POCKETNET: LIGAND-GUIDED POCKET PREDICTION FOR BLIND DOCKING

Matthew R. Masters, Amr H. Mahmoud, Markus A. Lill
Department of Pharmaceutical Sciences
University of Basel
Klingelbergstrasse 50, 4056
Basel, Switzerland
{matthew.masters, amr.abdallah, markus.lill}@unibas.ch

ABSTRACT

We introduce PocketNet, a novel method for identifying ligand binding sites (LBS). Unlike current methods, PocketNet is tailored to identify the binding site specifically associated with the target ligand. With most protein targets having multiple binding sites, the selection process becomes ambiguous without specific ligand information. This limitation negatively impacts downstream applications such as docking and virtual screening. PocketNet addresses this challenge by combining the ouput of multiple LBS prediction tools and utilizing a deep neural network that incorporates ligand information to re-rank the sites. Our results demonstrate that PocketNet outperforms the latest methods for both pocket prediction and blind docking tasks.

1 INTRODUCTION

Protein-ligand binding is essential for life and plays a critical role in many therapeutic interventions. Understanding how and where a ligand binds to a particular target is central to structure-based drug design (Pérot et al., 2010), virtual screening (Sunseri & Koes, 2021), and adverse effect prediction (Xie et al., 2011). Analysis of the Protein Data Bank (Berman et al., 2002) has shown that ligand binding sites (LBSs) are typically found within the largest and deepest pocket of the protein surface (Laskowski et al., 1996; Sotriffer & Klebe, 2002). However it is also common for large ligands to bind to shallow pockets on the solvent exposed surface (Stern & Wiley, 1994; Nisius et al., 2012).

Most proteins also contain multiple binding sites and can bind a diverse set of ligand structures (Ludlow et al., 2015). These sites can be broadly classified into two categories: orthosteric and allosteric. An orthosteric binding site is involved in the function of the protein and often binds to endogenous compounds or larger biomolecular structures. A protein usually has a single orthosteric binding site, also called the active site. A typical goal of structure-based drug design is to find an inhibitor which competes against the endogenous substrate for the active site. However, there is a rising interest in targeting allosteric sites which have often been overlooked in drug discovery projects of the past (Grover, 2013; Wagner et al., 2016). Allosteric sites are located elsewhere on the protein but can still be targeted using small molecules to modulate protein function. Moreover, while existing medications often target a single well-characterized pocket, it is estimated that over 50% of drugs interact with more than five different targets, leading to possible adverse effects (Wishart et al., 2008). For these reasons, there is a large interest in the identification and characterization of all possible ligand binding sites (Kufareva et al., 2012; Desaphy et al., 2015).

To this end, computational methods have been introduced which predict binding sites given a protein structure (Zhao et al., 2020). While these methods have been used for decades, the problem of identifying binding sites is still not fully solved. One limitation of these methods is that they only rely upon the protein structure which can contain multiple pockets, ranging from deep buried sites to shallow exposed sites. The main determinant in which site a ligand binds to is the ligand structure itself. Therefore, it makes sense to condition the prediction of binding sites with a ligand structure. Such a tool would aid in the discovery of both on- and off-target binding of drugs, especially for pockets which haven't been previously investigated. With the introduction of AlphaFold 2 and the

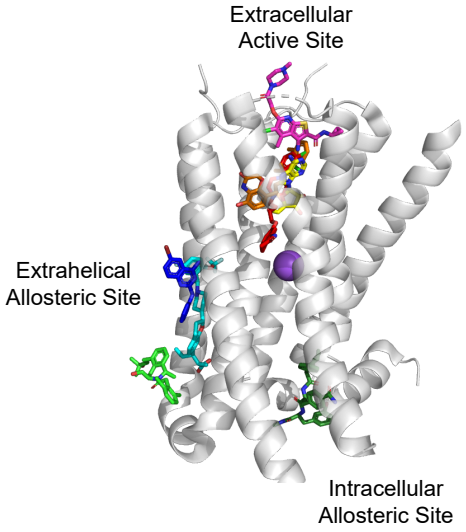

Figure 1: Example of a target with multiple binding sites targeted by different ligands. GPCRs generally feature an extracellular active site, extrahelical allosteric sites, and an intracellular allosteric site. Structure of GPCR Dopamine Receptor D1 is shown as gray (PDB ID: 7LJD) and ligands are shown as various colors (PDB IDs: 5TZY, 7LD3, 4MQT, 5X7D, 6OBA, 6N48, 7VOE).

associated proteome database, there is a plethora of new protein structures and new pockets which have yet to be explored (Jumper et al., 2021; Tunyasuvunakool et al., 2021).

This study aims to overcome this limitation by introducing PocketNet, a new method for LBS identification which is guided by a ligand structure in order to isolate a single pocket of interest. This is accomplished by constructing an ensemble of pocket predictions using existing LBS prediction tools followed by a re-ranking procedure with a deep neural network. The neural network is provided both pocket and ligand features and is trained to predict a score which can be used to re-rank the ensemble. This method surpasses the accuracy of existing LBS prediction tools and has the novel ability to differentiate pockets based on ligand structure. Furthermore, in order to demonstrate the applicability of PocketNet for downstream use-cases, we use its predictions for the task of blind docking where a ligand is docked to a protein without knowledge of the correct binding site. Again PocketNet shows superior performance to existing blind docking methods.

1.1 Related Work

Pocket prediction tools were introduced as early as 1985 and numerous diverse methods have been developed since (Zhao et al., 2020). The seminal works mostly used spatial geometry in order to locate accessible cavities of the protein. For example, in LIGSITE (Hendlich et al., 1997) a 3D grid is superimposed over the protein and at each point they look for intersections with the protein in 7 different directions (x, y, z axes and diagonals). If both ends intersect, then the point is assumed to be inside a pocket and its value is increased. Binding site residues are each assigned a score based on their nearby grid points to get the final pocket prediction. Other geometry-based methods include Fpocket (Le Guilloux et al., 2009), Patch-Surfer2.0 (Zhu et al., 2015), and CurPocket (Liu et al., 2020).

In addition to these geometry-based approaches, several works use an energy-based approach such as QSiteFinder (Laurie & Jackson, 2005), SITEHOUND (Hernandez et al., 2009; Ghersi & Sanchez, 2009), FTSite (Ngan et al., 2012), and SiteComp (Lin et al., 2012). In this method, the system is parameterized using a molecular force field and probe atoms are used to calculate interaction forces at each point in a grid. The grid is further processed to remove points which are associated with energetically unfavorable interactions. The remaining points are then clustered to generate the final pocket prediction.

Another class of LBS prediction methods relies on templates from known protein-ligand complexes. These template-based methods can be further divided into structure-based and sequence-based depending on whether the templates are based on 3D structures or sequence information. ConSurf (Glaser et al., 2003), FINDSITE (Brylinski & Skolnick, 2008), and S-SITE/TM-SITE (Yang et al., 2013) are just a few of these template-based methods.

The last class of LBS prediction methods employ machine learning algorithms in a variety of ways. For example, PRANK (Krivák & Hoksza, 2015a;b) provides physical descriptors to a random forest model which predicts a score to rank putative pockets. DEEPSite (Jiménez et al., 2017) and DeepCSeqSite (Cui et al., 2019) are two methods which utilize deep neural networks for their predictions. DEEPSite encodes the pocket as 3D atom density grids which it provides as input to a 3D convolutional neural network (CNN) for scoring. DeepCSeqSite uses a different approach by encoding residue features in sequential order which it provides as input to a 1D CNN.

Molecular docking is a widely-used computational technique to predict the binding pose of a ligand within a protein pocket. Since many drug discovery projects target well-known pockets, focused docking can be used to restrict the search to that area. However, there is also blind docking which predicts poses without knowledge of existing pockets. There are several approaches to perform blind docking. The simplest is to use a traditional docking program with a search box that encloses the full protein. Another approach is to predict likely sites using a pocket prediction tool and then to perform focused docking (Liu et al., 2020). Finally, there have been some recent developments which aim to solve blind docking using deep learning models. In particular, EquiBind (Stärk et al., 2022), TANKBind (Lu et al., 2022), and DiffDock (Corso et al., 2022) each employ unique deep neural network approaches for blind docking.

2 Materials and Methods

2.1 Pocket Ensemble

PocketNet leverages existing pocket prediction tools by creating an ensemble of LBS boxes from multiple programs. Since pocket prediction tools are relatively fast to run (i.e. generally faster than docking) and are highly diverse in terms of their detection method, they are well-suited for ensembling. Ideally this ensemble covers the full extent of druggable pockets within the protein. Each box is axis-aligned and defined by its centroid and size along the x, y, and z axes. The predictions of each program are converted into this standard representation when necessary. This enables the use of most LBS prediction tools when constructing the ensemble. Additionally, a buffer of 8 Å is added to the size of each box in order to fully enclose more ligands. In this study, four of these tools were utilized with further details described below.

SiteMap (Halgren, 2007; 2009) is a commercial software developed by Schrödinger which predicts binding sites based on physical descriptors and a druggability score. Since SiteMap results are given as volume points which occupy each predicted pocket, the points were converted to the standard box format. The geometric center of points defines the box center and the minimum and maximum coordinates define the extent of the box so that all points are contained inside. SiteMap version 4.9.012 was used with default settings.

P2Rank (Krivák & Hoksza, 2018) is an open-source software which is based on the PRANK algorithm (Krivák & Hoksza, 2015a;b). The algorithm uses physical descriptors which are provided to a random forest model which is then used to rank potential sites. Like SiteMap, the results are provided as a set of points so the same procedure was used to convert each pocket. P2Rank version 2.4 was used with default settings.

FPocket (Le Guilloux et al., 2009) is another open-source software which is based on Voronoi tessellation and simple geometric rules. While generally less accurate than P2Rank, FPocket is considerably faster and can be run on very large datasets. As with the two previous methods, the results were converted from points to boxes. FPocket version 4.0 was used with default settings.

CurPocket (Liu et al., 2020) is the newest of these programs and also utilizes its LBS predictions for docking in CB-Dock. CurPocket is also fast to run on large datasets since it is based on simple geometric rules which uses the curvature of the protein surface to detect pockets. The results are

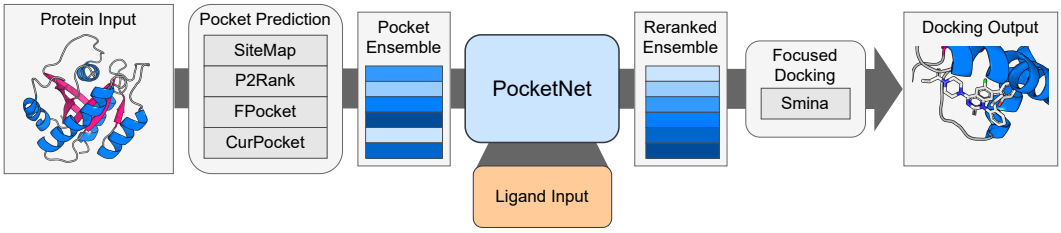

Figure 2: Overview of PocketNet blind docking workflow. Starting with a protein structure input, different pocket prediction tools are used to generate an ensemble of potential pockets. Geometric and chemical descriptors are then extracted from each predicted pocket and the ligand input. Then the PocketNet model scores each ligand-pocket pair to produce the reranked ensemble. The top pockets can then be used for focused docking with Smina.

provided in the desired format so no conversion was necessary. CurPocket version 1.5 was used with default settings.

While these pocket prediction tools often rely upon similar underlying descriptors, they each perform the best on different targets. Constructing an ensemble with all four programs achieves better coverage than an ensemble of any three (see Supporting Information Table 1).

2.2 POCKETNET

Once an ensemble of pockets has been constructed, a scoring function must be used to rerank the ensemble. To this end we introduce PocketNet, a deep neural network which reranks the pocket ensemble conditioned by a ligand structure (see Figure 2). The architecture of PocketNet is a hybrid graph neural network (GNN) which incorporates both graph and tabular features.

2.2.1 FEATURIZATION

Predicted pockets are encoded into a graph by extracting the residues inside the box and treating each residue as a node in the graph. Each node of the pocket graph contains features describing the residue type and chemical environment. It is possible to use all atoms within the predicted box, but we opted to use this residue-level coarse-grained approach for simplicity and reduce sensitivity towards small-scale conformational variations of the residues' side chains. The graph is locally connected with edges bridging nodes within 8Å of each other and uses the distance as an edge feature. Ligands are also encoded into a graph but include all heavy atoms and edges are assigned based on covalent bonds.

In addition to these graph features, several handcrafted tabular features are also used by the model. For the pocket, this includes geometrical features (e.g. box length-width-height ratio, volume, protein surface area), chemical descriptors (e.g. atom density), measurements of a pockets agreement with the rest of the ensemble (e.g. overlap with other boxes), and the original rank given by the pocket prediction tool. The ligand was featurized using mostly chemical descriptors (e.g. weight, rotatable bonds, H-bond acceptors/donors, polar surface area, etc.). A complete description of the graph and tabular features used for both pocket and ligand can be found in Supporting Information Table 2. Samples are provided to the model in the form of ligand-pocket pairs, with associated target variables discussed in Section 2.4.

2.2.2 MODEL

The PocketNet model takes the two graphs (pocket and ligand) and tabular features as input and output a single target variable. To accomplish this, we adopt a simple hybrid architecture which combines two graph neural networks and a series of feed-forward layers to produce the final prediction (see Supporting Information Figure 5).

The GNN layers are based on the graph attentional layer introduced by Veličković et al. (2017). Given a graph $\mathcal{G} = (\mathcal{X}, \mathcal{E})$ with node features $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}, \mathbf{x}_i \in \mathbb{R}^f$ and edge features

$\mathcal{E} = \{\mathbf{e}_{1,1}, \mathbf{e}_{1,2}, \ldots, \mathbf{e}_{N,N}\}, \mathbf{e}_{i,j} \in \mathbb{R}^d$ a GNN layer can be defined by an equation which updates $\mathcal{X}$ in successive layers. In the case of PocketNet, we are given two graphs $\mathcal{G}_p$ and $\mathcal{G}_l$ for the pocket and ligand respectively. The size of the feature space f in the input layer is equal to the number of input features for each graph. Using the parameters from our best model, f is then set to 256 for three hidden layers and 512 for the output layer. The edge features $\mathcal{E}$ are not changed by the GNN layer and thus the edge feature space d does not change from the input size.

The GNN layer update can be formalized by Equation 1 where Θ is a shared linear transformation applied to each node $\mathbf{x}_i$, and $\alpha_{i,j}$ is the attention coefficient between node i and node j. The attention coefficient $\alpha_{i,j}$ can be calculated according to Equation 2 where Θ_e is another shared linear transformation acting on each edge $\mathbf{e}_{i,j}$, $\mathbf{a}$ are the learnable attention parameters, $\cdot^T$ is the transposition operation, $||$ is the concatenation operation, and $\mathrm{LeakyReLU}$ is a non-linear activation function which scales down negative values. Through each GNN layer, the nodes gather more information about the collective graph features which can be used to predict the target variable.

$$\mathbf{x}'_i = \alpha_{i,i}\Theta\mathbf{x}_i + \sum_{j \in N} \alpha_{i,j}\Theta\mathbf{x}_j \tag{1}$$

$$\alpha_{i,j} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^\top \left[\mathbf{\Theta}\mathbf{x}_i \,\|\mathbf{\Theta}\mathbf{x}_j\| \,\mathbf{\Theta}_e\mathbf{e}_{i,j}\right]\right)\right)}{\sum_{k \in N} \exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^\top \left[\mathbf{\Theta}\mathbf{x}_i \,\|\mathbf{\Theta}\mathbf{x}_k\| \,\mathbf{\Theta}_e\mathbf{e}_{i,k}\right]\right)\right)} \tag{2}$$

Since these graph layers only learn 3D information through the edge distances, they have the property of being invariant to any rotation or translation of the pocket or ligand coordinates. This is desirable since these structures have no inherent orientation and the scoring of a ligand-pocket pair should not change depending on its orientation. For this reason, most recent studies applying GNNs to molecular structures require that the network is invariant (or equivariant) to these transformations (Jing et al., 2021; Satorras et al., 2021; Batzner et al., 2022).

Following the pocket and ligand GNN layers, we are provided with the node embeddings $\mathcal{X}'_p$ and $\mathcal{X}'_l$. In order to obtain fixed-size embeddings which represent the full pocket or ligand, each set of node embeddings is pooled by summation to form a single feature vector $\hat{\mathbf{x}} \in \mathbb{R}^f$ (here $f = 512$). These two feature vectors are concatenated together with the tabular features, denoted $\mathbf{x}_{tab}$, to form the complete ligand-pocket pair embedding. To compute the prediction $\hat{y}$ we introduce a feed-forward network ψ which takes the pair embedding and produces a scalar output:

$$\hat{y} = \sigma\left(\psi\left(\left[\hat{\mathbf{x}}_p \;||\; \hat{\mathbf{x}}_l \;||\; \mathbf{x}_{tab}\right]\right)\right) \tag{3}$$

The network is coupled with an activation function σ which is selected based on the target variable. Since there are multiple ways to evaluate how well a proposed pocket fits the native pose, we explored three different target variables during training: success, IoU, and centroid distance (see Section 2.4). As a result of our experiments, centroid distance was found to be the best target variable. Thus, σ was set to be the ReLU activation function to ensure the prediction is non-negative. The model was trained using an appropriate loss function for the given target: binary cross-entropy (BCE) for success and mean squared error (MSE) for IoU and distance. Additional training details can be found in the Supporting Information Section 4.4.

Once the model has been trained, it can be used for inference to score ligand-pocket pairs which can be used to re-rank the ensemble for each target. When the model is trained by distance, then the ensemble is ranked in ascending order such that the closest predicted pockets are ranked on top. Finally, these top-ranked pockets can be used for downstream tasks such as focused docking.

2.3 DOCKING

Once the pocket ensemble is re-ranked using PocketNet, a traditional docking program can be used for focused docking of the top-ranked pockets. To this end, we used Smina (Koes et al., 2013), a fork of AutoDock Vina with an improved scoring function. It has been widely used in docking studies (Xu et al., 2020; Gorgulla et al., 2020) and as a baseline for evaluating docking programs (Corso et al., 2022; Masters et al., 2022). Smina uses a Monte Carlo search algorithm and empirical scoring function which incorporates both force-field and knowledge-based terms. All default parameters were used except for the number of modes which was set to be 10. This same docking procedure was used for various other baselines described further in Section 2.6.

2.4 METRICS

To evaluate the pocket prediction task we have the known ligand pose and our ranked ensemble of predicted boxes. From this we define three metrics: success, intersection over union (IoU), and centroid distance. Success is a binary metric which indicates if the box (including added buffer) fully encloses the native ligand pose. IoU is a continuous metric between 0 and 1 which indicates the overlap between the predicted box and the bounding box of the ligand. It is defined as the intersection volume of the two boxes divided by their union volume which is trivial to calculate using the axis-aligned box definition. A value of 0 indicates there is no overlap between the two boxes while a value of 1 indicates perfect overlap. It may be a better metric than success for pocket prediction since it penalizes boxes that are unnecessarily large. Centroid distance is the final metric for pocket prediction. It is calculated as the Euclidean distance between the centroid of the predicted box and the bounding box of the ligand. When presented in the results, these metrics are shown as averages over the test set unless otherwise noted.

To evaluate the blind docking task, the top-ranked ligand pose with the lowest predicted energy is always used to compare against the native pose. As with pocket prediction, there are numerous metrics which can be used to assess docking performance. The predominant metric for pose quality, called root-mean-square deviation (RMSD), quantifies how close a predicted pose is to the native pose. However, some care should be taken with its calculation; many ligands have structural symmetry which should be taken into account during the calculation to prevent artificially higher RMSD values. To this end, sPyRMSD (Meli & Biggin, 2020) was used for symmetry-aware RMSD calculations. Docking success is another binary metric which simply indicates whether the RMSD falls below a certain threshold value, typically 2.0 - 5.0Å. Additional details on the metrics used can be found in Section 4.5 of the Supporting Information.

2.5 DATASET

This study utilizes the PDBbind dataset (Liu et al., 2017), a collection of nearly 20k protein-ligand complexes which has been widely used for training and benchmarking new methods. The dataset also contains experimentally measured binding affinity data for each complex, although affinity prediction was not assessed within this study. Several different dataset splitting strategies have been used in the context of machine learning with molecules, however there is still no consensus on the best method (Wu et al., 2018; Volkov et al., 2022). This study utilizes a time-based split which trains on older complexes (deposited from 1982 - 2018) and tests on new data (2019). This split is consistent with previous works on blind docking and allows for straightforward comparison of results with the baseline methods (Stärk et al., 2022; Lu et al., 2022; Corso et al., 2022).

2.6 BASELINES

Several baselines were established for the pocket prediction and blind docking tasks to compare against PocketNet. For both tasks, the four pocket prediction tools used in the ensembling process (SiteMap, P2Rank, FPocket, and CurPocket) were used as baselines. In order to compare these tools in terms of downstream docking performance, poses were generated by Smina as described in Section 2.3. These baselines maintain the original ranking of boxes given by each tool. Furthermore, we employ three additional benchmarks for the blind docking task: Blind, EquiBind, and DiffDock.

Blind is a baseline where the full protein structure is used for docking with Smina. The search box is set to enclose the entire protein plus the additional 8Å buffer. All other settings are kept consistent with the other Smina docking baselines.

EquiBind (Stärk et al., 2022) is one of the first deep learning models to directly predict the pose of protein-ligand complexes. The model aims to predict the pose directly in one-shot, bypassing the need for a search algorithm entirely. Pre-trained weights with recommended hyperparameters were used for inference.

DiffDock (Corso et al., 2022) is another recent deep learning model which relies upon a diffusion model which gradually moves the pose from a random starting point. While inference is significantly longer than EquiBind, DiffDock showed state-of-the-art performance on this test set. Pre-trained weights with recommended hyperparameters were used for inference and 40 poses were generated for each system, with the top-scoring pose being used for evaluation.

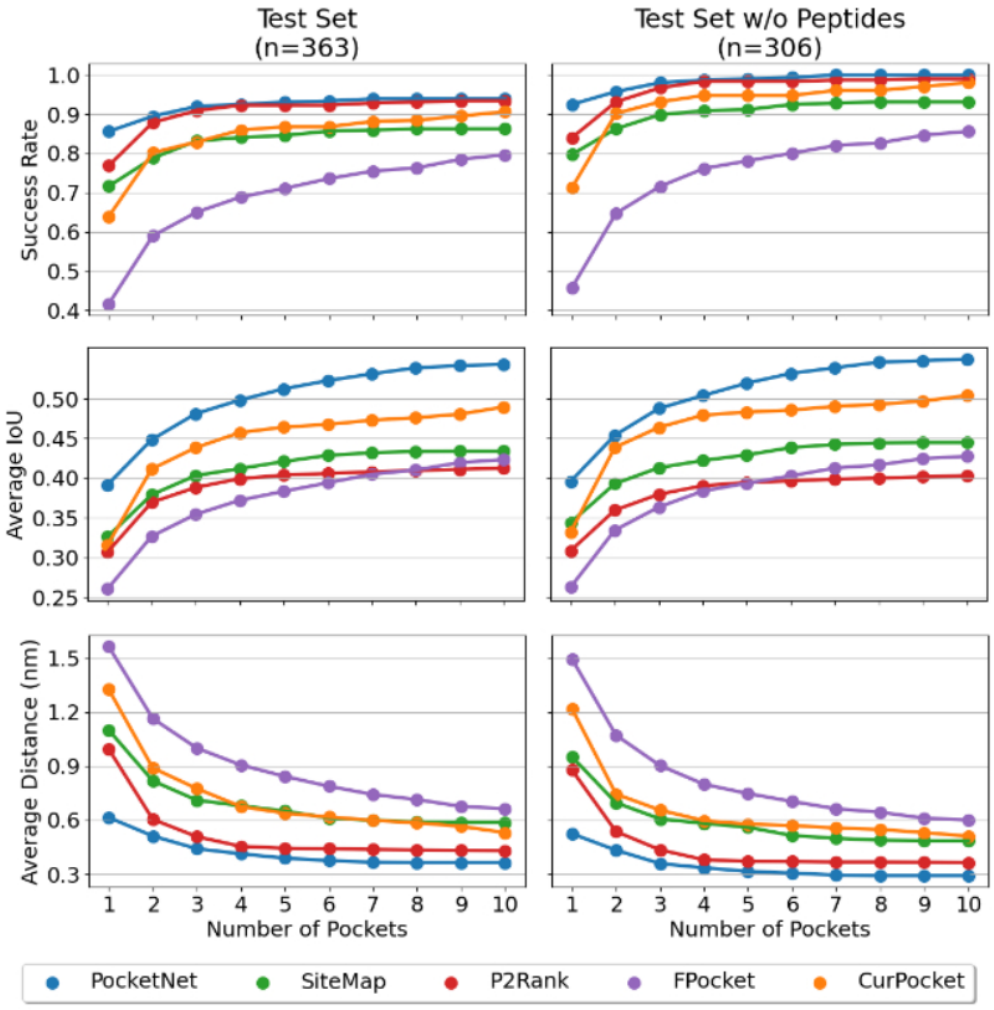

Figure 3: Pocket prediction task results including PocketNet and all pocket prediction baselines. Tools were evaluated by the three pocket predictions metrics introduced earlier (success, IoU, and cetroid distance) averaged over the test set. Due to the poor performance of existing pocket prediction tools for peptide binding sites, results are shown including and excluding these systems.

3 Results and Discussion

3.1 Pocket Prediction

Figure 3 shows the results for the pocket prediction task. PocketNet exceeds performance of all baseline tools, especially for the top-ranked pocket. In terms of the rate of success, PocketNet shows a 10-46% improvement over baseline methods. PocketNet also shows significant improvement in terms of average IoU and average centroid distance. However, the ranking is still not ideal since performance increases when more pockets are considered. Still, most of this gain in performance is seen within the top-3 pockets, with success nearing 93% (ideal ranking is 96%). Figure 3 also shows the results for the test set without peptides which nears 99% success within the top-3 pockets (ideal ranking is 100%). This gap in performance is due to the poor performance of existing pocket prediction tools for peptidic ligands. Typically these ligands bind to several shallow pockets on the surface which aren't scored highly or are subdivided into several different pockets.

3.2 BLIND DOCKING

Figure 4 shows the results for the blind docking task. As described in Section 2.6, we have seven total baselines to compare against our model. While EquiBind is able to reliably predict poses with RMSD < 20Å, the method does not achieve high success with RMSD < 2Å. It can also be seen that the strategy of blind docking to the full protein structure performs worse than using any one of the pocket prediction tools. These tools perform similarly to each other and their success is correlated with the pocket prediction metrics. For example, P2Rank exceeds the other tools in terms of centroid distance and performs similarly well at docking while FPocket performs the worse at pocket prediction and docking. The issue with all of these methods is that they often rank distant pockets highly and thus obtain very large RMSD values. The authors in Corso et al. (2022) provided a similar baseline strategy to the P2Rank baseline presented here. However, in their approach the box definition is set to be cubic and is not influenced by the length-width-height ratio of the binding site. Through our experiments, we found this negatively impacts docking performance and that using a box with different side lengths is superior. DiffDock itself was found to be the next best method for blind docking and produced very similar results to PocketNet. Nevertheless, PocketNet exceeded all baseline methods and achieved a new state-of-the-art performance for this test set.

3.3 DIFFERENTIAL POCKET PREDICTION

A unique advantage of PocketNet is its ability to differentiate pockets based on a ligand structure. To assess this capability, proteins which contain multiple binding sites with distinct ligands were selected for further analysis. The same pocket features were provided to PocketNet while swapping out the features for each active ligand. Visualizations of the pocket prediction and docking results for both systems can be seen in Supporting Information Figure 8. While PocketNet is able to predict the correct pocket in each case, the subsequent docking with Smina sometimes underperforms. This result highlights the limitation of traditional docking algorithms.

Androgen Receptor (AR) is a nuclear receptor which controls gene expression and is implicated in a number of diseases. AR contains a primary active site and several allosteric sites (Baek et al., 2006). The primary site binds testosterone and similar ligands with a steroid structure while the allosteric pockets bind various scaffolds. Therefore, two PDBs (2YLQ and 2PIX) were selected which contain ligands bound to both the active site and the allosteric Binding Function 3 (Bf3) site. When PocketNet is conditioned on the 2YLQ protein structure with testosterone, the primary active site is correctly ranked at the top. On the other hand, when the same protein structure is conditioned with the allosteric inhibitor (YLQ; Compound 2 from Lack et al. (2011)), the allosteric site is correctly ranked at the top. A similar result was seen when using PocketNet to predict the sites of dihydrotestosterone (DHT) and another Bf3 inhibitor, flufenamic acid (FLF) (Estébanez-Perpiñá et al., 2007). Top-1 RMSD values of 0.3, 0.4, 3.1, and 5.2Å were obtained for TES, DHT, YLQ, and FLF respectively.

Tyrosine-protein Kinase ABL1 (ABL) is an enzyme involved in cell differentiation, proliferation, and DNA repair. For these reasons, mutations in ABL are associated with cancers, especially Chronic Myelogenous Leukemia (CML). Therefore, numerous inhibitors have been designed to target ABL and downregulate its uncontrolled proliferative effect. The first successful inhibitor to target ABL, Imatinib, binds to and blocks the active ATP binding site. However, resistance to Imatinib has become a growing problem in the treatment of CML (Juan & Ong, 2012). Therefore, there has been an increased interest in targeting allosteric sites of ABL, such as the myristate pocket (Zhang et al., 2010; Jahnke et al., 2010; Yang et al., 2011). Here, we examined four PDBs (3K5V, 3PYY, 3MSS, and 3MS9) which contain both Imatinib (STI) and different allosteric inhibitors (STJ, 3YY, MS7, MS9) targeting the myristate pocket. In all four cases, PocketNet is able to differentiate between Imatinib and the allosteric inhibitors and predict correct pockets for each. Top-1 RMSD values of 11.9-12.1Å were obtained for Imatinib and 1.0-5.8Å for allosteric inhibitors.

3.4 LIMITATIONS

Here we will discuss some of the limitations of the PocketNet method and some opportunities for future development. The axis-aligned box definition utilized in this study is not ideal for fitting ligand binding sites which have no inherent orientation. Due to this definition, LBS predictions depend on the input orientation of the protein and IoU values are never optimal. Additionally,

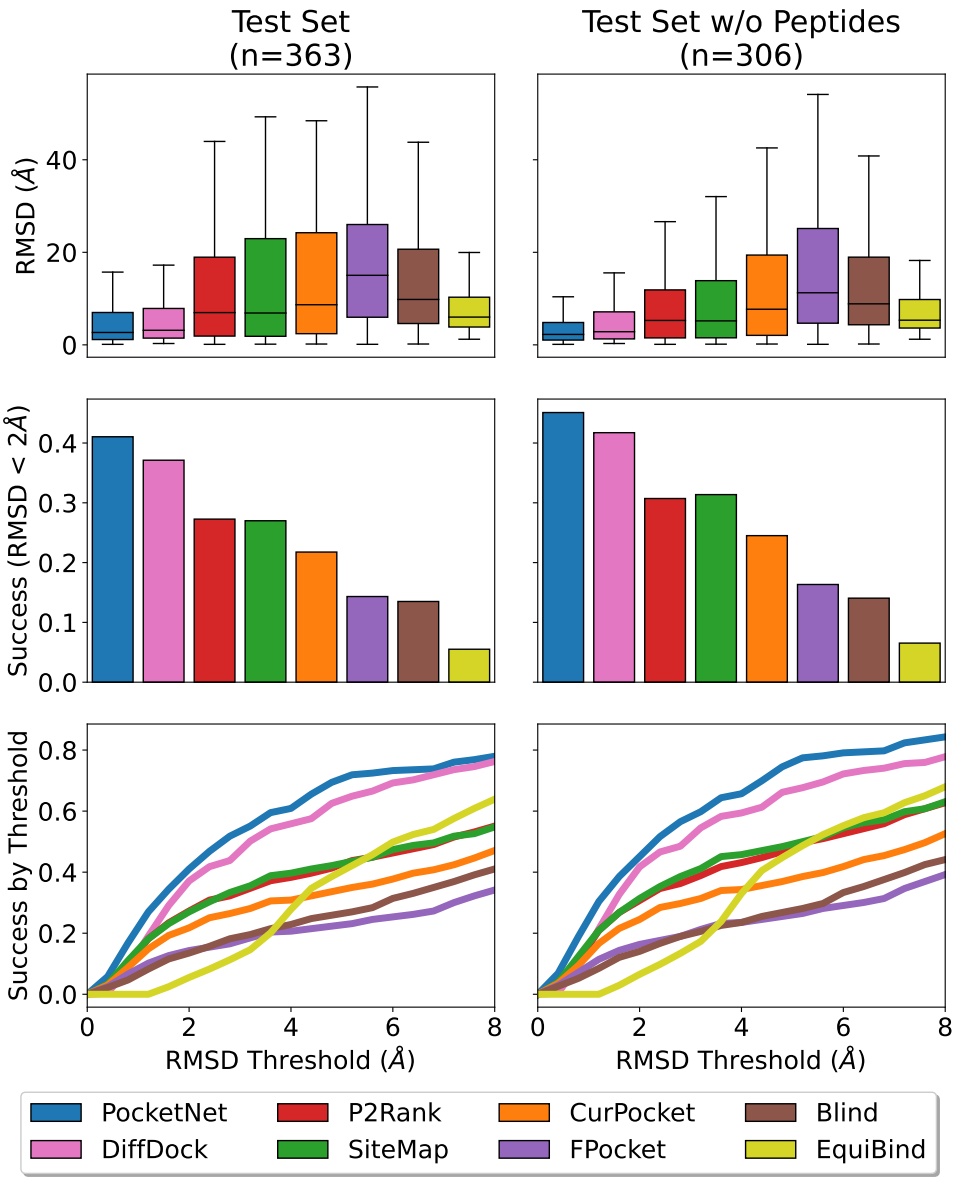

Figure 4: Blind docking task results including PocketNet and all blind docking baselines. Tools were evaluated by the docking metrics introduced earlier (RMSD and success) averaged over the test set.

since IoU is highly correlated with success of docking, we can assume this limitation also hinders downstream use of LBS predictions. This limitation is not simple to overcome because most LBS prediction and docking programs use this axis-aligned box definition.

Another limitation is our ensemble of pocket predictors. Still, success is limited to 96% due to poor performance of pocket prediction tools for peptidic ligands. No suitable tool could be found for this task. However, if a new tool emerges which performs accurately for these large ligands, it should be trivial to include the tool in the ensemble.

The final limitation we will discuss is the large gap between pocket prediction and docking accuracy. While PocketNet is successful over 86% of the time, the subsequent docking fails in more than half of these cases. Furthermore, in the differential pocket prediction analysis, PocketNet was able to predict the correct site in each examined structure but docking only succeeded in a third of cases. This high rate of failure can be attributed to issues with both the sampling and scoring of poses by Smina. Hence, by addressing the issues associated with focused docking we can narrow this gap and also improve blind docking results. Although there have been several recent developments which apply deep learning to molecular docking, each of these approaches aim to solve blind docking directly. Therefore, future research should investigate the application of these methods to focused docking and how to best incorporate the pocket prediction and docking tasks.

REFERENCES

Sung Hee Baek, Kenneth A Ohgi, Charles A Nelson, Derek Welsbie, Charlie Chen, Charles L Sawyers, David W Rose, and Michael G Rosenfeld. Ligand-specific allosteric regulation of coactivator functions of androgen receptor in prostate cancer cells. *Proceedings of the National Academy of Sciences*, 103(9):3100–3105, 2006.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1): 2453, 2022.

Helen M Berman, Tammy Battistuz, Talapady N Bhat, Wolfgang F Bluhm, Philip E Bourne, Kyle Burkhardt, Zukang Feng, Gary L Gilliland, Lisa Iype, Shri Jain, et al. The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, 58(6):899–907, 2002.

Michal Brylinski and Jeffrey Skolnick. A threading-based method (findsite) for ligand-binding site prediction and functional annotation. *Proceedings of the National Academy of sciences*, 105(1): 129–134, 2008.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv:2210.01776*, 2022.

Yifeng Cui, Qiwen Dong, Daocheng Hong, and Xikun Wang. Predicting protein-ligand binding residues with deep convolutional neural networks. *BMC bioinformatics*, 20(1):1–12, 2019.

Jérémy Desaphy, Guillaume Bret, Didier Rognan, and Esther Kellenberger. sc-pdb: a 3d-database of ligandable binding sites—10 years on. *Nucleic acids research*, 43(D1):D399–D404, 2015.

Eva Estébanez-Perpiñá, Leggy A Arnold, Phuong Nguyen, Edson Delgado Rodrigues, Ellena Mar, Raynard Bateman, Peter Pallai, Kevan M Shokat, John D Baxter, R Kiplin Guy, et al. A surface on the androgen receptor that allosterically regulates coactivator binding. *Proceedings of the National Academy of Sciences*, 104(41):16074–16079, 2007.

Dario Ghersi and Roberto Sanchez. Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites. *Proteins: Structure, Function, and Bioinformatics*, 74 (2):417–424, 2009.

Fabian Glaser, Tal Pupko, Inbal Paz, Rachel E Bell, Dalit Bechor-Shental, Eric Martz, and Nir Ben-Tal. Consurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics*, 19(1):163–164, 2003.

Christoph Gorgulla, Andras Boeszoermenyi, Zi-Fu Wang, Patrick D Fischer, Paul W Coote, Krishna M Padmanabha Das, Yehor S Malets, Dmytro S Radchenko, Yurii S Moroz, David A Scott, et al. An open-source drug discovery platform enables ultra-large virtual screens. *Nature*, 580 (7805):663–668, 2020.

Ashok Kumar Grover. Use of allosteric targets in the discovery of safer drugs. *Medical Principles and Practice*, 22(5):418–426, 2013.

Thomas A Halgren. Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling*, 49(2):377–389, 2009.

Tom Halgren. New method for fast and accurate binding-site identification and analysis. *Chemical biology & drug design*, 69(2):146–148, 2007.

Manfred Hendlich, Friedrich Rippmann, and Gerhard Barnickel. Ligsite: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of Molecular Graphics and Modelling*, 15(6):359–363, 1997.

Marylens Hernandez, Dario Ghersi, and Roberto Sanchez. Sitehound-web: a server for ligand binding site identification in protein structures. *Nucleic acids research*, 37(suppl_2):W413–W416, 2009.

Wolfgang Jahnke, Robert M Grotzfeld, Xavier Pelle, Andre Strauss, Gabriele Fendrich, Sandra W Cowan-Jacob, Simona Cotesta, Doriano Fabbro, Pascal Furet, Jurgen Mestan, et al. Binding or bending: distinction of allosteric abl kinase agonists from antagonists by an nmr-based conformational assay. *Journal of the American Chemical Society*, 132(20):7043–7048, 2010.

José Jiménez, Stefan Doerr, Gerard Martínez-Rosell, Alexander S Rose, and Gianni De Fabritiis. Deepsite: protein-binding site predictor using 3d-convolutional neural networks. *Bioinformatics*, 33(19):3036–3042, 2017.

Bowen Jing, Stephan Eismann, Pratham N Soni, and Ron O Dror. Equivariant graph neural networks for 3d macromolecular structure. *arXiv preprint arXiv:2106.03843*, 2021.

Wen Chun Juan and S Tiong Ong. The role of protein phosphorylation in therapy resistance and disease progression in chronic myelogenous leukemia. *Progress in molecular biology and translational science*, 106:107–142, 2012.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

David Ryan Koes, Matthew P Baumgartner, and Carlos J Camacho. Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. *Journal of chemical information and modeling*, 53(8):1893–1904, 2013.

Radoslav Krivák and David Hoksza. Improving protein-ligand binding site prediction accuracy by classification of inner pocket points using local features. *Journal of cheminformatics*, 7(1):1–13, 2015a.

Radoslav Krivák and David Hoksza. P2rank: Knowledge-based ligand binding site prediction using aggregated local features. In *Algorithms for Computational Biology: Second International Conference, AlCoB 2015, Mexico City, Mexico, August 4-5, 2015, Proceedings 2*, pp. 41–52. Springer, 2015b.

Radoslav Krivák and David Hoksza. P2rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *Journal of cheminformatics*, 10:1–12, 2018.

Irina Kufareva, Andrey V Ilatovskiy, and Ruben Abagyan. Pocketome: an encyclopedia of small-molecule binding sites in 4d. *Nucleic acids research*, 40(D1):D535–D540, 2012.

Nathan A Lack, Peter Axerio-Cilies, Peyman Tavassoli, Frank Q Han, Ka Hong Chan, Clementine Feau, Eric LeBlanc, Emma Tomlinson Guns, R Kiplin Guy, Paul S Rennie, et al. Targeting the binding function 3 (bf3) site of the human androgen receptor through virtual screening. *Journal of medicinal chemistry*, 54(24):8563–8573, 2011.

Roman A Laskowski, Nicholas M Luscombe, Mark B Swindells, and Janet M Thornton. Protein clefts in molecular recognition and function. *Protein science: a publication of the Protein Society*, 5(12):2438, 1996.

Alasdair TR Laurie and Richard M Jackson. Q-sitefinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics*, 21(9):1908–1916, 2005.

Vincent Le Guilloux, Peter Schmidtke, and Pierre Tuffery. Fpocket: an open source platform for ligand pocket detection. *BMC bioinformatics*, 10(1):1–11, 2009.

Yingjie Lin, Seungyeul Yoo, and Roberto Sanchez. Sitecomp: a server for ligand binding site analysis in protein structures. *Bioinformatics*, 28(8):1172–1173, 2012.

Yang Liu, Maximilian Grimm, Wen-tao Dai, Mu-chun Hou, Zhi-Xiong Xiao, and Yang Cao. Cb-dock: A web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacologica Sinica*, 41(1):138–144, 2020.

Zhihai Liu, Minyi Su, Li Han, Jie Liu, Qifan Yang, Yan Li, and Renxiao Wang. Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of chemical research*, 50 (2):302–309, 2017.

Wei Lu, Qifeng Wu, Jixian Zhang, Jiahua Rao, Chengtao Li, and Shuangjia Zheng. Tankbind: Trigonometry-aware neural networks for drug-protein binding structure prediction. *bioRxiv*, pp. 2022–06, 2022.

R Frederick Ludlow, Marcel L Verdonk, Harpreet K Saini, Ian J Tickle, and Harren Jhoti. Detection of secondary binding sites in proteins using fragment screening. *Proceedings of the National Academy of Sciences*, 112(52):15910–15915, 2015.

Matthew Masters, Amr H Mahmoud, Yao Wei, and Markus Alexander Lill. Deep learning model for flexible and efficient protein-ligand docking. In *ICLR2022 Machine Learning for Drug Discovery*, 2022.

Rocco Meli and Philip C. Biggin. spyrmsd: symmetry-corrected rmsd calculations in python. *Journal of Cheminformatics*, 12(1):49, 2020.

Chi-Ho Ngan, David R Hall, Brandon Zerbe, Laurie E Grove, Dima Kozakov, and Sandor Vajda. Ftsite: high accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics*, 28(2):286–287, 2012.

Britta Nisius, Fan Sha, and Holger Gohlke. Structure-based computational analysis of protein binding sites for function and druggability prediction. *Journal of biotechnology*, 159(3):123–134, 2012.

Stéphanie Pérot, Olivier Sperandio, Maria A Miteva, Anne-Claude Camproux, and Bruno O Villoutreix. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug discovery today*, 15(15-16):656–667, 2010.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Christoph Sotriffer and Gerhard Klebe. Identification and mapping of small-molecule binding sites in proteins: computational tools for structure-based drug design. *Il Farmaco*, 57(3):243–251, 2002.

Hannes Stärk, Octavian Ganea, Lagnajit Pattanaik, Regina Barzilay, and Tommi Jaakkola. Equibind: Geometric deep learning for drug binding structure prediction. In *International Conference on Machine Learning*, pp. 20503–20521. PMLR, 2022.

Lawrence J Stern and Don C Wiley. Antigenic peptide binding by class i and class ii histocompatibility proteins. *Structure*, 2(4):245–251, 1994.

Jocelyn Sunseri and David Ryan Koes. Virtual screening with gnina 1.0. *Molecules*, 26(23):7369, 2021.

Kathryn Tunyasuvunakool, Jonas Adler, Zachary Wu, Tim Green, Michal Zielinski, Augustin Žídek, Alex Bridgland, Andrew Cowie, Clemens Meyer, Agata Laydon, et al. Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873):590–596, 2021.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Mikhail Volkov, Joseph-André Turk, Nicolas Drizard, Nicolas Martin, Brice Hoffmann, Yann Gaston-Mathé, and Didier Rognan. On the frustration to predict binding affinities from protein–ligand structures with deep neural networks. *Journal of Medicinal Chemistry*, 2022.

Jeffrey R Wagner, Christopher T Lee, Jacob D Durrant, Robert D Malmstrom, Victoria A Feher, and Rommie E Amaro. Emerging computational methods for the rational discovery of allosteric drugs. *Chemical reviews*, 116(11):6370–6390, 2016.

David S Wishart, Craig Knox, An Chi Guo, Dean Cheng, Savita Shrivastava, Dan Tzur, Bijaya Gautam, and Murtaza Hassanali. Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research*, 36(suppl_1):D901–D906, 2008.

Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chemical science*, 9(2):513–530, 2018.

Lei Xie, Li Xie, and Philip E Bourne. Structure-based systems biology for analyzing off-target binding. *Current opinion in structural biology*, 21(2):189–199, 2011.

Zhijian Xu, Cheng Peng, Yulong Shi, Zhengdan Zhu, Kaijie Mu, Xiaoyu Wang, and Weiliang Zhu. Nelfinavir was predicted to be a potential inhibitor of 2019-ncov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. *BioRxiv*, pp. 2020–01, 2020.

Jianyi Yang, Ambrish Roy, and Yang Zhang. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*, 29(20):2588–2595, 2013.

Jingsong Yang, Nino Campobasso, Mangatt P Biju, Kelly Fisher, Xiao-Qing Pan, Josh Cottom, Sarah Galbraith, Thau Ho, Hong Zhang, Xuan Hong, et al. Discovery and characterization of a cell-permeable, small-molecule c-abl kinase activator that binds to the myristoyl binding site. *Chemistry & biology*, 18(2):177–186, 2011.

Jianming Zhang, Francisco J Adrián, Wolfgang Jahnke, Sandra W Cowan-Jacob, Allen G Li, Roxana E Iacob, Taebo Sim, John Powers, Christine Dierks, Fangxian Sun, et al. Targeting bcr–abl by combining allosteric with atp-binding-site inhibitors. *Nature*, 463(7280):501–506, 2010.

Jingtian Zhao, Yang Cao, and Le Zhang. Exploring the computational methods for protein-ligand binding site prediction. *Computational and structural biotechnology journal*, 18:417–426, 2020.

Xiaolei Zhu, Yi Xiong, and Daisuke Kihara. Large-scale binding ligand prediction by improved patch-based method patch-surfer2. 0. *Bioinformatics*, 31(5):707–713, 2015.

4 SUPPORTING INFORMATION

4.1 POCKET ENSEMBLE ANALYSIS

Test Set

Ensemble	Success Rate	Average IoU	Average Dis. (nm)
SM	86.2%	0.43	0.59
P2R	**93.4%**	0.41	**0.43**
FP	79.6%	0.42	0.66
CP	90.6%	**0.49**	0.53
SM + P2R	94.5%	0.50	0.36
SM + FP	93.1%	0.50	0.39
SM + CP	92.3%	0.52	0.38
P2R + FP	**95.6%**	0.47	0.37
P2R + CP	95.3%	0.52	**0.35**
FP + CP	93.9%	**0.52**	0.40
SM + P2R + FP	95.6%	0.52	0.33
SM + P2R + CP	95.6%	**0.55**	**0.31**
SM + FP + CP	93.9%	0.54	0.33
P2R + FP + CP	**96.1%**	0.54	0.32
SM + P2R + FP + CP	**96.1%**	**0.56**	**0.29**

Test Set w/o Peptides

Ensemble	Success Rate	Average IoU	Average Dis. (nm)
SM	93.1%	0.44	0.49
P2R	**99.0%**	0.40	**0.36**
FP	85.6%	0.43	0.60
CP	98.0%	**0.50**	0.51
SM + P2R	99.7%	0.49	**0.29**
SM + FP	99.3%	0.51	0.32
SM + CP	98.7%	0.54	0.35
P2R + FP	**100%**	0.47	0.31
P2R + CP	**100%**	0.53	0.31
FP + CP	99.3%	**0.54**	0.37
SM + P2R + FP	**100%**	0.52	**0.26**
SM + P2R + CP	**100%**	0.55	0.27
SM + FP + CP	99.3%	**0.56**	0.29
P2R + FP + CP	**100%**	0.55	0.27
SM + P2R + FP + CP	**100%**	**0.56**	**0.25**

Table 1: Ensemble analysis with top-10 results from SiteMap (SM), P2Rank (P2R), FPocket (FP), CurPocket (CP). All metrics are calculated by using the best matching box regardless of rank. Success rate is the percent of cases where the box fully encloses the native pose, average IoU is the mean intersection over union (IoU), and average distance is the mean distance between the box and ligand centroids. Best results in each tier is bolded.

4.2 MODEL

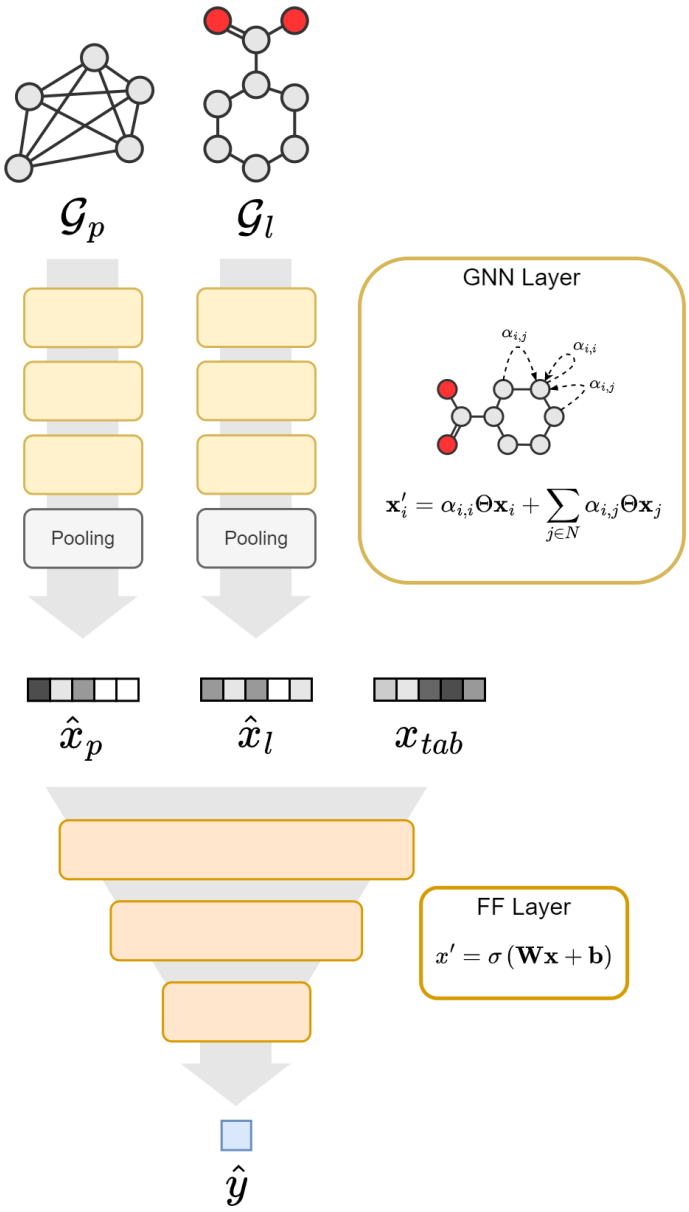

Figure 5: PocketNet model overview. Protein and ligand graphs are prepared prior to training and encode information about the atoms in each system. Ligand edges are maintained between covalently bonded atoms and the pocket graph is locally connected with a cutoff of 8Å. In addition to the graph structures, several tabular features which describe the entire system are also provided as input to the network (2). The network architecture is composed of two graph neural networks which process the ligand and pocket graphs independently before joining the embeddings with the tabular features and feeding them through a feed-forward network. In order to learn a fixed-sized embedding for systems of variable size, a pooling operation reduces the nodes features of the graph into one fixed-size embedding. Generated embeddings are concatenated with tabular features and fed to the feed-forward network. This network gradually shrinks the dimensions of the embedding until reaching a scalar value prediction

4.3 FEATURIZATION DETAILS

Data Structure	Descriptor	Encoding
Pocket Graph Nodes	Residue Type	One-Hot
	Secondary Structure	One-Hot
	Total SASA	Z-Score
	P2Rank Score	Z-Score
Pocket Graph Edges	Distance	None
Ligand Graph Nodes	Atom Type	One-Hot
	Hybridization Type	One-Hot
	Chirality Type	One-Hot
	Number of Bonds	One-Hot
	Number of Hydrogens	One-Hot
	Number of Rings	One-Hot
	Ring Sizes	One-Hot
	Formal Charge	One-Hot
	Partial Charge	None
	Aromatic Flag	None
Ligand Graph Edges	Bond Type	One-Hot
	Distance	None
Pocket Tabular	Source Program	One-Hot
	Source Rank	Log
	Ensemble IoU Sum	Z-Score
	Ensemble IoU Mean	Z-Score
	P2Rank Score	Z-Score
	Sorted Side Lengths	Z-Score
	Side Length Ratios	Z-Score
	3D Box Volume	Z-Score
	Atom Count / Volume	Z-Score
	Residue Count / Volume	Z-Score
	Protein Chain Count	None
	Fraction of Residues Inside Box	None
	Fraction of SiteMap Points Inside Box	None
	Fraction of P2Rank Points Inside Box	None
	Fraction of FPocket Points Inside Box	None
Ligand Tabular	Molecular Weight	Z-Score
	Number of Rotatable Bonds	Z-Score
	Number of H-Bond Donors	Z-Score
	Number of H-Bond Acceptors	Z-Score
	Number of Total Rings	Z-Score
	Number of Aliphatic Rings	Z-Score
	Number of Aromatic Rings	Z-Score
	Number of Saturated Rings	Z-Score
	Number of Amide Bonds	Z-Score
	Solvent Accessible Surface Area	Z-Score
	Total Polar Surface Area	Z-Score
	Crippen logP	Z-Score

Table 2: Graph and tabular feature details. One-Hot indicates a categorical feature encoding where a vector contains one non-zero element indicating the category. After the first neural network layer, this sparse encoding is transformed into a dense feature vector. Z-score indicates that a scalar feature has been normalized using the mean and standard deviation across the entire dataset. Log indicates that a scalar feature has been log-transformed in order to attenuate larger values. Some scalar features were left unnormalized.

4.4 TRAINING DETAILS

The model was trained using an appropriate loss function for each target variable (e.g. MSE for IoU and distance, BCE for success). Hyperparameter optimization using the validation set alone was performed to find the optimal set of hyperparameters (see Table 3 below for further details). The best model was trained for 100 epochs using the Adam optimizer with a learning rate of 0.0001. Training samples were shuffled and provided to the network as batches of 256.

	Parameter	Search Space	Best Value
Training	Target Variable	[Success, IoU, Distance]	Distance
	Learning rate	Log Uniform ($10^{-1} - 10^{-5}$)	10^{-4}
Pocket GNN	Hidden layers	[1, 3, 5]	3
	Hidden dim.	[64, 128, 256, 512]	256
	Output dim.	[64, 128, 256, 512]	512
	Dropout	Uniform (0.0 - 0.25)	0.0
Ligand GNN	Hidden layers	[1, 3, 5]	3
	Hidden dim.	[64, 128, 256, 512]	256
	Output dim.	[64, 128, 256, 512]	512
	Dropout	Uniform (0.0 - 0.25)	0.0
Feed-Forward NN	Hidden layers	[1, 3, 5]	3
	Hidden dim.	[64, 128, 256, 512]	256
	Dropout	Uniform (0.0 - 0.25)	0.1

Table 3: Search space and results for hyperparameter optimization (HPO). For faster training, the validation set alone was used for HPO. While the target variable used for training was itself a hyperparameter, the HPO aimed to optimize the success for the top-1 predicted pocket. By this evaluation, training by success had about the same performance as training by distance, the latter was found to be superior in terms of the IoU and distance metrics. Therefore, the best model was trained by centroid distance.

4.5 METRICS

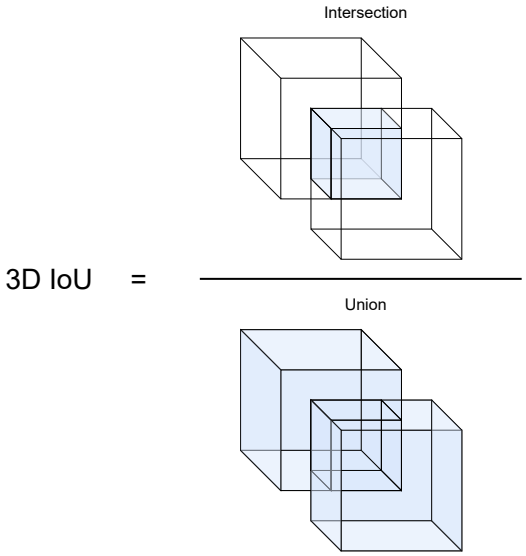

Figure 6: Depiction of 3D Intersection over Union (IoU) calculation. Using the axis-aligned box definition, the calculation of intersection and union of the boxes becomes trivial.

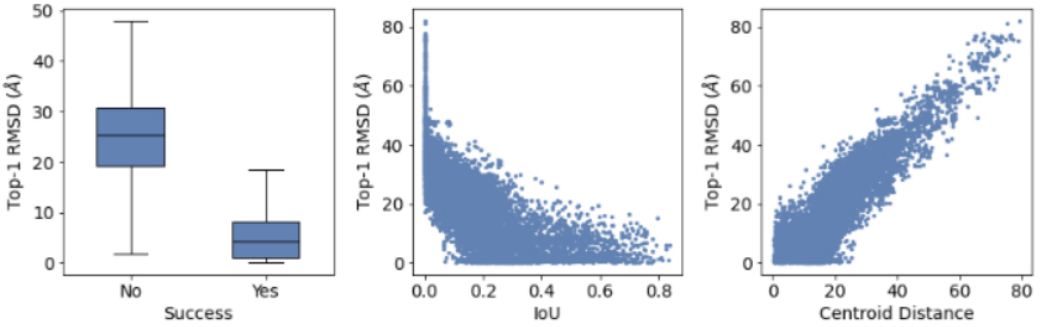

Figure 7: Correlation between the three pocket prediction metrics (success, IoU, and centroid distance) and top-1 docking RMSD (Å). Success guarantees the native pose is contained within the search box, but the docking algorithm can still find to sample or accurately score this pose. IoU has a negative correlation with RMSD (R=-0.73). Boxes with IoU = 0 (no overlap with native pose), the RMSD is at least 20Å. Boxes with IoU $>$ 0.5 (high overlap with native pose), the RMSD is at most 20Å. The metric, centroid distance, has the highest correlation with RMSD (R=0.89). This coincides with findings from the hyperparameter optimization which found centroid distance to be the best target variable.

4.6 ADDITIONAL RESULTS

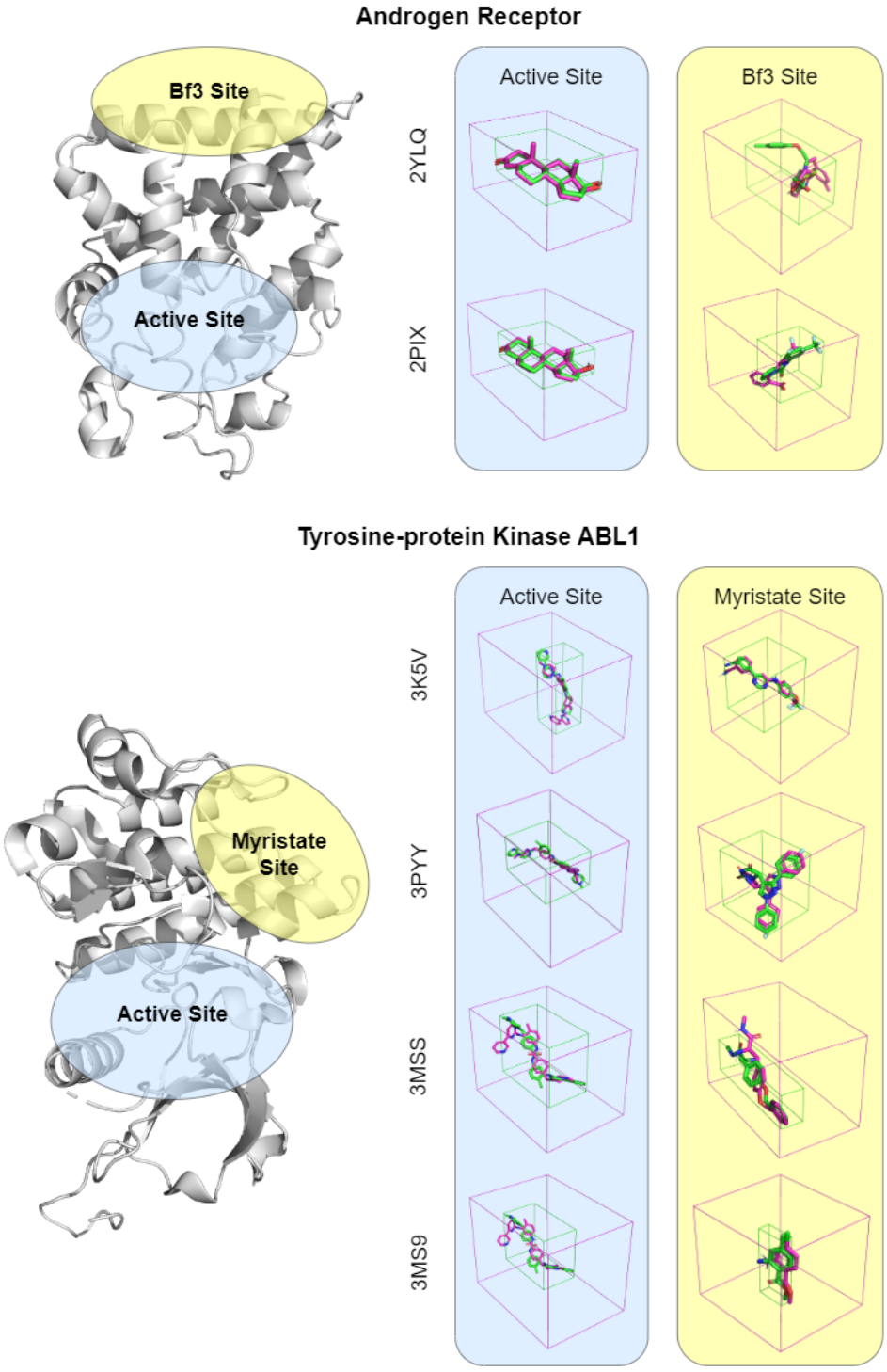

Figure 8: Visualized results from the differential pocket prediction analysis for Androgen Receptor and Tyrosine-protein Kinase ABL1. Predicted pocket box and docked pose shown in magenta. Native pose and its bounding box is shown in green.