurobiol.* **2004**, *14*, 370−378.
(53) Fani, L.; Bak, S.; Delhanty, P.; Van Rossum, E.; Van Den Akker, E. The melanocortin-4 receptor as target for obesity treatment: a systematic review of emerging pharmacological therapeutic options. *Int. J. Obes.* **2014**, *38*, 163−169.
(54) Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. In *Grad-cam: Visual Explanations from Deep Networks via Gradient-Based Localization*, Proceedings of the IEEE International Conference on Computer Vision; IEEE, 2017; pp 618−626.
(55) Lu, T.; Soll, R. M.; Illig, C. R.; Bone, R.; Murphy, L.; Spurlino, J.; Salemme, F. R.; Tomczuk, B. E. Structure-activity and crystallographic analysis of a new class of non-amide-based thrombin inhibitor. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 79−82.
(56) Chirgadze, N.; Sall, D.; Briggs, S.; Clawson, D.; Zhang, M.; Smith, G.; Schevitz, R. The crystal structures of human -thrombin complexed with active site-directed diamino benzo [b] thiophene derivatives: a binding mode for a structurally novel class of inhibitors. *Protein Sci.* **2008**, *9*, 29−36.
(57) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455−461.
(58) Valdés-Tresanco, M. S.; Valdés-Tresanco, M. E.; Valiente, P. A.; Moreno, E. AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4. *Biol. Direct* **2020**, *15*, 1−12.
(59) *The PyMOL Molecular Graphics System*; Schrödinger, LLC, 2022. https://pymol.org/2/.
(60) Pattanaik, L.; Ganea, O. E.; Coley, I.; Jensen, K. F.; Green, W. H.; Coley, C. W. Message Passing Networks for Molecules with Tetrahedral Chirality. 2020, arXiv:2012.00094 2020. arXiv.org e-Print archive. https://arxiv.org/abs/2012.00094.
(61) Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. 2017, arXiv:1710.10903. arXiv.org e-Print archive. https://arxiv.org/abs/1710.10903.

https://doi.org/10.1021/acs.jcim.2c00327
J. Chem. Inf. Model. 2022, 62, 2341−2351