Article

Ligand Binding Prediction Using Protein Structure Graphs and Residual Graph Attention Networks

Mohit Pandey [1], Mariia Radaeva [1] , Hazem Mslati [1], Olivia Garland [1] , Michael Fernandez [1], Martin Ester [2] and Artem Cherkasov [1,*]

[1] Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

[2] School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

[*] Correspondence: acherkasov@prostatecentre.com

Abstract: Computational prediction of ligand–target interactions is a crucial part of modern drug discovery as it helps to bypass high costs and labor demands of in vitro and in vivo screening. As the wealth of bioactivity data accumulates, it provides opportunities for the development of deep learning (DL) models with increasing predictive powers. Conventionally, such models were either limited to the use of very simplified representations of proteins or ineffective voxelization of their 3D structures. Herein, we present the development of the PSG-BAR (Protein Structure Graph-Binding Affinity Regression) approach that utilizes 3D structural information of the proteins along with 2D graph representations of ligands. The method also introduces attention scores to selectively weight protein regions that are most important for ligand binding. Results: The developed approach demonstrates the state-of-the-art performance on several binding affinity benchmarking datasets. The attention-based pooling of protein graphs enables identification of surface residues as critical residues for protein–ligand binding. Finally, we validate our model predictions against an experimental assay on a viral main protease (Mpro)—the hallmark target of SARS-CoV-2 coronavirus.

Keywords: deep learning; drug–target interaction; graph attention network; computer-aided drug discovery; protein–ligand binding; virtual screening; SARS-CoV-2

1. Introduction

The drug discovery field is constantly evolving to yield safe and potent drugs in the most time-, cost- and labor-efficient way. Traditionally, wet-lab high-throughput screening (HTS) was performed with libraries of compounds tested against a target of interest (typically a protein) to identify potential drug candidates [1]. The drawbacks of such an approach are high cost and labor demand, further hindered by a modest (on average 0.03%) hit rate [2–4]. Thus, effective computational methods are urgently needed to help accelerate drug discovery [2,5].

For instance, virtual screening helps to narrow down the score of compounds for experimental validation from large chemically diverse databases, thereby saving time and resources [6,7]. The mostly used virtual screening tool, molecular docking, relies mainly on physics-based or statistically derived scoring functions (SFs) to predict the binding affinity of ligands. Docking has achieved impressive advances in refining the drug-discovery pipeline and helped identify many potent and selective drug candidates [8,9]. Although docking is significantly faster than other virtual screening tools such as quantum mechanics-based approaches and free-energy perturbation simulations, its speed is limited and does not allow covering the wealth of available chemical structures. For example, the new release of the ZINC database [10]. containing over a billion molecules is impossible to screen with conventional tools and thus large chemical space remains unexplored. To overcome this issue, machine learning has been widely integrated into drug-discovery pipelines. As such,

DeepDocking [7]. integrates quantitative structure–activity relationship deep models with conventional docking software and achieves a 50-fold acceleration. Other machine learning (ML) models designed to predict drug–target interactions were shown to be capable of capturing the non-linear patterns and inferring complex binding rules [11,12]. In fact, some of the traditional ML models and recently developed deep learning (DL) models outperformed classical virtual high throughput screening (vHTS) approaches in terms of speed and predictive performance [9].

Several previous studies have leveraged DL-based models to predict drug–target interaction (DTI) and interaction binding affinity [13]. An efficient model should integrate both protein target and small molecule information and go beyond traditional pharmacophore descriptors [14]; one-dimensional descriptors learned from protein sequences and ligand atoms, respectively, have gained significant traction [15–17]. Thus, the DeepDTI approach predicted DTI using a deep belief network model fed with extended-connectivity fingerprints for drug representations and sequence composition descriptors for proteins [18]. DeepDTA [19] and DeepConv-DTI [20] exploit a CNN architecture with protein and ligand sequences as inputs to predict binding affinity for kinases.

However, complex patterns of protein–ligand interactions can only be captured with more realistic 3D structures and 3D and higher-order descriptors that have been in development over several decades by many groups around the world [21], including our own [22–24]. Importantly, databases of experimentally derived protein crystal structures also grow by the day [25], while powerful predictive tools such as AlphaFold [26] further supplement the wealth of available protein structural data. To this end, recent studies more readily employ structural protein information alongside ligands. These methods demand faithful representation of biological targets—such as proteins—and drug compounds. Inspired by the advances in computer vision and 3D object recognition, researchers have explored 3DCNN for modeling protein structures [11,27,28]. As such, Jimenez et al. (2018) [27] employed 3D convolutional neural networks (3D-CNN) to predict ligand-binding affinities. Jones et al. (2021) [11] developed a model to predict ligand–protein binding affinity using a fusion of 3D-CNN and spatial graph convolutional neural network.

While these approaches yield excellent predictive results, grids with atomic voids within the structure makes them susceptible to sparsity, memory bottleneck, and inefficient computations. Furthermore, CNNs are sensitive to rotation and orientation. While this may be useful for different conformers of ligands, proteins are typically considered as fixed rigid bodies for binding. To overcome this expensive computation of 3DCNN for large protein structures, alternative rotation invariant representations, namely Graph Neural Networks (GNN), have been investigated.

In contrast to CNNs, graphs provide efficient and succinct representations of 3D structures. Several studies have reported superior performance on binding affinity prediction by representing proteins as a network of their amino acid residues [29–33]. These methods model protein–ligand complexes as graphs and learn meaningful features from them. In this spirit, GraphBAR [31] modelled protein–ligand complexes as adjacency matrices at various spatial resolution of atoms of the binding site. Similarly, Lim et al. (2019) [29] proposed a distance-aware representation of the protein–ligand complex using two adjacency matrices. While the spatial features from the protein–ligand complex are certainly most relevant to binding affinity prediction, long-range interactions between protein and the ligand are also critical [34]. In their recent work, Li et al. proposed the Structure-aware Interactive Graph Neural Network (SIGN). In addition to learning the complex's spatial features, SIGN also considered long-range protein–ligand interactions within a certain cutoff distance by building an atom co-occurrence matrix.

These methods report exceptional results on the affinity prediction task on PDBBind or CSAR-HiQ [35]. However, in these datasets, the active pocket structure of the protein and the protein–ligand complex is known. Obtaining such information is expensive and limited to relatively smaller datasets and hence the applicability of such methods is restricted to a few carefully curated datasets. Our method does not explicitly take protein–ligand complex

structures as input, thereby expanding its generalizability to many larger benchmarking datasets. Inspired by Li et al.'s work on capturing long-range protein–ligand interactions, we define an attention-based strategy to learn important residues for binding.

Similar to our work, GEFA [30] fuses two separate graphs (i) for protein using amino acid sequence and contact map and (ii) for the drug molecules from their constituent atoms. Their strategy of masking-out the predicted unimportant regions based on self-attention scores is restrictive as it relies strongly on the accuracy of their method's binding region prediction. On the other hand, attention scores of our proposed method are used to soft-guide the protein graph's pooling in accordance with residues' importance concerning binding.

To alleviate these limitations and challenges of classical docking as well as previous deep-learning approaches, we introduce a graph-Siamese-like network with attention-based graph pooling. Benefitting from the advantages of protein crystallographic structures, the developed method applies a graph attention network to all the protein structures to predict several measures of ligand–protein binding affinities [36]. The graph-Siamese-like network facilitates concurrent feature extraction and fusion of the protein and ligand graphs by the predictive model. The use of attention enables our model to locate the likely active residues of the protein for ligand binding, hence reducing the overhead of time-consuming manual preparation of protein structures necessary for docking. For the model evaluation, we employed publicly accessible databases—PDBbind 18, BindingDB 19, KIBA [33], and DAVIS [37].—containing affinity measures for 9777, 123,904, 23,732 and 41,142 complexes, respectively (Table 1). We consequently retrieved protein structural information from the PDB database and for the KIBA dataset supplemented missing structures with those predicted by AlphaFold [26,38]. We demonstrate that the use of AlphaFold-predicted structures enhances the performance of the model and that our approach effectively predicts ligand–protein binding affinity and outperforms several existing baselines and circumvents expensive target site preparation needed for docking.

Table 1. Data statistics for datasets used for evaluation of PSG-BAR.

Dataset	Unique Targets	Unavailable PDBs	Unique Ligands	Unique Interactions
PDBBind	9619	0	7981	9777
KIBA	467	181	2356	124,374
DAVIS	442	139	68	20,604
BindingDB	1038	100	13,222	41,142
AID 1706	1	0	290,765	290,765

Our main contributions in this paper are

(i) We propose an effective graph-Siamese-like network, PSG-BAR, to simultaneously operate on structural graphs of proteins and molecular graphs of drugs for binding affinity prediction.

(ii) We introduce an attention-based readout strategy to generate graph-level embedding for proteins. These attention scores are shown to correlate with some known physical properties of binding sites. We call this learning component the interaction attention module.

This paper is organized as follows: In Section 2, we describe the datasets, their statistics, and the details of our proposed method. Section 3 features predictive results of our method on four benchmarking datasets and its application for several pertinent drug-discovery tasks. Further in the same section, we present a multi-view analysis of results and investigate the effect of architectural sub-components on the overall performance of our method. Finally, Section 4 discusses future directions and implications of this work.

2. Methods

2.1. Dataset

In this work, we used four publicly available datasets on protein–ligand binding affinity including BindingDB [39], PDBBind v2016 [40], KIBA [33]., and DAVIS [37]. that contain various binding affinity (Kd, Ki) datasets. First, we retrieved all protein–ligand pairs with the associated dissociation constant (Kd) from the BindingDB database [39]. Following data curation procedures such as removing duplicates, inorganic compounds, invalid SMILES, and Kds outside of typical range, and deriving associated structure files for each protein (PDBs with lowest resolution per protein target), the dataset was reduced to 41,142 protein–ligand pairs with 1038 unique proteins and 13,222 unique ligands. The pKd values had a mean of 6 and a standard deviation of 1.57. Similarly, we derived 13,478 unique interactions from PDBBind [40]. The 2016 dataset version of the PDBBind was used in this study in order to provide a consistent comparative analysis against a set of aforementioned studies. This PDBBind dataset includes three subsets: a core, a general, and a refined database totaling 10,790 unique protein–ligand complexes which were narrowed down to 8000 unique complexes after filtering on entries associated with Kd values. The Kd values were also transformed to pKd and had a mean STD of 6.35 and 1.89, respectively.

In addition to those values, we also retrieved data from the KIBA dataset [33] which consists of a drug–target matrix with information on three major bioactivity types—IC50, Ki, and Kd. Protein targets from the KIBA were labelled with UniProt IDs, which were then mapped to their respective PDB IDs to derive crystal structure. KIBA scores were transformed similarly to Shim et al. (2021) [41] and He et al. (2017) [42]. In particular, drugs and targets with less than 10 interactions were filtered out; then, KIBA scores were transformed to resemble pKd. Of the original 467 targets, after filtering and mapping to available PDB IDs, 286 targets remained. Of the 52,498 drugs, after filtering, there remained 2356. This yielded 124,374 interaction pairs.

We also evaluated the proposed model on the DAVIS dataset [37]. containing 72 kinase inhibitors tested over 442 kinases which represent over 80% of the human catalytic protein kinome. The activity number distribution was similar to BindingDB and PDBBind with a mean of 5.48 and STD of 0.928. Protein targets were labelled by name, so as to retrieve corresponding PDB structures, the RCSB PDB [43] Search API was used to perform a keyword search. Of the 442 proteins, 303 could be mapped to a PDB structure covering 68 of the initial 72 kinase inhibitors. In total, this yielded 20,604 drug–target interaction pairs.

To further evaluate the applicability of our method for the SARS coronavirus (SARS-CoV-2), we attempted to predict inhibitors of the SARS-CoV-2 3C-like protease (3CLpro), also known as the main protease (Mpro). The bioassay record (AID 1706) [44] by Scripps Research Institute provides a PubChem Activity Score normalized to 100% observed primary inhibition. As suggested by the authors of AID 1706, a threshold for the activity was at 15 to get 444 inhibitors for 3CLpro along with 290,321 inactive. The protein structure corresponding to the crystal structure of SARS-CoV-2 Mpro in the complex with an inhibitor N1 (1WOF [45]) was used.

For each of the datasets, certain targets and ligands were dropped on account of the unavailability of the crystal structures and un-processability by RDKit, respectively. Table 1 lists the counts of eventually used protein–ligand interaction pairs for each dataset.

### 2.2. Protein and Ligand Graph Construction

The intricate 3D structure of proteins that develop as a result of complex folding largely determine the functions and properties of protein molecules. Geometric deep learning provides an efficient way to represent such 3D structures as translation and rotation invariant graphs. These protein graphs can be created at the residual or atomic level. We conduct all our experiments on residual-level graphs where each amino acid corresponds to a unique attributed node in the protein graph $\mathcal{G}_p$. We choose residue-level graphs over atomic ones because the latter require more expensive computational processing. Formally, these graphs are constructed by connecting the five nearest neighbors for each residue from the protein's

contact map. The features utilized for the nodes of PSG are one-hot encoded amino acids, 3D coordinates of the alpha carbon atom of each residue, secondary structural features, Meiler embeddings [46], and expasy embeddings [47]. Further, the edges of these multirelational graphs are defined by intermolecular interactions and attributed by Euclidean distance between corresponding nodes. In particular, these interactions include peptide bonds, hydrogen bonds, ionic interactions, aromatic sulphur interactions, hydrophobic interactions, and cation–Π interactions as provided by Jamasb et al. (2020) [48].

To define molecular graphs for ligands $\mathcal{G}_l$, we utilize the Pytorch-geometric approach in which nodes represent atoms of a ligand, and edges denote molecular bonds. The 3D graph for ligand conformers may be promising; however, the optimal ligand conformation for the binders in most of the benchmarking datasets is unknown. To circumvent this conformational challenge we used a simplified 2D graph representation. The adjacency matrix indicates all node pairs whose corresponding atoms share a covalent bond between the respective atoms in the ligand. Similar to the protein graphs, the nodes and edges of the ligand graphs are attributed. We use the same atomic features for nodes as Kearnes et al. (2016) [49]: atom type, formal charge, hybridization ('sp', 'sp2', 'sp3'), hydrogen bond acceptor/donor, membership to aromatic ring, degree, connected hydrogen count, chirality, and partial charge. For the edges, bond type, bond conjugation, stereo configuration and membership to a shared ring were used as attributes. The ligand molecular graph and the PSGs are batched together similar to a graph matching or a Siamese network in order to serve as input to our predictive model.

2.3. Model Architecture

PSG-BAR receives two attributed graphs $\mathcal{G}_p$ and $\mathcal{G}_l$ corresponding to the protein 3D structure and molecular graph of ligand's 2D representation, respectively (Figure 1). The construction of $\mathcal{G}_p := (\mathcal{V}_p, \mathcal{E}_p, \mathcal{A}_p)$ and $\mathcal{G}_l := (\mathcal{V}_\ell, \mathcal{E}_\ell, \mathcal{A}_\ell)$ is as defined in Section 2.2. The model architecture follows an encoder–decoder paradigm with the encoder comprising of (i) protein encoder $\mathcal{F}_p$, (ii) cross attention head, and (iii) ligand encoder $\mathcal{F}_\ell$. We describe the cross-attention mechanism in Section 2.4. The decoder is a multilayer feed forward neural network with LeakyReLU activation.
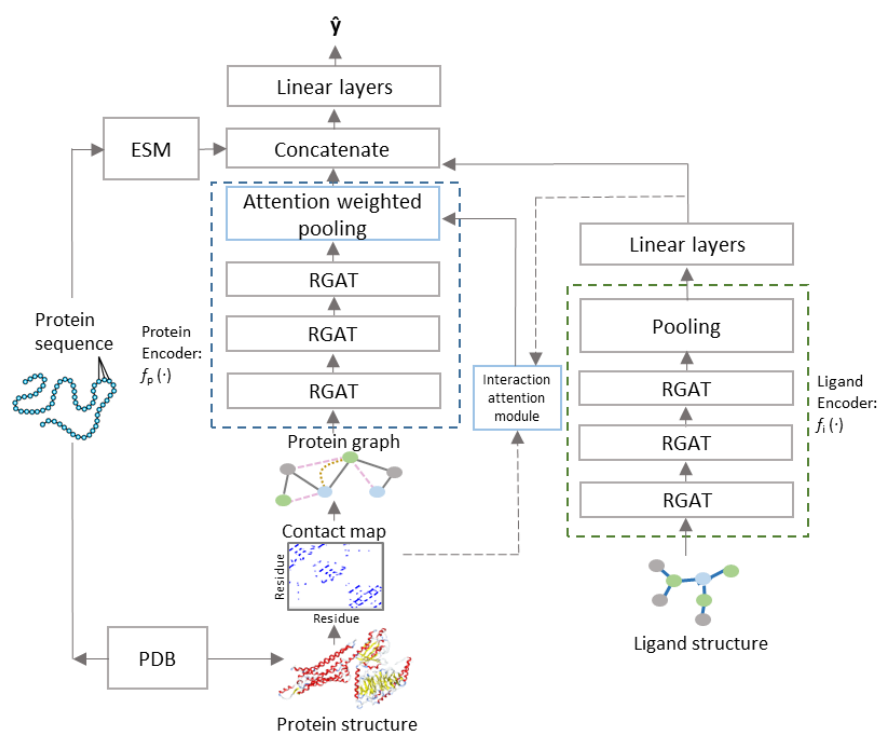
Figure 1. PSG-BAR Architecture.

$\mathcal{G}_p$ and $\mathcal{G}_l$ are processed by independent and architecturally identical protein and ligand encoders ($\mathcal{F}_p$ and $\mathcal{F}_\ell$, respectively) that stack several layers of GAT with skip connection. For the ligand encoder ($\mathcal{F}_\ell$), the vector representations learned by GAT layers are aggregated over all nodes of the graphs using a readout function $r(.)$ that combines global max and average pooling over all nodes of the graph to generate graph level representations (x'_l). The readout function for the protein encoder performs global max pooling over the node embeddings weighted by the attention map learned by the interaction attention module. The encoded latent representations for protein are further enriched with continuous dense embedding for the protein backbone, provided by a pretrained language model trained on amino acid sequences [50]. $x'_p = [\mathcal{F}_p(x_p) || x_p{}^{seq}]$ and $x'_l = \mathcal{F}_\ell(x_l)$. The protein and ligand representations interact with each other in a Siamese-like fusion approach (Figure 2). The decoder comprises a multilayered perceptron with LeakyReLU activation and forms the predictor function.

$$x = \left[x'_p \middle|\middle| x'_l \right]$$
$$\hat{y} = LeakyReLU\left(\mathbf{W}^T x \right)$$
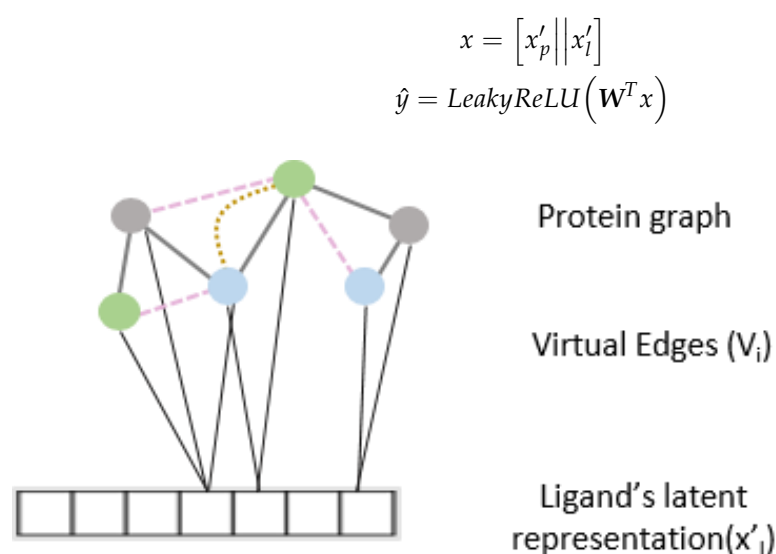
Figure 2. Virtual edge set to calculate cross-attention between protein graph nodes and learned representation of the drug.

2.4. Interaction Attention Module

To further enhance the predictive capability of our model and to learn from the relationship between protein and ligand interaction, we propose an interaction attention module based on the cross attention mechanism (Figure 2). Cross attention learns to selectively attend to the nodes of the protein structure graph and hence identify principal nodes for a given protein–ligand interaction pair. We create the virtual edge set $V = \{uv \mid u \in \mathcal{V}_p, v = x_l\}$ to connect all nodes of the protein graphs to the drug representation out of the ligand encoder. V_i is calculated as

$$V_i = \text{LeakyReLU}\left(U_1^T x'_l + U_2^T u_i \right)$$

where, U_1, U_2 are trainable model parameters and $u_i \in \mathcal{V}_p$.

2.5. Training and Hyperparameter Tuning

In all our experiments we employed a base model with three stacked GAT layers as encoder (Figure 1). LeakyReLU activation was used throughout the model, except in the interaction attention module where *tanh* was used. All GAT layers had a dropout rate of 10% and batch normalization to avoid overfitting. Further, early stopping conditioned on validation loss was also adopted. The Adam optimizer with starting learning rate of 0.07 and a decaying learning rate scheduler was employed for all our experiments. A minibatch size of 256 was found to be the most suitable on our Nvidia Tesla V100 GPUs.

Mean squared error and cross entropy loss were used as objectives for regression and classification, respectively. All hyperparameters were empirically chosen to maximize the Pearson correlation metric on the validation set. Our methods were implemented using Pytorch-geometric and Pytorch.

3. Results

Evaluation metrics: Our models were evaluated for regression using the Pearson coefficient and mean squared error (MSE). Additionally, for activity classification on SARS-CoV datasets, we report cross entropy loss and area under the curve of receiver operating characteristic curve.

3.1. Binding Affinity Prediction Results

We benchmark PSG-BAR against several state-of-the-art methods (Table 2) that consider spatial information or are based on GNNs. For KIBA, we compare against published results with the lowest reported MSE. For BindingDB, we chose methods that reported regression results for binding affinity on this dataset. Additionally, to test the robustness and generalizability of our method, we test the performance under the following settings for each benchmark dataset (Table 3). These settings are based on the different stratification criteria of proteins and ligands for the train–test split.

Table 2. Comparison of PSG-BAR to other state-of-the-art methods on popular benchmark datasets. For each of the 4 datasets, we compare PSG-BAR to the best reported performance on that dataset that we find in the literature survey. For brevity, RSME data were consolidated into MSE when not directly available from the authors. Since most works report their performance on warm setting, we present our benchmarking results in the same setting. For GEFA and DeepPurpose, we were able to reproduce respective implementations and report results on the same subset of the dataset as PSG-BAR. For other works, we report the best reported performance metric in the literature for these methods.

Dataset	Architecture	MSE ($\downarrow$)	Pearson ($\uparrow$)
DAVIS	GCNConvNet [51]	0.284	0.804
	GINConvNet [51]	0.257	0.824
	DGraphDTA [32]	0.241	0.837
	GEFA [30]	**0.228**	0.846
	PSG-BAR	0.237	**0.856**
KIBA	KronRLS [52]	0.261	0.752
	GANsDTA [53]	0.387	0.662
	SimCNN-DTA [41]	0.257	0.757
	SimBoost [42]	0.204	-
	PSG-BAR	0.200	0.850
	PSG-BAR w/AF	**0.177**	**0.865**
BindingDB	DeepAffinity [54]	1.212	0.700
	DeepPurpose [55]	0.765	0.836
	PSG-BAR	**0.651**	**0.864**
PDBBind	GAT [56]	3.115	0.601
	SGCN [57]	2.505	0.686
	SIGN [56]	1.731	0.797
	KDeep [27]	**1.612**	**0.820**
	PSG-BAR	1.660	0.762

Table 3. PSG-BAR performance on various dataset stratification schemes.

Dataset	Architecture	MSE ($\downarrow$)	Pearson ($\uparrow$)
DAVIS	Warm	0.237	0.856
	Cold Drug	0.902	0.456
	Cold Protein	0.436	0.612
	Cold Protein–Ligand	0.910	0.357
KIBA	Warm	0.200	0.850
	Cold Drug	0.362	0.601
	Cold Protein	0.298	0.756
	Cold Protein–Ligand	0.415	0.360
BindingDB	Warm	0.651	0.864
	Cold Drug	1.353	0.720
	Cold Protein	1.811	0.540
	Cold Protein–Ligand	2.102	0.515
PDBBind	Warm	1.660	0.762
	Cold Drug	1.895	0.694
	Cold Protein	2.011	0.602
	Cold Protein–Ligand	2.100	0.599

Warm setting: No splitting restriction imposed on proteins and ligands. Any protein (or ligand) may be repeated in training and test split; however, interactions are not duplicated across the two splits.

Cold protein setting: Each unique protein in the dataset has restricted membership to either the training or the test set. No restrictions on ligands.

Cold ligand setting: Each unique drug in the dataset has restricted membership to either the training or the test set. No restrictions on proteins.

Cold protein–ligand setting: Any protein or ligand seen during training is absent in the test set. This is the most stringent measure for our model.

3.2. SARS Inhibitor Prediction

To perform a binary classification on the SARS-CoV Mpro inhibitor and to overcome the problem of class imbalance between actives and inactives, we oversampled the actives and randomly subsampled inactives to construct a class-balanced dataset of 26,400 interactions. With a train–test split of 80–20% and five-fold CV, our method yielded 0.72 ROC-AUC (Figure 3A). The sparsity of actives within an expansive and diverse chemical space warrants identification of even slightly probable hit compounds for further evaluation. As such, we optimize for recall by using weighted binary cross entropy loss.

*3.3. SARS CoV-2 MPro Experimental Validation*

To validate the utility of PSGBAR method, we additionally perform testing on the novel coronavirus, SARS-CoV-2, the causative agent of the ongoing COVID-19 global pandemic. We rank PSG-BAR scores against docking screen results for the Mpro target. This docking screen [58] was performed on the Mpro crystal structure (6W63 [59]) using the DeepDocking [7] tool coupled with the commercial software Glide [60]. Importantly, after the docking ranking, the final set of 1200 molecules was selected by expert chemists. These selected compounds were then experimentally tested and 116 actives were confirmed. We observe that our model highly ranks the compounds that the human expert decided are worth purchasing. We show the normalized histograms of predicted scores for docking screen compounds (blue) and decoys (orange). The details of our findings are as follows:

The predicted score distribution for hits from docking is right skewed, meaning that the model highly ranks most compounds in agreement with the docking program's ranking.

The flat, near-zero left tail indicates that the tested set has almost no compounds which are predicted to be extremely poor binders by our method. This aligns with our a priori knowledge that these tested compounds are expert-selected following docking screen.

A set of decoys of a random sample of 20,000 poorly ranked compounds by docking yields low score on our model too. Figure 4 illustrates that, on average, hit compounds exhibit higher PSG-BAR scores than non-hits. This corroborates the utility of our method as a standalone method to screen large virtual libraries. However, the most lucrative positioning of our method would be as a filter succeeding docking and preceding hit compound purchase.

A)   B)
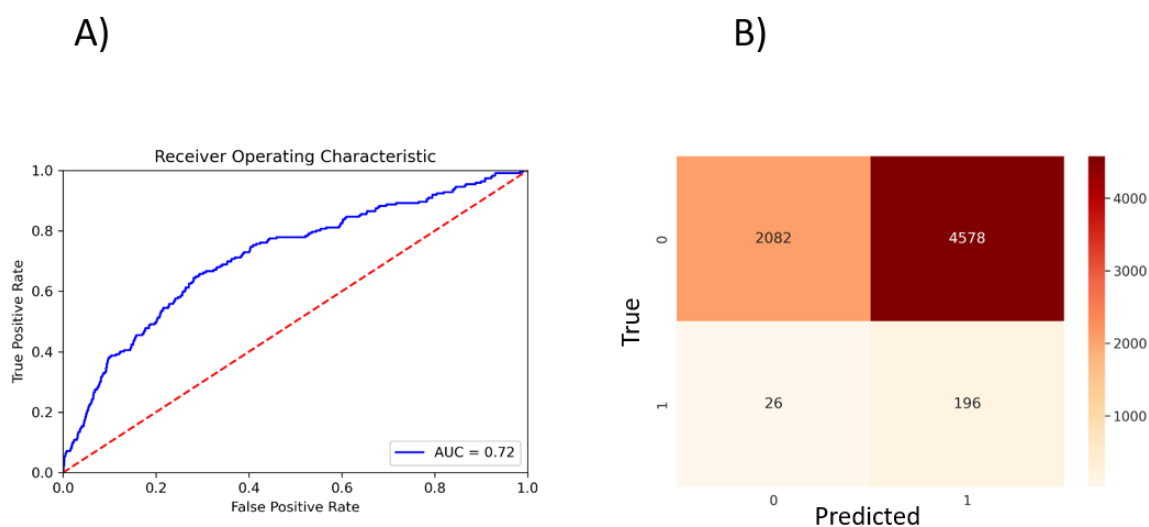
Figure 3. PSG-BAR predictions for SARS-CoV inhibitors. (**A**) ROC plot. (**B**) Confusion matrix. 0: inactives; 1: actives.
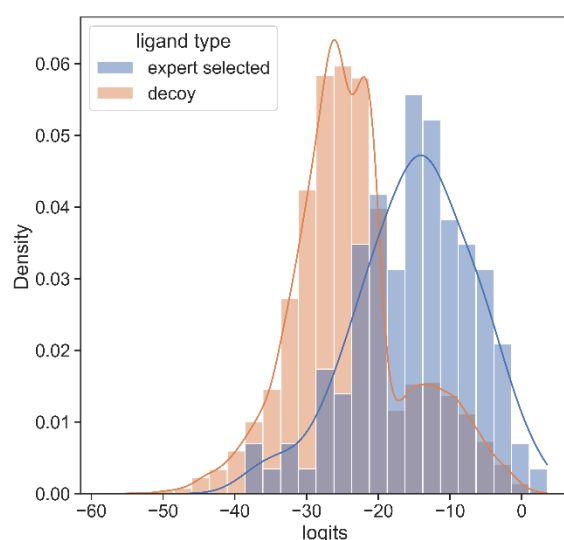
Figure 4. Histogram of PSG-BAR scores for docked compounds suggests that our method assigns higher-score hits from docking and low scores to poor binders. The distribution to the right shows the PSG-BAR scores for top ranking hits of docking screen for MPro target. The left distribution is of PSG-BAR scores for 20,000 randomly chosen compounds from docking screen that were not considered as hits. These compounds on average had lower PSG-BAR score (−30) compared to scores of hit compounds (−15). PSG-BAR scores are unnormalized logits of our trained deep neural network.

3.4. Attention Centrality

Thus far, we have evaluated the performance of our model as a whole on several datasets and use cases. We are further interested in investigating the interpretability of the method's predictions and understanding the cause of its superior performance. The relevance of the attention scores was evaluated on a randomly selected sample of 50 proteins from six protein families, namely, protease, kinase, polymerase, transferase, phosphodiesterase, and hydrolase, from the PDBBind dataset. We focus on the amino acid residues with the highest attention scores, which should be relevant for predicting binding. Figure 5 illustrates the distributions of the average Solvent Accessible Area (SAA) for all protein residues and the top-10 residues with the highest learned interaction attention scores. Notably, residues with the highest attention scores displayed higher mean SAA for all protein families. We also observed that, on average, surface residues have 22% higher attention scores than core residues. This aligns with our domain knowledge on protein–ligand binding that suggests the pocket residues are located on the protein surface or are usually somehow accessible to the solvent.
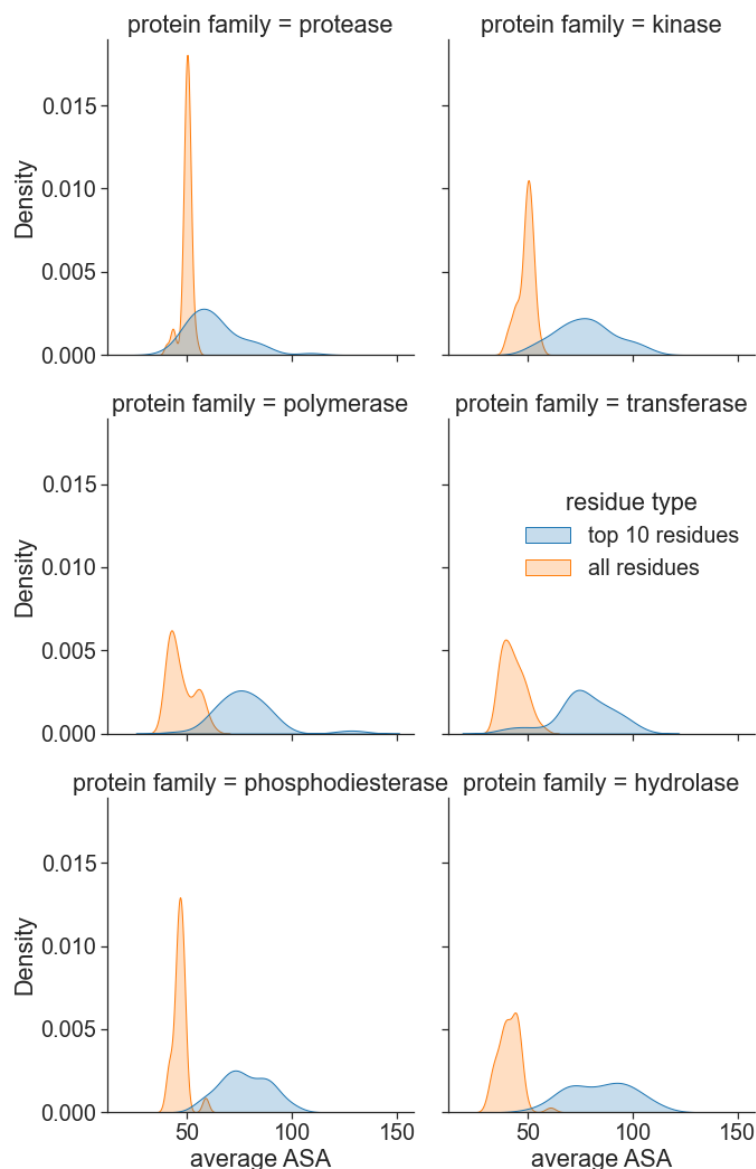
Figure 5. Interaction attention scores predict residues on the protein surface. For the 6 most frequent protein families in the PDBBind dataset, the 10 highest scored residues by PSG-BAR have mean SAA higher than the mean SAA of the entire protein.

3.5. Drug Promiscuity

PSG-BAR operates as a protein target agnostic method for ligand binding prediction. As such, it could be utilized to study the off-target effects of drug compounds on a multitude of human proteins. To this end, we evaluate the promiscuity of the hit compounds discovered in Section 3.3. To train a model that predicts activity across major off-target proteins, we used a dataset presented by Ietswaart et al. (2020) [61] that contains in vitro activity of 1866 marketed drugs across 35 protein targets (38,091 protein–ligand interactions). These targets are adverse-drug reaction (ADR)-related and include well-known proteins such as hERG (induces cardiotoxicity), nuclear receptors (carcinogenic effects), and COX enzymes (intestinal bleeding). Thus, our model predicts whether a molecule is likely to bind any of these off-target proteins. The modified PSG-BAR method used to perform classification for activity on the ADR dataset yields an ROC-AUC value of 0.88 (Figure 6A).
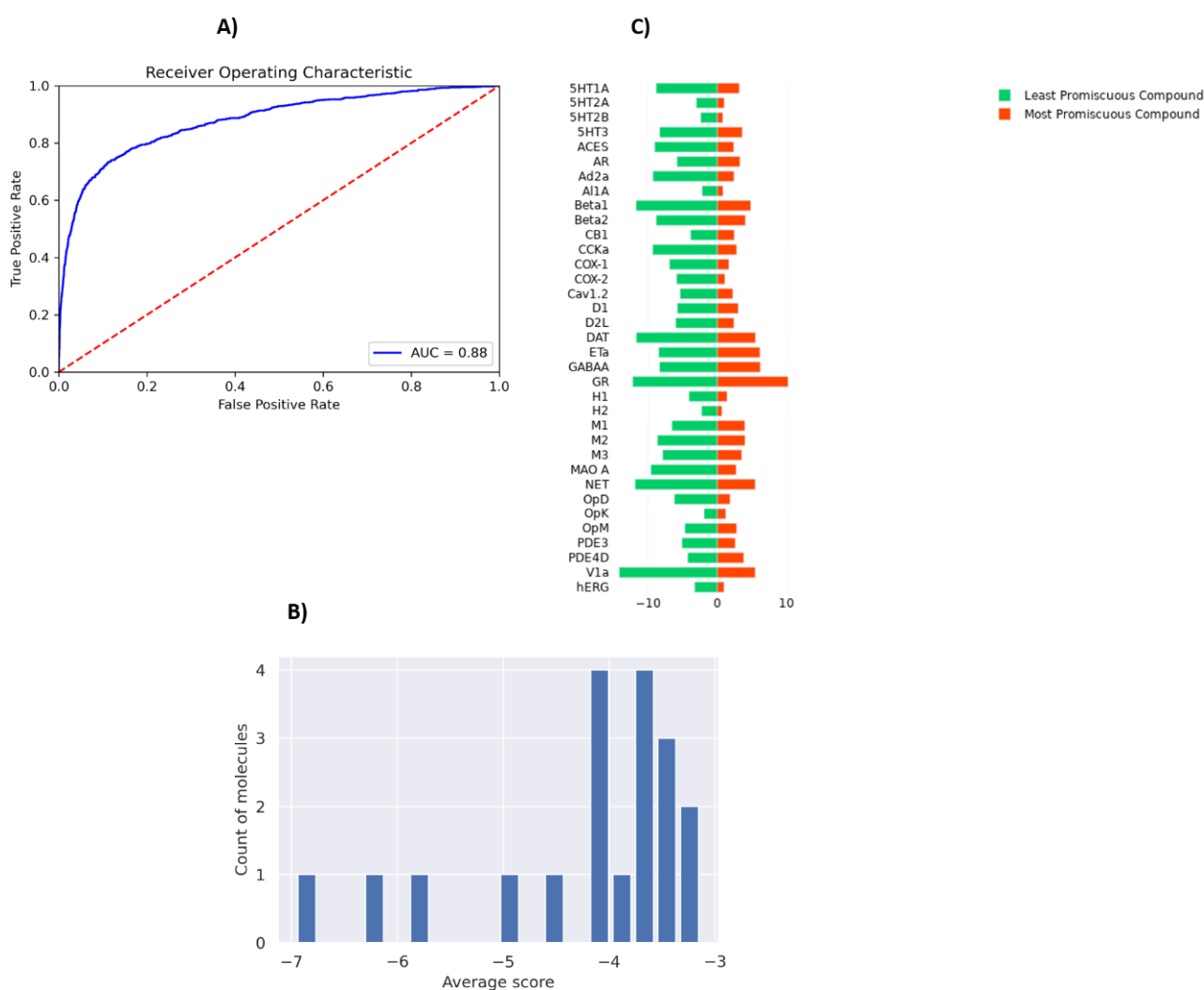
Figure 6. Evaluation of off-target effects of MPro hits on human proteins using PSG-BAR. (**A**) ROC-AUC on ADR dataset (**B**) Histogram of average scores across 35 ADR-related proteins for 19 predicted Mpro binders. The average score is a proxy for compound promiscuity. (**C**) Predicted likelihood of binding to ADR proteins for the most promiscuous molecule (red) and least promiscuous molecule (green). Lower values (closer to the center) indicate a high likelihood of binding.

To evaluate effects on ADR proteins, we chose 19 compounds ranked within the top 25th percentile by both PSG-BAR and docking screen. These compounds are predicted to be the most likely binders to Mpro by our pipeline. We averaged scores across all proteins to assess the general promiscuity of a molecule, i.e., the likelihood of binding to many

proteins. The average score ranged from -3 to -7, where a higher score means a more promiscuous compound (Figure 6B). We examined two molecules from both extremes. The most promiscuous molecule was predicted to most likely bind alpha-1A adrenergic receptor (Al1A), histamine H2 receptor (H2), COX-2, hERG, and 5-hydroxytryptamine receptor 2A (5T2A) and 2B (5HT2B) (Figure 6C). Thus, if this molecule in the process of drug development turned out to be problematic (e.g., exert toxicity in mice), our model can hint at the possible underlying off-targets. Such information might be extremely useful for structure-based optimization of the lead compound.

3.6. Ablation Studies

We critically analyze the contribution of key components of our model towards predictive performance. These studies were conducted on the PDBBind dataset under warm setting, train/validation split of 0.8/0.2 ratio, and with a fixed random seed to impose fair uniformity across experiments. We first study the effect of the skip connections in GATs. The protein and ligand encoders are equipped with a varying number of GAT layers while iteratively ablating the skip connections for each constituent layer.

We observe that stacking graph layers have diminishing returns (Figure 7A). This is consistent with the observations made by previous works as most state-of-the-art GNNs are shallower than four layers. Our experiments show similar peak performance for a three-layered network. In each of the stacking modes, the GAT network equipped with the skip connection performed better than the corresponding non-residual GAT version. Further stacking layers reduce the overall efficacy of our model; however, it continues to benefit from the addition of the skip connection. Furthermore, we noticed that over the early epochs of model training, non-residual GAT yields slightly better performance than the skip-connection variant and converges faster (Figure 7B). This may be attributed to (i) the comparatively simpler architecture of the former and (ii) the averaging of node representations to nearly the same vector to cause oversmoothing occurs over the due course of passes over the data.
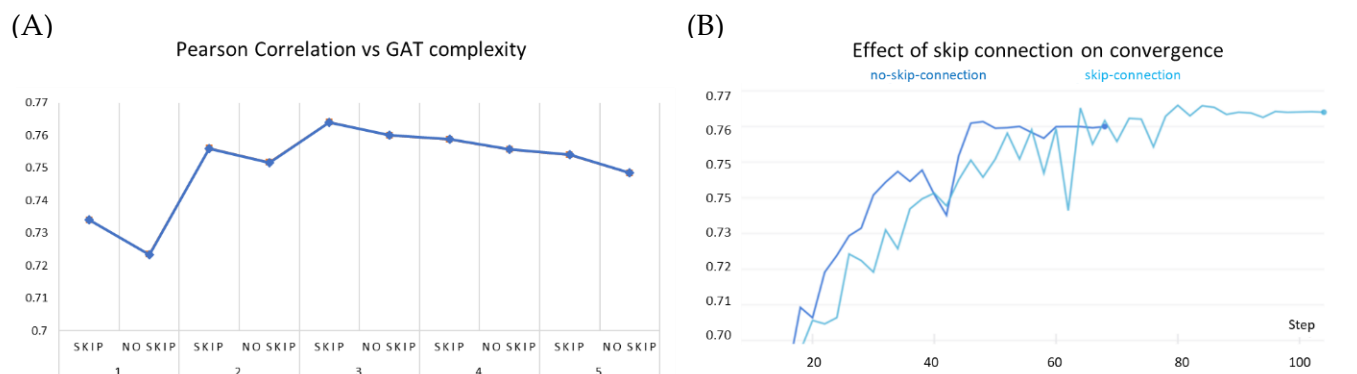
Figure 7. Effect of skip connection on model performance on PDBBind dataset. (**A**) skip connection on stacking GAT layers compared to GAT models of same complexity without skip connection. (**B**) skip connection vs no skip connection on 3-layered GAT: the early success of the no-skip variant is superseded by the skip connection variant over increasing epochs.

We further investigated the effect of the pre-trained amino acid embeddings on the overall performance. In the model without amino acid embeddings, we observed a 3.47% decrease in the Pearson correlation and 4.93% increase in the MSE on the PDBBind dataset. This is attributed to the fact that protein homology is driven by sequence similarity and a well-trained embedding on large amounts of amino acid sequences effectively captures such protein similarity.

3.6.1. Effect of Augmenting KIBA Dataset with AlphaFold Structures

As described in the dataset section, utilization of any protein–ligand interaction in the training (or validation) process was contingent upon the availability of the crystallographic structure of the corresponding protein target. For the KIBA dataset we supplemented the proteins with missing PDB with their corresponding structures from AlphaFold. Consequently, the additional interaction samples led to a 43.4% increase in available proteins in the usable dataset and a 19.67% increase in total interactions. The model performance also rose by 1.72% in terms of Pearson correlation and 11.1% for MSE (Table 2).

3.6.2. Effect of Secondary Structure Features of Proteins

We hypothesized that additional protein structure descriptors derived from the DSSP module will help to ameliorate the definition of binding pockets. Indeed, Stank et al. highlighted examples where secondary and tertiary features aided the classification of various protein pockets [62]. Furthermore, topological, solvation, and hydrophobicity descriptors may help determine the druggability of protein sites. For instance, it is known that most druggable sites are highly hydrophobic and relatively deep [63]. As expected, the aforementioned features boosted the performance of PSG-Bar with a Pearson correlation increase in the range of 1.3–1.76% and an MSE decrease in the range of 4.00% to 4.81% (Table 4). This could indicate the benefit of implicit learning of the connection between properties of amino acid residues in the pocket and the binding free energy of ligands.

Table 4. Protein surface features improve PSG-BAR predictions across all 4 benchmarked datasets.

	With Surface Features		Without Surface Features	
Dataset	**MSE**	**Pearson**	**MSE**	**Pearson**
BindingDB	0.651	0.864	0.678	0.851
PDBBind	1.660	0.762	1.744	0.749
KIBA	0.200	0.850	0.209	0.837
DAVIS	0.237	0.856	0.249	0.845

3.7. Error Analysis of Prediction of Effective Binders

It should be noted that drug discovery practitioners tend to care more for prediction concerning good binders rather than for the overall performance of computational approaches. To this extend, we stratify the affinity score range into four intervals as $\leq$10, 10–12, 12–14, and >14 for the KIBA dataset (Figure 8). A significantly large majority (78.4%) of interactions lie in the score 10–12 interval. These are weak binders, and the model performance is at its peak for these interactions. The MSE for this stratum is 0.093 compared to the population mean MSE of 0.2. Further, the moderate binders (score 12–14) and strong binders (score > 14) span a much smaller proportion of total interactions at 18.4% and 2.7%, respectively. For the poor binders, the predictive performance is considerably inferior to any other strata.

It should be mentioned, however, that these interactions are of relatively low significance to the design and discovery of drugs. Strong binders (score > 14) are unarguably the interval of highest interest and difficulty to predict. This, coupled with their sparsity in the evaluated subset of the data, leads to a mediocre result for strong binders (MSE 1.2 compared to MSE 0.2 overall). However, further elaborating the KIBA dataset with AlphaFold structures shows a marked improvement of 57.9% in MSE to bring MSE down to 0.50. This is due to the presence of more structurally similar proteins that enable the model to learn complex binding features, central to strong protein–ligand binding.
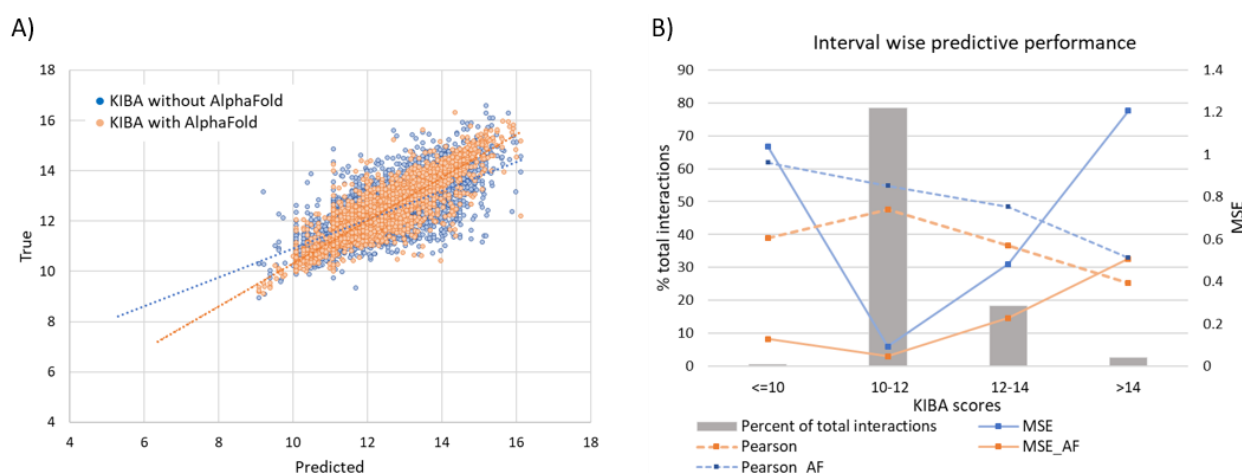
Figure 8. (**A**) Linear correlation of experimental vs predicted pKd values from the KIBA dataset. (**B**) Interval wise predictive performance on KIBA dataset. The most common interval of 10–12 has highest performance while the extremes on both sides have very poor performance.

4. Discussion and Conclusions

In this study we report the development of PSG-BAR, a ligand-binding affinity prediction method leveraging protein 3D structures, 2D graph representations of a drug molecule, and attention-based interaction scoring. The integration of protein structures helps to achieve better predictive results on protein–ligand binding. This is mainly because 3D structures contain relevant information on the actual configuration of the binding pockets, which have immediate implications for the ligand binding. Many reported studies concerning the protein–ligand interaction fail to consider the complex folded 3D structure of the proteins and employ just the primary protein sequence. Yet others that consider protein 3D structure fall short of considering the relevance of active sites of the protein and view all the protein residues as equally important. Some other studies utilizing protein–ligand complex information rely on computationally expensive methods to determine binding sites a priori and hence restrict their applicability to limited datasets.

In this work, we considered the residue-level graphs of the 3D structures of proteins in conjunction with molecular topological graphs of the ligands. Our Siamese-like architecture encodes both these graphs simultaneously. Additionally, we also proposed an interaction attention module that learns attention scores between the ligand's latent representation and nodes (amino acids) of the protein structural graph. This enabled our model to ascribe higher weights to residues critical concerning binding. Through a comparative analysis, we demonstrate that the learning from the proposed interaction attention module concurs with domain knowledge about SAA. The learned critical residues had higher SAA than average SAA of the protein for most frequent protein families in the PDBBind dataset.

Unlike most previous studies, PSG-BAR does not rely on expensive precomputation of binding sites and does not expect protein–ligand complex information explicitly. This interaction inference is conducted implicitly through the attention mechanism. As a result, it may be applied to many diverse and extensive binding datasets. Our experiments yielded state-of-the-art results on BindingDB, KIBA, and DAVIS datasets and comparable to state-of-the-art results on the PDBBind dataset.

This applicability of our method on diverse datasets enables its use as an "off-target detection tool". The training on the abundance of diverse proteins makes the model highly generalizable across different protein families. Thus, the model could be applied for predictions across a set of selected proteins such as ADR-relevant targets. We evaluated the selected Mpro lead compounds using our model built for prediction across ADR-related proteins. We found that the molecules range in their predicted promiscuity. Thus, these predictions might help to guide future lead optimization of the drug candidates.

We acknowledge other works with better-reported MSE, especially on the PDBBind dataset (KDeep). This gain is largely attributed to the utilization of protein–ligand complexes in their model, indicating that the binding pockets (active sites) of the proteins are most critical towards this downstream prediction. However, obtaining such a protein–ligand complex is expensive and hence limits the applicability of many of these methods to smaller datasets with selective protein targets. Similarly, conventional docking approaches depend on computationally expensive preparation of binding sites. Our attention-based method learns surface residues without direct supervision while simultaneously predicting binding affinity. This identification of key binding residues diminishes the need for expensive binding site preparation and makes our model accessible to minimally preprocessed protein structures.

In conclusion, our work represents a step in the direction of alleviating the problem of a priori knowledge of a binding site which is an expensive prerequisite to all protein–ligand interaction studies. The road to explainable AI, such as in attention-based visual question answering, is expected to reform deep learning for DTI. We expect more protein–ligand complexes to be experimentally resolved, so the next generation of deep models will be able to learn such attention scores more accurately. In fact, with soft-supervision these attention scores may even lead to reliable identification of binding sites. In this study, we work under the rigidity assumption of protein and ligand (by using only 2D molecular structures). Further research should investigate 3D ligand conformers in conjunction with flexible protein 3D structures.

References

1. Yang, X.; Kui, L.; Tang, M.; Li, D.; Wei, K.; Chen, W.; Miao, J.; Dong, Y. High-throughput transcriptome profiling in drug and biomarker discovery. *Front. Genet.* **2020**, *11*, 19. [CrossRef] [PubMed]
2. Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203–214. [PubMed]
3. Dara, S.; Dhamercherla, S.; Jadav, S.S.; Babu, C.M.; Ahsan, M.J. Machine Learning in Drug Discovery: A Review. *Artif. Intell. Rev.* **2022**, *55*, 1947–1999. [CrossRef] [PubMed]
4. Zhu, T.; Cao, S.; Su, P.-C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E. Hit identification and optimization in virtual screening: Practical recommendations based on a critical literature analysis: Miniperspective. *J. Med. Chem.* **2013**, *56*, 6560–6572. [CrossRef]

5. Medina-Franco, J. Grand challenges of computer-aided drug design: The road ahead. *Front. Drug Discov.* **2021**, *1*, 728551.

6. Maia, E.H.B.; Assis, L.C.; De Oliveira, T.A.; Da Silva, A.M.; Taranto, A.G. Structure-based virtual screening: From classical to artificial intelligence. *Front. Chem.* **2020**, *8*, 343. [CrossRef]

7. Gentile, F.; Agrawal, V.; Hsing, M.; Ton, A.-T.; Ban, F.; Norinder, U.; Gleave, M.E.; Cherkasov, A. Deep Docking: A Deep Learning Platform for Augmentation of Structure Based Drug Discovery. *ACS Cent. Sci.* **2020**, *6*, 939–949. [CrossRef]

8. Bender, B.J.; Gahbauer, S.; Luttens, A.; Lyu, J.; Webb, C.M.; Stein, R.M.; Fink, E.A.; Balius, T.E.; Carlsson, J.; Irwin, J.J.; et al. A practical guide to large-scale docking. *Nat. Protoc.* **2021**, *16*, 4799–4832. [CrossRef]

9. Wójcikowski, M.; Ballester, P.J.; Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci. Rep.* **2017**, *7*, 46710. [CrossRef]

10. Irwin, J.J.; Tang, K.G.; Young, J.; Dandarchuluun, C.; Wong, B.R.; Khurelbaatar, M.; Moroz, Y.S.; Mayfield, J.; Sayle, R.A. ZINC20—A free ultralarge-scale chemical database for ligand discovery. *J. Chem. Inf. Model.* **2020**, *60*, 6065–6073. [CrossRef]

11. Jones, D.; Kim, H.; Zhang, X.; Zemla, A.; Stevenson, G.; Bennett, W.F.D.; Kirshner, D.; Wong, S.E.; Lightstone, F.C.; Allen, J.E. Improved Protein–Ligand Binding Affinity Prediction with Structure-Based Deep Fusion Inference. *J. Chem. Inf. Model.* **2021**, *61*, 1583–1592. [CrossRef] [PubMed]

12. Pandey, M.; Fernandez, M.; Gentile, F.; Isayev, O.; Tropsha, A.; Stern, A.C.; Cherkasov, A. The transformational role of GPU computing and deep learning in drug discovery. *Nat. Mach. Intell.* **2022**, *4*, 211–221. [CrossRef]

13. Kim, J.; Park, S.; Min, D.; Kim, W. Comprehensive Survey of Recent Drug Discovery Using Deep Learning. *Int. J. Mol. Sci.* **2021**, *22*, 9983. [CrossRef] [PubMed]

14. Cortés-Ciriano, I.; Ain, Q.U.; Subramanian, V.; Lenselink, E.B.; Méndez-Lucio, O.; IJzerman, A.P.; Wohlfahrt, G.; Prusis, P.; Malliavin, T.E.; van Westen, G.J. Polypharmacology modelling using proteochemometrics (PCM): Recent methodological developments, applications to target families, and future prospects. *MedChemComm* **2015**, *6*, 24–50. [CrossRef]

15. Gao, K.; Nguyen, D.D.; Sresht, V.; Mathiowetz, A.M.; Tu, M.; Wei, G.-W. Are 2D fingerprints still valuable for drug discovery? *Phys. Chem. Chem. Phys.* **2020**, *22*, 8373–8390. [CrossRef]

16. Tetko, I.V.; Karpov, P.; Van Deursen, R.; Godin, G. State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis. *Nat. Commun.* **2020**, *11*, 5575. [CrossRef]

17. Wen, M.; Zhang, Z.; Niu, S.; Sha, H.; Yang, R.; Yun, Y.; Lu, H. Deep-learning-based drug–target interaction prediction. *J. Proteome Res.* **2017**, *16*, 1401–1409. [CrossRef]

18. Tian, Q.; Bilgic, B.; Fan, Q.; Liao, C.; Ngamsombat, C.; Hu, Y.; Witzel, T.; Setsompop, K.; Polimeni, J.R.; Huang, S.Y. DeepDTI: High-fidelity six-direction diffusion tensor imaging using deep learning. *NeuroImage* **2020**, *219*, 117017. [CrossRef]

19. Öztürk, H.; Özgür, A.; Ozkirimli, E. DeepDTA: Deep drug–target binding affinity prediction. *Bioinformatics* **2018**, *34*, i821–i829. [CrossRef]

20. Lee, I.; Keum, J.; Nam, H. DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. *PLoS Comput. Biol.* **2019**, *15*, e1007129. [CrossRef]

21. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R. QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **2014**, *57*, 4977–5010. [CrossRef] [PubMed]

22. Hilpert, K.; Fjell, C.D.; Cherkasov, A. Short linear cationic antimicrobial peptides: Screening, optimizing, and prediction. In *Peptide-Based Drug Design*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 127–159.

23. Cherkasov, A.; Shi, Z.; Fallahi, M.; Hammond, G.L. Successful in silico discovery of novel nonsteroidal ligands for human sex hormone binding globulin. *J. Med. Chem.* **2005**, *48*, 3203–3213. [CrossRef] [PubMed]

24. Cherkasov, A. Inductive QSAR descriptors. Distinguishing compounds with antibacterial activity by artificial neural networks. *Int. J. Mol. Sci.* **2005**, *6*, 63–86. [CrossRef]

25. Liu, Z.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.; Nie, W.; Liu, Y.; Wang, R. PDB-wide collection of binding data: Current status of the PDBbind database. *Bioinformatics* **2015**, *31*, 405–412. [CrossRef]

26. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef]

27. Jiménez, J.; Skalic, M.; Martinez-Rosell, G.; De Fabritiis, G. K deep: Protein–ligand absolute binding affinity prediction via 3d-convolutional neural networks. *J. Chem. Inf. Model.* **2018**, *58*, 287–296. [CrossRef]

28. Stepniewska-Dziubinska, M.M.; Zielenkiewicz, P.; Siedlecki, P. Development and evaluation of a deep learning model for protein–ligand binding affinity prediction. *Bioinformatics* **2018**, *34*, 3666–3674. [CrossRef]

29. Lim, J.; Ryu, S.; Park, K.; Choe, Y.J.; Ham, J.; Kim, W.Y. Predicting Drug–Target Interaction Using a Novel Graph Neural Network with 3D Structure-Embedded Graph Representation. *J. Chem. Inf. Model.* **2019**, *59*, 3981–3988. [CrossRef]

30. Nguyen, T.M.; Nguyen, T.; Le, T.M.; Tran, T. GEFA: Early fusion approach in drug-target affinity prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**, *19*, 718–728. [CrossRef]

31. Son, J.; Kim, D. Development of a graph convolutional neural network model for efficient prediction of protein-ligand binding affinities. *PLoS ONE* **2021**, *16*, e0249404. [CrossRef]

32. 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. [CrossRef] [PubMed]

33. Tang, J.; Szwajda, A.; Shakyawar, S.; Xu, T.; Hintsanen, P.; Wennerberg, K.; Aittokallio, T. Making sense of large-scale kinase inhibitor bioactivity data sets: A comparative and integrative analysis. *J. Chem. Inf. Model.* **2014**, *54*, 735–743. [CrossRef] [PubMed]

34. Leckband, D.; Israelachvili, J.; Schmitt, F.; Knoll, W. Long-range attraction and molecular rearrangements in receptor-ligand interactions. *Science* **1992**, *255*, 1419–1421. [CrossRef] [PubMed]
35. Dunbar Jr, J.B.; Smith, R.D.; Yang, C.-Y.; Ung, P.M.-U.; Lexa, K.W.; Khazanov, N.A.; Stuckey, J.A.; Wang, S.; Carlson, H.A. CSAR benchmark exercise of 2010: Selection of the protein–ligand complexes. *J. Chem. Inf. Model.* **2011**, *51*, 2036–2046. [CrossRef]
36. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
37. Davis, M.I.; Hunt, J.P.; Herrgard, S.; Ciceri, P.; Wodicka, L.M.; Pallares, G.; Hocker, M.; Treiber, D.K.; Zarrinkar, P.P. Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2011**, *29*, 1046–1051. [CrossRef]
38. Sussman, J.L.; Lin, D.; Jiang, J.; Manning, N.O.; Prilusky, J.; Ritter, O.; Abola, E.E. Protein Data Bank (PDB): Database of three-dimensional structural information of biological macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1998**, *54*, 1078–1084. [CrossRef]
39. Gilson, M.K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res.* **2016**, *44*, D1045–D1053. [CrossRef]
40. Wang, R.; Fang, X.; Lu, Y.; Yang, C.-Y.; Wang, S. The PDBbind database: Methodologies and updates. *J. Med. Chem.* **2005**, *48*, 4111–4119. [CrossRef]
41. Shim, J.; Hong, Z.-Y.; Sohn, I.; Hwang, C. Prediction of drug–target binding affinity using similarity-based convolutional neural network. *Sci. Rep.* **2021**, *11*, 4416. [CrossRef]
42. He, T.; Heidemeyer, M.; Ban, F.; Cherkasov, A.; Ester, M. SimBoost: A read-across approach for predicting drug–target binding affinities using gradient boosting machines. *J. Cheminform.* **2017**, *9*, 24. [CrossRef] [PubMed]
43. Kouranov, A.; Xie, L.; de la Cruz, J.; Chen, L.; Westbrook, J.; Bourne, P.E.; Berman, H.M. The RCSB PDB information portal for structural genomics. *Nucleic Acids Res.* **2006**, *34* (Suppl. 1), D302–D305. [CrossRef] [PubMed]
44. Tokars, V.; Mesecar, A. QFRET-Based Primary Biochemical High Throughput Screening Assay to Identify Inhibitors of the SARS Coronavirus 3C-Like Protease (3CLPro). Available online: https://pubchem.ncbi.nlm.nih.gov/bioassay/1706 (accessed on 1 July 2022).
45. Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* **2005**, *3*, e324.
46. Yang, K.K.; Wu, Z.; Bedbrook, C.N.; Arnold, F.H. Learned protein embeddings for machine learning. *Bioinformatics* **2018**, *34*, 2642–2648. [CrossRef] [PubMed]
47. Duvaud, S.; Gabella, C.; Lisacek, F.; Stockinger, H.; Ioannidis, V.; Durinx, C. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res.* **2021**, *49*, W216–W227. [CrossRef] [PubMed]
48. Jamasb, A.R.; Lió, P.; Blundell, T.L. Graphein-a python library for geometric deep learning and network analysis on protein structures. *bioRxiv* **2020**. [CrossRef]
49. Kearnes, S.; McCloskey, K.; Berndl, M.; Pande, V.; Riley, P. Molecular graph convolutions: Moving beyond fingerprints. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 595–608. [CrossRef]
50. Rives, A.; Meier, J.; Sercu, T.; Goyal, S.; Lin, Z.; Liu, J.; Guo, D.; Ott, M.; Zitnick, C.L.; Ma, J. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2016239118. [CrossRef]
51. Nguyen, T.; Le, H.; Venkatesh, S. GraphDTA: Prediction of drug–target binding affinity using graph convolutional networks. *bioRxiv* **2019**, 684662. [CrossRef]
52. Nascimento, A.C.; Prudêncio, R.B.; Costa, I.G. A multiple kernel learning algorithm for drug-target interaction prediction. *BMC Bioinform.* **2016**, *17*, 46. [CrossRef]
53. Zhao, L.; Wang, J.; Pang, L.; Liu, Y.; Zhang, J. GANsDTA: Predicting drug-target binding affinity using GANs. *Front. Genet.* **2020**, *10*, 1243. [CrossRef] [PubMed]
54. Karimi, M.; Wu, D.; Wang, Z.; Shen, Y. DeepAffinity: Interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics* **2019**, *35*, 3329–3338. [CrossRef] [PubMed]
55. Huang, K.; Fu, T.; Glass, L.M.; Zitnik, M.; Xiao, C.; Sun, J. DeepPurpose: A deep learning library for drug–target interaction prediction. *Bioinformatics* **2020**, *36*, 5545–5547. [CrossRef] [PubMed]
56. Li, S.; Zhou, J.; Xu, T.; Huang, L.; Wang, F.; Xiong, H.; Huang, W.; Dou, D.; Xiong, H. Structure-aware interactive graph neural networks for the prediction of protein-ligand binding affinity. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Singapore, 14–18 August 2021; pp. 975–985.
57. Danel, T.; Spurek, P.; Tabor, J.; Śmieja, M.; Struski, Ł.; Słowik, A.; Maziarka, Ł. Spatial graph convolutional networks. In *International Conference on Neural Information Processing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 668–675.
58. Gentile, F.; Fernandez, M.; Ban, F.; Ton, A.-T.; Mslati, H.; Perez, C.F.; Leblanc, E.; Yaacoub, J.C.; Gleave, J.; Stern, A. Automated discovery of noncovalent inhibitors of SARS-CoV-2 main protease by consensus Deep Docking of 40 billion small molecules. *Chem. Sci.* **2021**, *12*, 15960–15974. [CrossRef]
59. Mesecar, A. A taxonomically-driven approach to development of potent, broad-spectrum inhibitors of coronavirus main protease including SARS-CoV-2 (COVID-19). *Be Publ.* **2020**.
60. *Schrödinger Release 2020–4: Glide*; Schrödinger, LLC: New York, NY, USA, 2020.

61. Ietswaart, R.; Arat, S.; Chen, A.X.; Farahmand, S.; Kim, B.; DuMouchel, W.; Armstrong, D.; Fekete, A.; Sutherland, J.J.; Urban, L. Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology. *EBioMedicine* **2020**, *57*, 102837. [CrossRef]

62. Stank, A.; Kokh, D.B.; Fuller, J.C.; Wade, R.C. Protein binding pocket dynamics. *Acc. Chem. Res.* **2016**, *49*, 809–815. [CrossRef]

63. Liu, T.; Altman, R. Identifying druggable targets by protein microenvironments matching: Application to transcription factors. *CPT Pharmacomet. Syst. Pharmacol.* **2014**, *3*, 1–9. [CrossRef]