

Graph Neural Networks Bootstrapped for Synthetic Selection and Validation of Small Molecule Immunomodulators

Prageeth R. Wijewardhane,^{#,1} Krupal P. Jethava,^{#,1} Jonathan A. Fine,¹ Gaurav Chopra^{*,1,2}

¹Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN, USA

²Purdue Institute for Drug Discovery, Integrative Data Science Institute, Purdue Center for Cancer Research, Purdue Institute for Inflammation, Immunology and Infectious Disease, Purdue Institute for Integrative Neuroscience, West Lafayette, IN, USA

[#]These authors share an equal contribution to this work

^{*}Corresponding author

ABSTRACT

The Programmed Cell Death Protein 1/Programmed Death-Ligand 1 (PD-1/PD-L1) interaction is an immune checkpoint utilized by cancer cells to enhance immune suppression. There is a huge need to develop small molecule drugs that are fast acting, cost effective, and readily bioavailable compared to antibodies. Unfortunately, synthesizing and validating large libraries of small-molecules to inhibit PD-1/PD-L1 interaction in a blind manner is both time-consuming and expensive. To improve this drug discovery pipeline, we have developed a machine learning methodology trained on patent data to identify, synthesize, and validate PD-1/PD-L1 small molecule inhibitors. Our model incorporates two features: docking scores to represent the energy of binding (E) as a global feature and sub-graph features through a graph neural network (GNN) of molecular topology to represent local features. This interaction energy-based Graph Neural Network (EGNN) model outperforms traditional machine learning methods and a simple GNN with a F1 score of 0.9524 and Cohen's kappa score of 0.8861 for the hold out test set, suggesting that the topology of the small molecule, the structural interaction in the binding pocket, and chemical diversity of the training data are all important considerations for enhancing model performance. A Bootstrapped EGNN model was used to select compounds for synthesis and experimental validation with predicted high and low potency to inhibit PD-1/PD-L1 interaction. The potent inhibitor, (4-((3-(2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)-2-methylbenzyl)oxy)-2,6-dimethoxybenzyl)-D-serine, is a hybrid of two known bioactive scaffolds, with an IC₅₀ of 339.9 nM that is comparatively better than the known bioactive compound. We conclude that our bootstrapped EGNN model will be useful to identify target-specific high potency molecules designed by scaffold hopping, a well-known medicinal chemistry technique.

INTRODUCTION

Programmed cell death protein 1 (PD-1) is an immune checkpoint receptor implicated for the creation of new cancer therapeutics.¹ The prolonged interaction between the T-cell receptor and the major histocompatibility complex (MHC) leads to upregulation of PD-1 on the activated T-cell surface.² Activated T cells produce cytokines, such as Interferon- γ , which in turn cause tumor cells to express programmed death ligand 1 (PD-L1) on their cell surface.² Tumors escape the action of immune system by utilizing the interaction between PD-1 and ligand PD-L1 resulting in lower effector T-cell function and survival, as such resulting in a suppressive immune response in the tumor microenvironment.² The inhibition of the PD-1/PD-L1 interaction can enhance anti-tumor immunity and a large amount of work has been done to develop monoclonal antibodies as inhibitors of PD-1/PD-L1 interaction inhibitors.^{3,4} For example, pembrolizumab, cemiplimab, and nivolumab are three FDA approved anti-PD-1 monoclonal antibodies.⁴ The discovery of small-molecule inhibitors would be an advantageous over monoclonal antibodies, such as being fast-acting, simple for *in vivo* administration, ability to penetrate through cell membranes and interact with the cytoplasmic domains of cell surface receptors.⁵ Since a few years, there has been significant development in designing PD-1/PD-L1 inhibitors.^{6,7} Specifically, Bristol-Myers Squibb (BMS) discovered a set of potent PD-1/PD-L1 small molecule inhibitors based on the peptidomimetic molecules and non-peptidic small molecules.^{6,7} In particular, BMS revealed a 2-methyl-3-biphenyl-methanol scaffold containing chemical libraries. Later, Holak *et al.* studied the interaction of BMS molecules with PD-L1 suggesting that BMS molecules induce PD-L1 dimerization and also reported crystal structures of compounds with dimeric PD-L1.^{8,9}

Based on these findings, we envisioned to develop a machine learning (ML) framework for selecting and testing new PD-1/PD-L1 inhibitors.

Traditionally, the development of small-molecule inhibitors requires high throughput screening of a large library of diverse drug-like compounds¹⁰ or a medicinal chemist iterating over a scaffold with weak receptor activity to enhance potency.¹¹ This entire process is – (i) time consuming; (ii) needs expensive instrumentation and robotics; (iii) based on trial-and-error; and (iv) highly inefficient to identify several new scaffolds rapidly.¹² In addition, virtual screening using docking methods have been developed to improve this process but with limited success.¹³ Further, ML architectures such as Support Vector Machine (SVM)^{14–16}, Random Forest (RF)^{17–19}, Graph Convolution Network²⁰, and Graph Neural Networks (GNN)^{21,22} have been used for drug design and predicting drug-target interactions^{23,24}. Recently, new architectures utilizing a combination of graph features in the binding site of a protein have shown great promise for calculating binding affinities and determining whether a compound will bind to a target.^{20,22}

Several new neural network-based architectures have also been proposed which promise to identify potent scaffolds, but many have not been tested experimentally,^{15,16,25–28} and developments in the ability to mine and characterize protein crystallography data hopes to drive the creation of these models.²⁹ Recently, it has been shown that molecular sub-graph features incorporated through a GNN and protein features encoded by their sequence can be combined to predict if a compound can target a given protein.²⁴ Inspired by this work and based on our interest in developing methods for drug design and immunology^{29–36}, we have developed a new machine learning model to predict if a compound can inhibit the PD-1/PD-L1 interaction. Our method replaces the protein sequence features with docking scores representing the free energy

Fluorescence (HTRF) assay. We also tested negative predictions suggesting the utility of the model to be used for selecting potent leads as PD-1/PD-L1 inhibitors.

RESULTS

Patent Data for Training the EGNN Model

We used PD-1/PD-L1 small molecule inhibition data for 762 compounds from four patents to train our models: WO 2015/034820 A1⁷ and WO 2015/160641 A2³⁷ by BMS (674 compounds),

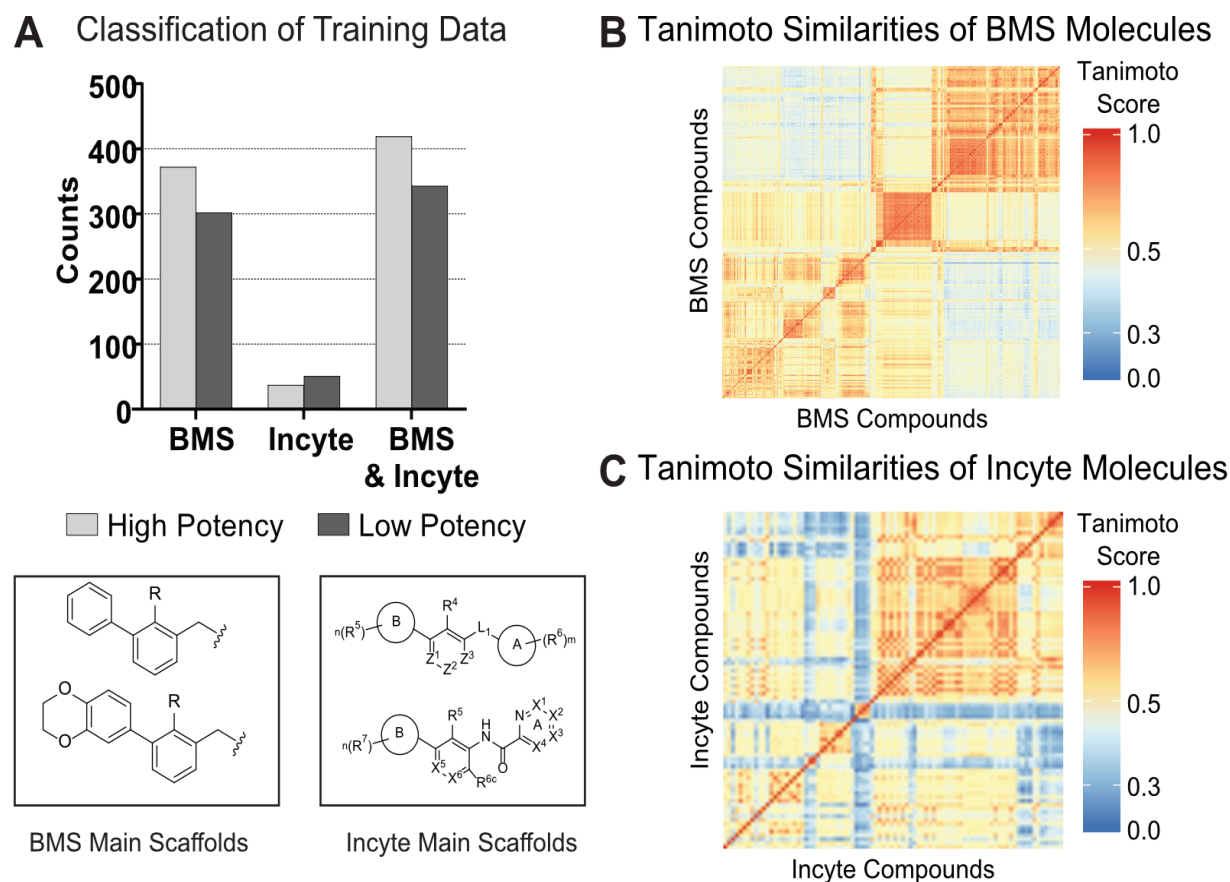


Figure 2. (A) Upper: Classification of Training Data in BMS and Incyte Patents. Bottom: Left: Main PD-L1 inhibitor scaffolds of BMS patents. R group can be CN, Cl, Br, or CH₃. Right: Main PD-L1 inhibitor scaffolds of Incyte patents. Here A and B denote sub-scaffolds. (B) & (C) Heatmaps of pairwise Tanimoto similarity scores of BMS and Incyte compounds, respectively.

and WO 2018/119263 A1³⁸ and US 2018/0273519 A1³⁹ by Incyte Corporation (88 compounds). A homogeneous time-resolved fluorescence (HTRF) binding assay was used to show activity against PD-1/PD-L1 interaction in the patents. However, the patents did not list individual IC₅₀ values for all compounds but provided a range of inhibition with different molecules. Therefore, we trained a binary classifier with cutoffs for both datasets to treat a molecule as “High potency” or “Low potency” (**Figure 1**). If the reported IC₅₀ of a molecule is less than or equal to 100 nM in the patent it was considered as a “High potency” molecule, otherwise it was considered as a “Low potency” molecule. This threshold was selected as it is the only common threshold value among four patents (**Table S4**). It should be noted that the actual value of IC₅₀ should not be considered here as our experiments with multiple replicates were not able to obtain exactly reported results for some molecules in the patents (see IC₅₀ value of compound **4a** in **Table 2**, BMS-1 annotated with 6-100 nM in the WO 2015/034820 A1 patent⁷) since our experimental conditions were different from the conditions reported in patents (See HTRF assay part in the experimental section). Therefore, we consider positive prediction (high potency) based on our experimental IC₅₀ value as compared to the upper limit of a BMS control molecule (compound **4a**/BMS-1) in WO 2015/034820 A1 patent⁷. The training dataset of 762 small molecules with the BMS or Incyte annotation is shown in Supporting Information File (**TrainingData.xlsx**).

We selected BMS and Incyte patents to include chemical diversity of the molecules in the training data set. **Figure 2A** shows the distribution of low and high potency molecules and general scaffolds in the BMS and Incyte patents. The BMS patents have 372 high potency compounds and 302 low potency compounds while the Incyte patents have 47 high potency compounds and 41 low potency compounds respectively. The BMS patent scaffolds contains 417 derivatives of (2-

methyl-3-biphenyl)methanol and 257 derivatives of [3-(2,3-dihydro-1,4-benzodioxin-6-yl)-2-methylphenyl]methanol shown in **Figure 2A (bottom - left)** with R groups as CN, Cl, Br, and CH₃. On the other hand, Incyte patent scaffolds have distinct sub-scaffolds, denoted as A and B in **Figure 2A (bottom – right)**. For Incyte scaffolds, X denotes for either N or C—R groups (R: Alkyl groups). These scaffolds suggest that the chemical diversity of Incyte compounds is higher than that of the BMS compounds because the general structures of the compounds in Incyte patents have more structural diversity for sub-scaffolds and atoms. We validated this observation using pairwise Tanimoto similarity scores of BMS and Incyte compounds as shown as heatmaps in **Figure 2B** and **2C**, respectively. Morgan fingerprints with radius of 2 and bit length of 1024 were used to calculate pairwise Tanimoto similarities. High red color areas in the BMS heatmap indicates that the molecular pairs are structurally similar to each other. Low red areas in the Incyte heatmap suggests it has more chemical diversity in molecular structures. Furthermore, the average pairwise Tanimoto⁴⁰ similarity score of all BMS compounds was found to be 0.4434 and 0.3920 for all Incyte compounds, confirming higher chemical diversity in Incyte compounds when compared to BMS compounds.

PD-L1 homodimer and PD-1/PD-L1 Crystal Structures Reveals a Binding Site for Docking

It has been shown previously that BMS compounds inhibit the PD-1/PD-L1 interaction by inducing dimerization of PD-L1.^{8,9} Therefore, a PD-L1 homodimer crystal structure (PDB ID: 5N2F) was selected for docking all the compounds in this manuscript. A PD-1/PD-L1 crystal structure (PDB ID: 4ZQK) was also used to check whether the binding site location of PD-L1 in the homodimer

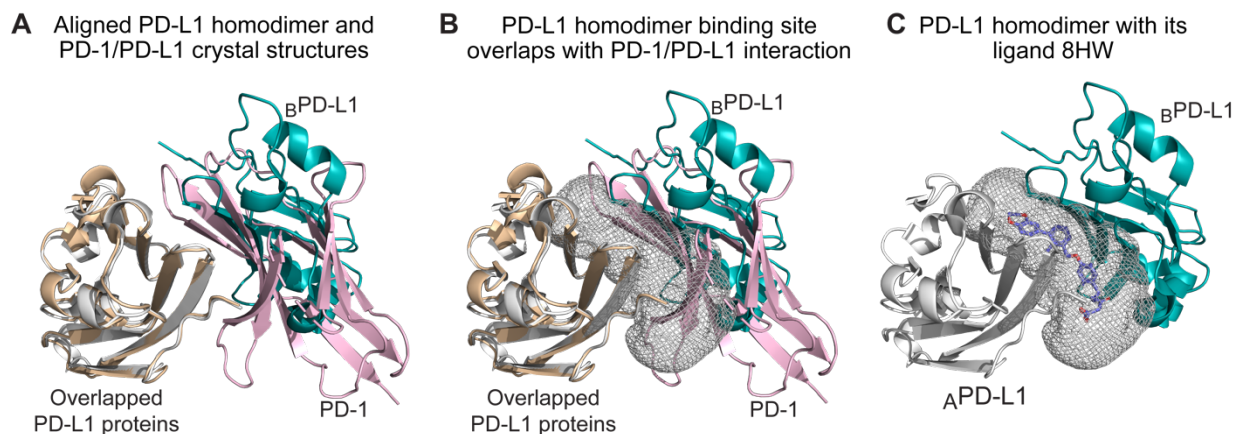


Figure 3. The light pink chain represents the PD-1 protein and the pale cyan chain represents the PD-L1 protein in the PD-1/PD-L1 complex crystal structure (PDB ID: 4ZQK). The wheat color chain represents the PD-L1 chain A and the blue white color represents the PD-L1 chain B in the PD-L1 homodimer crystal structure (PDB ID: 5N2F) (A) Overlapped and aligned PD-1/PD-L1 (4ZQK) and PD-L1 dimer (5N2F) crystal structures. (B) Overlapped and aligned two crystal structures with the determined binding site (grey color mesh) of the PD-L1 dimer (5N2F). (C) The PD-L1 dimer (5N2F) crystal structure with the small molecule (ligand ID: 8HW) at its binding site (grey color mesh).

crystal structure (5N2F) overlapped and aligned with each other using the PyMol software package⁴¹(**Figure 3A**). In **Figure 3B**, the selected binding site of the PD-L1 homodimer on the overlapped and aligned crystal structures is shown to indicate that the formation of the homodimer of PD-L1 with small molecules blocks the PD-1/PD-L1 interaction. A known inhibitor of the PD-1/PD-L1 interaction (ligand ID: 8HW)⁸ in the selected binding site (**Figure 3C**) suggests that the selected binding site corresponding to PD-L1 homodimers is relevant to develop PD-1/PD-L1 inhibitors. Therefore, the docking interactions of the PD-L1 homodimer will be relevant towards identifying PD-1/PD-L1 inhibitors. Also, direct docking with the PD-1/PD-L1 was not carried out since the binding site in between the PD-1 and the PD-L1 is filled with interacting amino acid residues from both proteins. Therefore, there is no space to dock a small compound with the PD-1/PD-L1 complex.

CANDOCK³¹ was used to generate docking conformations of small molecules with PD-L1 homodimer (see Experimental Section on **Generation of Energy Features with Docking and Energy Vector (E) in EGNN** for details). Before developing a machine learning method, we also assessed the ability of only using the docking scores for compounds in the training set for each of the 96 potential energy scoring functions³¹ in CANDOCK to classify the high potency vs low potency molecules. Cohen's Kappa scores were used to select the best scoring functions to differentiate between two classes. (**Table S1**). The scoring function, radial cumulative complete 15 (RCC15) acquired the highest Cohen's kappa score of 0.41447. However, RCC15 scores were not able to clearly separate all the high and low potent classified molecules in the training data (see Violin plots in **Figure S1**). Using only one scoring function is not sufficient to capture the different states of PD-1/PD-L1 inhibition with small molecules. Therefore, we developed an EGNN model using top scoring function of each class which demonstrated a positive Cohen's kappa value (**Table S1**) to represent the global features (see Experimental Section on **Generation of Energy Features with Docking and Energy Vector (E) in EGNN** for details). This included RCR15 (kappa = 0.37746) and RCC15 (kappa = 0.41447) scoring functions. A model with kappa score between 0.21 - 0.40 is considered as a fair agreement model and if the kappa score is between 0.41 - 0.60, then the model is considered as a moderate agreement model.⁴²

EGNN Model with Hyperparameter Optimization Outperforms GNN and Other Baseline Models

A detailed description of the EGNN model including a combination of molecular GNN combined with docking is given in the **Experimental Section**. **Figure 1** shows that the EGNN model is a combination of local features of the small molecule represented as a GNN (see **Graph Neural Network for Molecular Graphs in EGNN**) along with global features of protein-ligand interaction

represented as docking scores (see section **Generation of Energy Features with Docking and Energy Vector (E) in EGNN**). The EGNN was trained with 88 small molecules with high and low potency for PD-1/PD-L1 inhibition extracted from two Incyte patents (see **Patent Data for Training the EGNN Model**). We calculated variation in the average F1 score (over five cross-validated folds) versus the number of epochs for different hyperparameters (**Figure S2**). Also, model's performance with different datasets with changing hyperparameters are shown in **Table S6**. Optimal hyperparameters were selected to avoid overfitting and underfitting for EGNN. Dimension of the hidden molecular vector (dim) = 10, sub-graph radius = 2, and number of hidden layers = 1 were selected as optimum hyperparameters for the EGNN model (see Experimental Section on **EGNN Training and Hyperparameter Optimization**).

The EGNN and GNN models were trained with different training sets to examine the effect of chemical diversity on model performance for classification of high and low potency molecules. Two datasets (BMS and Incyte) were used separately and in combination to train the EGNN model and determine the best dataset to predict PD-1/PD-L1 inhibitors. Splitting of the dataset into training-validation set and test set (4:1) were carried out using two different methods: (1) using a random splitter on shuffled data and (2) using a scaffold splitting method by DeepChem library.⁴³ Then training was carried out with fivefold cross validation and test sets were used to evaluate the models' prediction ability. Here, Cohen's kappa, F1 score and Area Under the Receiver Operator Characteristic Curves (AUROC) were measured to compare three models trained with BMS data only, Incyte data only and BMS-Incyte combined data. Further, as a separate experiment, all the measures were obtained for EGNN and GNN models trained only on BMS data while predicting for Incyte data, and vice versa as well.

Figure 4A shows how data sets were split within the train-validation-test set scheme. Initial dataset was split into two sets with the 4:1 ratio based on the scaffold splitting or random shuffled splitting. Then the 80% dataset was used as the training and validation dataset while the 20% dataset was used as the hold out test set to evaluate model performances.

GNN models with scaffold splitting appeared to generate comparable results with the EGNN (**Figure S4**). However, this was expected since the graph neural network uses the two-dimensional molecular framework/topology in training. When the framework distributions of the compounds are similar in the train-validation and test sets, GNN performs well. However, our intension is to develop a model which could be used to screen a large compound library which would not be necessary to share the same distribution of scaffolds with training sets (i.e. Incyte or BMS). Hence, we selected the random splitter with shuffling to create the test set for performance evaluations to develop a more generalized model.

Cohen's kappa scores of different test sets (hold out test set is based on random splitting) for both models trained with BMS compounds, Incyte compounds, and the union of these sets are shown in **Figure 4B**. The kappa scores of the EGNN and GNN models trained with Incyte data and tested on the hold out test set were 0.8861 and 0.4304, respectively (**Figure 4B**). This result suggests that the EGNN trained model with Incyte data that contains diverse chemical scaffolds (**Figure 2C**) performs much better than the GNN trained with the same data set. However, when the same test was done with only BMS compounds with lower chemical diversity than Incyte, the Cohen's kappa score is comparable for both models with 0.6416 for the EGNN model and 0.7164 for the GNN model. This suggests that the GNN model performs well with smaller chemical diversity in the training and test data as compared to larger chemical diversity. Both models show

comparable performances with the combined datasets as well. When both BMS and Incyte datasets were combined, the kappa score for the hold out test set of the EGNN model was 0.6072 and 0.6729 for the GNN model. A similar trend is observed for the F1 scores for the three different training set comparisons (**Figure 4C**). These results suggest that the EGNN model outperforms the GNN model for chemically diverse data sets such as Incyte data. We believe this to be due to the addition of ‘global’ energy features captured by the docking scores of PD-L1 homodimers as training data in EGNN compared to only the ‘local’ structural features of small molecules in the training data for the GNN model.

We also investigated the ability of the EGNN and GNN models trained on one compound set to predict high and low potency inhibitors of PD-1/PD-L1 in the other compound set. These results are represented in **Figure 4B** and **4C** (Kappa scores and F1 scores respectively) with different bar patterns to represent different test sets. Tanimoto similarities between Incyte and BMS compounds are also shown in the **Figure 4D** heat map. The average pairwise Tanimoto similarity score of 0.3044 shows that compounds in these two datasets are very dissimilar to each other. When the EGNN and GNN models are trained on BMS compounds and used to predict Incyte compounds, a Cohen’s kappa score of 0.1505 for the EGNN and 0.1200 for the GNN was observed and F1 score of 0.3810 and 0.2264 were observed, respectively. On the other hand, both F1 and kappa scores for both models improved when they were trained with Incyte data and used to predict the BMS compounds (kappa score of EGNN = 0.3852, GNN = 0.3196 and F1 score of EGNN = 0.7400, GNN = 0.6958). These results show that there is a marked improvement in F1 scores and Cohen’s kappa scores for both EGNN and GNN models when trained on Incyte data and tested on BMS. However, AUROC score cannot distinguish these models correctly

(**FigureS3**). This was expected as the AUROC is a not good measure to evaluate models trained on skewed/unbalanced datasets and it can hinder the poor performance of a model.⁴⁴ Also, these results suggest that the EGNN model outperforms the GNN model in both cases. (see **Table S2** for details). This highlights the importance of chemical diversity in training data even though there is not much compound similarity between the training and test sets. Therefore, it is not suitable to use only BMS or combined BMS and Incyte data to train the final model to make predictions for unknown molecules. Hence, we selected only the Incyte dataset to train the EGNN model which improved the EGNN model performance significantly (**Table S2**).

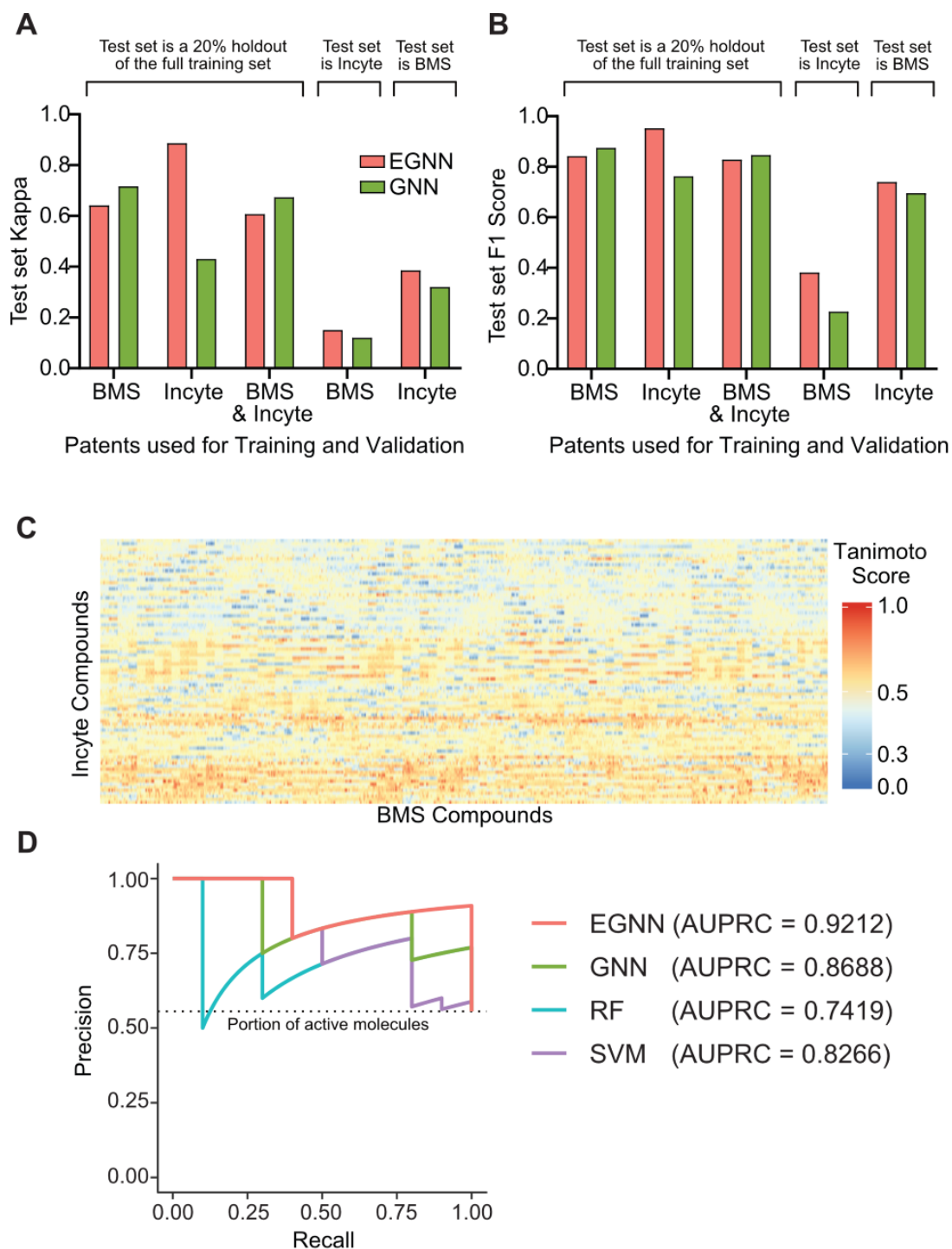


Figure 4. (A) Training-validation and Test scheme used for models. (B) Cohen's kappa scores for EGNN and GNN with different training-validation and test sets. (C) F1 scores for EGNN and GNN models with different training-validation and test sets. (D) Heatmap of pairwise Tanimoto similarity scores between BMS and Incyte compounds (E) Precision-recall curves for EGNN, GNN, RF and SVM models trained with Incyte data.

Also, we have checked the ability to classify training set compounds into low and high potency classes by just comparing the Tanimoto 2D similarities. Violin plots showing the distributions of Tanimoto 2D similarity scores for low potency, high potency and all compounds are shown in **figure S5**. This clearly shows that compounds either in the high potency class or the low potency class shows a high probability to have a low pairwise similarity score and that is true even when all the compounds are considered as well. So, it is not enough to just consider the topological similarity to select potent PD-1/PD-L1 inhibitors.

Finally, we compared the cross-validated EGNN model with GNN, Support Vector Machine (SVM), and Random Forest (RF) baseline models trained with Incyte training data, using their test set performances. Both SVM and RF models are trained on local and global features as well. Extracted fingerprints from a molecular graph with the radius of 2 using Weisfeiler-Lehman algorithm⁴⁵ were used as 'local' features similar to EGNN and GNN models. Here we have padded zeros up to the maximum fingerprint length to maintain the same fingerprint dimension. Same pre-selected docking scores (RCR15 and RCC15) obtained by CANDOCK were used as 'global' energy features. Obtained AUROC, AUPRC, Precision, Recall, F1 Score and Cohen's kappa values are tabulated in the **Table 1** for all four models. The SVM model was trained using the 'svm' package in scikit-learn library⁴⁶ with the "linear" kernel and the RF was trained using the 'RandomForestClassifier' in scikit-learn library⁴⁶ with 500 trees. The 'metrics' module in scikit-learn package⁴⁶ was used for statistics AUROC, precision, recall, F1 score and Cohen's kappa. Precision-recall curves for models and AUPRC values were obtained using the 'precrec' library⁴⁷ in R programming language. The EGNN model outperforms all the other models with values of 0.9250, 0.9212, 0.9091, 1.0000, 0.9524, and 0.8861 for AUROC, AUPRC, precision, recall, F1 score,

and Cohen’s kappa respectively (**Table 1**). Comparing precision-recall curves of these four models (**Figure 4E**) also confirmed that the EGNN model outperforms all three other models. Taken together, the combined local and global features in EGNN gives the best performance with the Incyte dataset.

Table 1. AUROC, AUPRC, Precision, Recall, F1 Score and Cohen’s kappa of the EGNN for PD-1/PD-L1 inhibitor predictions compared to other baseline models, such as, Random Forest, SVM, and GNN models. All models were trained on the Incyte dataset and evaluated based on the same hold out test set.

Measure	Model			
	Random Forest	SVM	GNN	EGNN
AUROC	0.8125	0.7750	0.8625	0.9250
AUPRC	0.7419	0.8266	0.8688	0.9212
Precision	0.7692	0.8000	0.7273	0.9091
Recall	1.0000	0.8000	0.8000	1.0000
F1 Score	0.8696	0.8000	0.7619	0.9524
Cohen’s Kappa	0.6494	0.5500	0.4304	0.8861

Synthetic Selection and Validation of EGNN Predictions for PD-1/PD-L1 Inhibition

The EGNN model trained with optimum hyperparameters and the Incyte dataset was used to get predictions for an in-house database of small molecular designs. We developed a bootstrapped EGNN model to predict compounds with high and low potency for PD-1/PD-L1 inhibition using 100 EGNN models (see section **Bootstrapping the EGNN model**). Bootstrapped EGNN models gave a F1 score of 0.91 and Kappa of 0.77 while bootstrapped GNN models gave 0.87 and 0.65

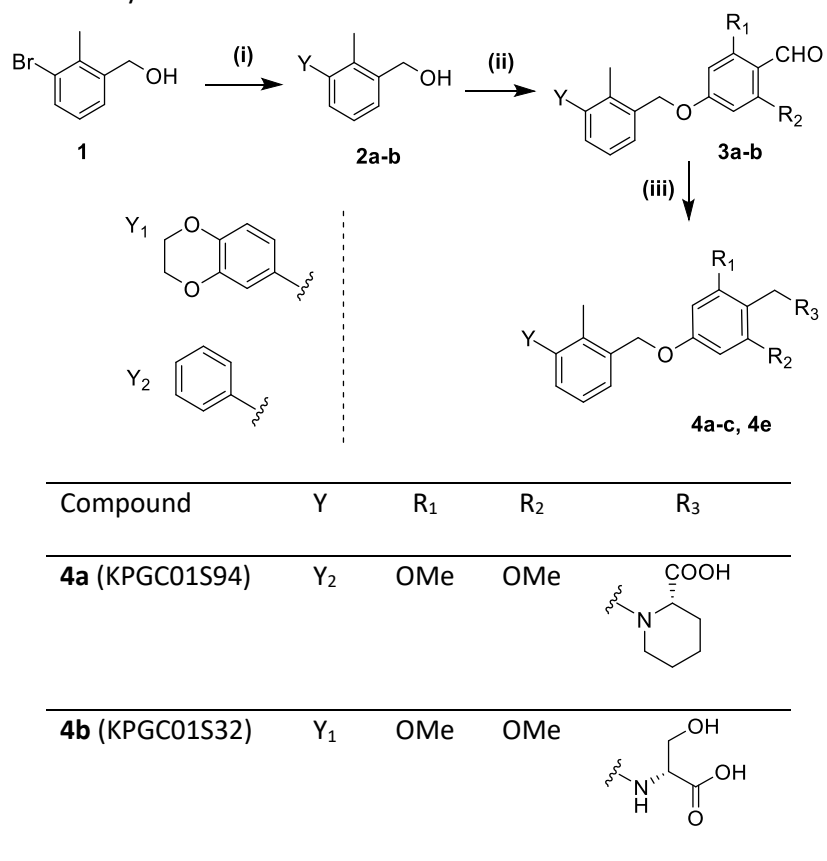
F1 score and Kappa respectively (**Table S5**). Bootstrapping is an essential statistical technique that can be used to select confident molecules for synthesis and experimental validation based on agreements among multiple models. The bootstrapped EGNN model identified high and low potency small molecules as PD-1/PD-L1 inhibitors that were synthesized and then experimentally verified with HTRF binding assay (see **Table 2** for summary). Specifically, we selected 4 molecules predicted to be high or low potent for PD-1/PD-L1 inhibition for testing based on bootstrapped EGNN SoftMax average scores and standard deviation (see EGNN SoftMax scores in **Table 2**).

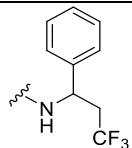
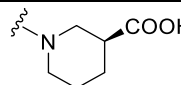
Out of EGNN bootstrapped predictions, we have selected 1 molecule as highly potent (compound **4b**) and 3 low potency molecules with different scaffolds (compound **4c**, **4d** and **4e**) for further testing. We have defined a new parameter called ‘counts’, which records the number of models out of 100 models which gives a SoftMax score of 0.5 or above for the molecule of interest. Specifically, the compound **4b** was predicted to be a high potency PD-1/PD-L1 inhibitor with 99 counts and an average SoftMax score of 0.7771 (± 0.1193). In contrast, only 69 counts and average SoftMax svalue of 0.5786 (± 0.1406) were resulted for compound **4c**, only 5 counts and an average SoftMax value of 0.1821 (± 0.1514) were resulted for compound **4d** and 62 counts and an average SoftMax value of 0.5280 (± 0.1259) were resulted for compound **4e** suggesting low potency predictions. We also synthesized a BMS scaffold (compound **4a** a known PD-1/PD-L1 inhibitor) for use as a positive control for our HTRF experiments. The compound structures are shown in **Scheme 1** and **2** (see **Experimental Section** for procedures and characterization). The predicted high potency molecule (compound **4b**) is (4-((3-(2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)-2-methylbenzyl)oxy)-2,6-dimethoxybenzyl)-D-serine, a hybrid of two BMS molecules, **4a** (BMS-1) and BMS-1002 containing (2-methyl-3-biphenyl)lmethanol and [3-(2,3-dihydro-1,4-

benzodioxin-6-yl)-2-methylphenyl]methanol, respectively (**Figure 5A**) and suggests the ability of EGNN model to do scaffold hopping.

The EGNN trained with Incyte data only had three [3-(2,3-dihydro-1,4-benzodioxin-6-yl)-2-methylphenyl]methanol scaffold containing compounds. As a separate experiment, we have removed these and predicted for our synthesized library and still the EGNN model was able to predict the compound **4b** (A compound based on [3-(2,3-dihydro-1,4-benzodioxin-6-yl)-2-methylphenyl]methanol scaffold) as a high potency compound with a SoftMax score of 0.8285 ± 0.1396 and 971 counts. This result demonstrates that the EGNN model can identify high potency PD-1/PD-L1 inhibitors with [3-(2,3-dihydro-1,4-benzodioxin-6-yl)-2-methylphenyl]methanol scaffolds even without being learned on similar scaffolds.

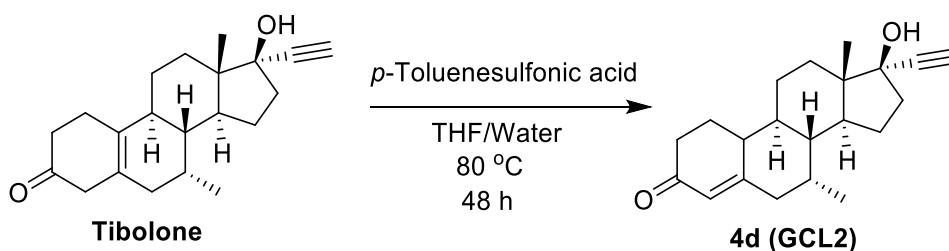
Scheme 1. Representative Synthesis Scheme^a



4c (KPGC01S138)	Y ₂	OMe	OMe	
4e (KPGC01S42)	Y ₂	OMe	OMe	

^a Reaction conditions: (i) BH₃•THF complex (1.0 M in THF), Anhydrous THF, 0 °C to rt, 2 days; (ii) PPh₃, DIAD, 0 °C to rt, 20 h, anhydrous THF; (iii) amine component, NaBH₃CN, cat. AcOH, DMF, 80 °C or room temperature, 1h or 3 h or overnight.

Scheme 2. Synthesis of **4d** (GCL2)



The top docked pose in PD-L1 homodimer (PDB ID: 5N2F) for compound **4b** interacts in a similar manner as shown previously for the co-crystal structures⁸ (**Figure 5B**). Specifically, for compound **4b**, the 2,3-Dihydro-1,4-benzodioxine group facilitated the movement of the amino acid residue Tyr56 in chain A of the PD-L1 homodimer (_ATyr56). It is known that this _ATyr56 does not close the hydrophobic pocket from one end if this 2,3-Dihydro-1,4-benzodioxine group is present⁸ creating a hydrophobic tunnel (**Figure 5B** inset) rather than a hydrophobic cleft in the docked conformation. Additionally, the aromatic ring of compound **4b** (2,3-Dihydro-1,4-benzodioxine) was stabilized by π - π stacking interactions with the amino acid residue _ATyr56 (**Figure 5C**). The central methylbenzyl ring (magenta color in **4b** in **Figure 5A**) in the structure is rotated by approximately 30° to 2,3-Dihydro-1,4-benzodioxine ring and the methyl group of the

methylbenzyl ring point towards chain B of the PD-L1 homodimer. This orientation results in hydrophobic interactions with Met115 of both chain A and B of the homodimer and with _BAla121. The D-serine end of the **4b** compound forms hydrogen bonds with _AThr20 and _AAla121 along with a plausible hydrogen bond formation between backbone NH of _ATyr123 and the oxygen in one of the two methoxy groups of the **4b** molecule (**Figure 5C**). These results suggest favorable interactions of compound **4b** that could dimerize PD-L1 will result in PD-1/PD-L1 inhibition.

The HTRF assay confirmed that compound **4b** has an IC_{50} of 339.9 nM (see **Experimental Section** for details) to inhibit PD-1/PD-L1 interaction (**Figure 5D**). This is comparatively better

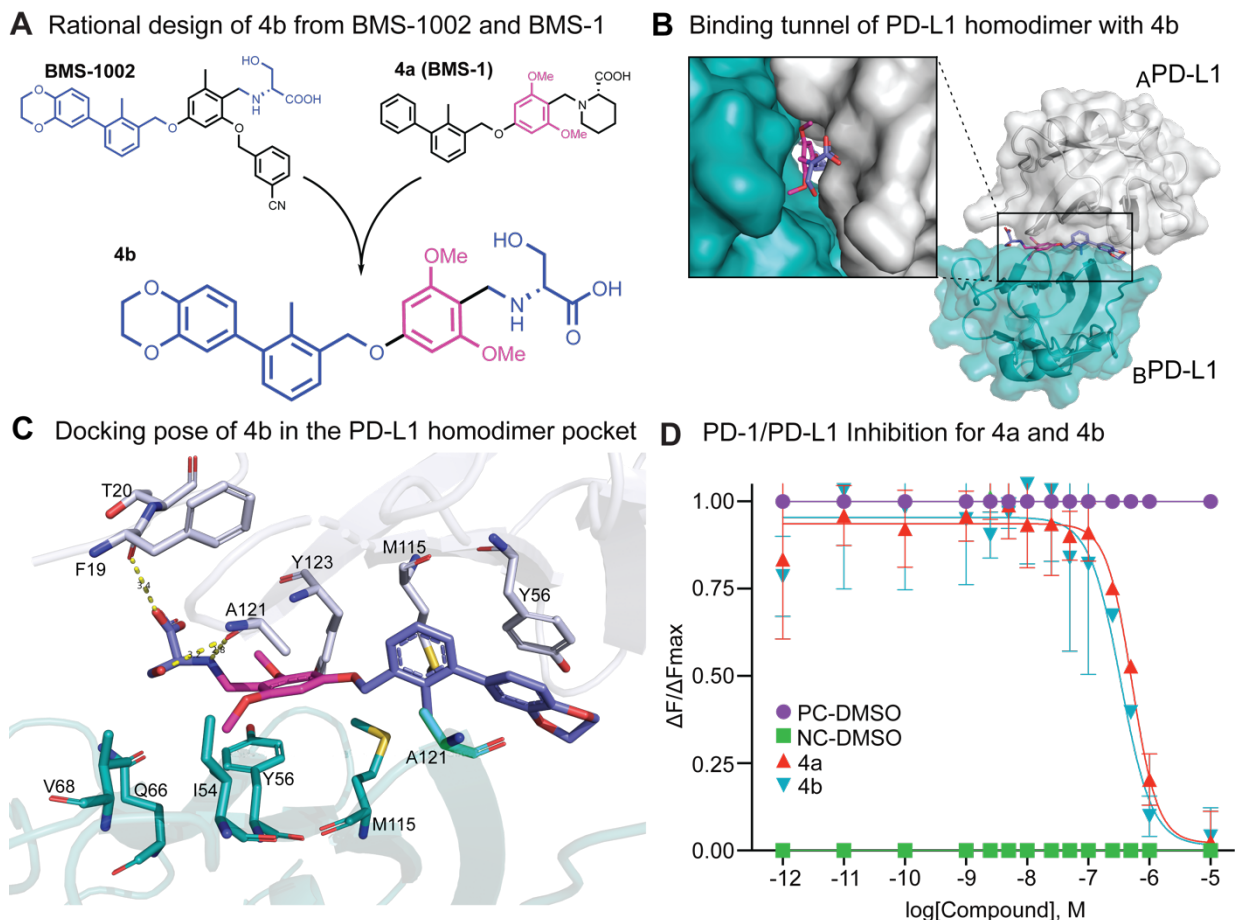


Figure 2. A. EGNN predicted a new PD-1/PD-L1 inhibitor, compound **4b**, by scaffold hopping of BMS compounds, **4a** or **BMS-1** and **BMS-1002**. Blue colored parts of the **4b** are added from the **BMS-1002** and pink color part was added from the **4a** (**BMS-1**). B. Showing location of top docked pose of the compound **4b** in PD-L1 homodimer crystal structure (PDB ID: 5N2F). Inset showing hydrophobic tunnel for compound **4b**. C. Showing chemical interactions of top docked pose interactions of the compound **4b** in PD-L1 homodimer. Blue and pink colored parts are shown as sticks for **4b**. The dotted yellow lines between the compound and the residues $_A$ Thr20 and $_A$ Ala121 represent hydrogen bonding. The orientation of the aromatic ring of tyrosine, $_A$ Tyr56, suggests a plausible π - π interaction with 2,3-dihydro-1,4-benzodioxin blue colored aromatic ring in the compound **4b**. D. Comparison of IC_{50} values of **4a** (**BMS-1** control compound, red color) and new compound **4b** (blue color). The DMSO controls for positive (PC-DMSO, purple color) and negative controls (NC-DMSO, green color) of the assay are shown for each tested concentration.

than the IC₅₀ of 521.5 nM for the BMS compound **4a** that was synthesized and tested in our lab (BMS-1 molecule in the BMS patent WO 2015/034820 A1). It should be noted that the BMS-1 molecule was denoted with the IC₅₀ of 6-100 nM with HTRF assay in the BMS patent⁷. However, multiple replicates of our experiments did not result in the IC₅₀ value less than 100 nM to inhibit PD-1/PD-L1 interaction (see **Calculation of IC₅₀ values** section and Supporting File **HTRF_IC50_Data.xlsx**). As mentioned previously, this result does not affect our machine learning method since we have classified molecules based on high and low potency rather than estimating the specific IC₅₀ value. A possible explanation of this difference in experimental results between our work and the patent could be differences in protocols used to perform the HTRF assay and calculation of IC₅₀ values. For this reason, we have included a detailed account of HTRF assay protocol, analysis of data for calculation of IC₅₀ and supporting data files to be used by the scientific community (see **Experimental Section**). In order to test the validity of our bootstrapped EGNN model to correctly identify low potency predictions, we also tested compounds **4c**, **4d** and **4e** resulting in no/poor inhibition to PD-1/PD-L1 interaction (**Table 2**). The IC₅₀ plots for each compound tested (**Figure S6**) as well as the ¹³C and ¹H NMR spectra are provided as **Supporting Information**. Pairwise Tanimoto similarity scores between these 4a-4e (**Table S3**) show that the EGNN model's capability of identifying high potency and low potency inhibitors regardless of the structural similarity. The compound **4e** shows a high similarity with the control BMS compound (4e) with a Tanimoto similarity score of 0.8018. However, the model recognized it as a low potency molecule and the actual test showed that it is a poor inhibitor for PD1/PD-L1 with an IC₅₀ of 1261 nM. On the other hand, the model recognized compound **4b** as a high potency PD1/PD-L1 inhibitor and the HTRF assay confirmed it with a very good IC₅₀ of 339.9 nM. However,

it's pairwise Tanimoto similarity score with the control compound (**4a**) is only 0.5074. Taken together, these results suggest that the bootstrapped EGNN model can be used to select molecules for synthesis and experimental validation of PD-1/PD-L1 inhibition and it can identify low potency molecules which are structurally similar to the control compound **4a**.

Table 2. IC₅₀ values for predicted high and low potency compounds with EGNN SoftMax scores

Compound	IC ₅₀ value	SoftMax score	Prediction
4a (KPGC01S94)	521.5 nM	Control*	Control*
4b (KPGC01S32)	339.9 nM	0.7771 +/- 0.1193	High potency
4c (KPGC01S138)	no inhibition	0.5786 +/- 0.1406	Low potency
4d (GCL.2)	no inhibition	0.1821 +/- 0.1514	Low potency
4e (KPGC01S42)	1261 nM (poor inhibition)	0.5280 +/- 0.1259	Low potency

*denotes BMS high potency compound

DISCUSSION AND CONCLUSION

Cancer immunotherapy marks a major step in treating cancer and the development of PD-1/PD-L1 immune checkpoint inhibitors have been an important area of research for treatment of several tumors. Currently, six therapeutic antibodies targeting both PD-1 (pembrolizumab, nivolumab, and cemiplimab) and PD-L1 (atezolizumab, durvalumab, and avelumab) have been approved by U.S. FDA. Recently, several new small molecules PD-1/PD-L1 inhibitors have been developed⁴⁸ along with structure determination of human PD-1/PD-L1 complex and co-crystals of inhibitory ligands^{49–51}. Still the field is very active in search for new small molecules to inhibit this important checkpoint and we hope to enhance the speed of this search with the use of new structure-based ML methods that have been benchmarked extensively and tested prospectively.

We have developed a new ML methodology, EGNN, based on combining local features of the small molecule topology and global features of the small molecule interacting within the binding pocket as energetic scores to select, synthesize and experimentally validate potent inhibitors of PD-1/PD-L1 interaction. Specifically, EGNN outperforms traditional ML architectures, such as, RF, SVM that include both local and global features, as well as the GNN model that uses only local features of small molecular topology. When benchmarked with known PD-1/PD-L1 inhibitors from BMS and Incyte patent data, we concluded that topology of the small molecule, the structural interaction in the binding pocket, and chemical diversity of the training data are all important considerations for enhancing model performance.

We used a bootstrapped EGNN model (based on 1000 EGNN models) for prediction and confident selection of new molecules for chemical synthesis and subsequent testing of inhibition using HTRF PD-1/PD-L1 inhibition assay. We believe that bootstrapping is an important statistical technique to use with ML methods to confidently select molecules for experimental validation in drug design. The predicted high potency molecule, (4-((3-(2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)-2-methylbenzyl)oxy)-2,6-dimethoxybenzyl)-D-serine, is a hybrid of two BMS high potency molecular scaffolds, and has an IC₅₀ value of 339.9 nM for inhibiting PD-1/PD-L1 interaction, suggesting the ability of EGNN model to do scaffold hopping to identify new inhibitors. Accurate selection of low potency molecules with different scaffolds suggests practical utility of our bootstrapped model for selection of compounds for synthesis, a hard problem in the field of ML based drug design.

Our EGNN methodology can be further developed with the addition of more chemically diverse data, and incorporating reinforcement iterative learning with experiments performed in

each step for developing a library of structurally diverse small molecule inhibiting PD-1/PD-L1 interaction to guide structure-activity relationships. Given the general nature of the machine learning model and docking methodology that is readily available for use, this approach can be adapted to identify small molecule immunomodulators by targeting other immune checkpoints, as well as, generally used to include local and global features for target-based drug design.

EXPERIMENTAL SECTION

Homogenous Time-Resolved Fluorescence (HTRF) Assay to Test Inhibition of Predicted Compounds

Inhibition of PD-1/PD-L1 interaction was tested for 4 high and low potent predicted compounds using the PD1/PD-L1 HTRF assay kit from Cisbio US, Inc. The assay protocol was used as mentioned in the kit for each predicted compound (**4b**, **4c**, **4d** and **4e**) and the BMS control compound (**4a**). Briefly, 2 μ L of the compound, 4 μ L from a 25 nM Tag1-PD-L1 protein solution and 4 μ L from a 250 nM Tag2-PD1 protein were added into a Cisbio's HTRF 96-well low volume white plate. Then, the plate was incubated for 15 minutes at room temperature. Next, 10 μ L from pre-mixed anti-tag detection reagents (5 μ L from 1X anti-Tag1-Eu³⁺ and 5 μ L from 1X anti-Tag2-XL665) were added and the sealed plate was incubated for 2 hours at room temperature. Finally, the plate sealer was removed, and measurements were taken using a HTRF[®] compatible reader. This protocol used 12 different concentrations of each compound where the maximum and minimum assay concentrations are 10,000 nM and 0.001 nM respectively. Several replicates at different concentration were done for high potent prediction compound **4b** (36 data points) and positive control compound **4a** (48 data points). The fitted curve for normalized signal denoted by

$\Delta F/\Delta F_{max}$ (calculated using HTRF ratio 665 nm/620 nm) versus log[concentration] was used to determine the 50% inhibitory concentrations (IC_{50}) of the compounds (see next section on **Calculation of IC_{50} values**).

To calculate $\Delta F/\Delta F_{max}$, first the HTRF ratio is calculated as follows;

$$HTRF\ ratio = \frac{Signal\ 665\ nm}{Signal\ 620\ nm} \times 10000$$

A multiplication factor of 10000 factor was used to not deal with decimal values that improves data accuracy during calculation. The ΔR ratio indicating “specific signal” of the compound disrupting the PD-1/PD-L1 interaction was calculated by subtracting background HTRF ratio (negative DMSO control in our work) from each compound (sample) HTRF ratio as follows;

$$\Delta R = HTRF\ ratio\ (sample) - HTRF\ ratio\ (background)$$

Next, data normalization was done to minimize variation in values on different days, different plate reader instruments, or if the assay was done by different individuals. The normalization was done with respect to the background HTRF ratio and was calculated as follows;

$$\Delta F = \frac{HTRF\ ratio\ (sample) - HTRF\ ratio\ (background)}{HTRF\ ratio\ (background)} \times 100\%$$

Finally, the $\Delta F/\Delta F_{max}$ ratio was calculated to enable comparison of values between multiple experiments.

$$\Delta F/\Delta F_{max} = \frac{\Delta F\ (sample)}{\Delta F\ max}$$

where $\Delta F\ max$ is taken as the ΔF of the positive DMSO control in the assay.

Calculation of IC₅₀ values

The IC₅₀ value for PD-1/PD-L1 inhibition was determined by analyzing the log of the concentration–response curves to fit a sigmoid curve with four-parameter logistic (4PL) regression using the GraphPad Prism Software version 8.3.0 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. The IC₅₀ values are provided in **Table 2**. Following equation defines the regression curve.

$$Y = Bottom + \frac{(Top - Bottom)}{(1 + 10^{((LogIC50 - X) \times HillSlope)})}$$

where X = Log of concentration, Y = Response $\Delta F / \Delta F_{max}$, Top and $Bottom$ = Plateaus in same units as Y , $LogIC50$ = Same log units as X , $HillSlope$ = Slope factor or Hill Slope, Unitless. Using the above equation, $LogIC50$, is calculated to obtain the IC₅₀ value for each compound. The **HTRF_IC50_Data.xlsx** data file with all replicates is provided as a Supporting File for use in GraphPad Prism Software to calculate IC₅₀ values.

Machine Learning Architecture of the EGNN model

The EGNN model was developed using PyTorch⁵². All scripts for implementing the machine learning model and results are provided on GitHub at <https://github.com/chopralab/egnn>. The **Figure 1** shows the overview of the EGNN machine learning architecture. We implemented the Graph Neural Networks for the molecular graph by Tsubaki and coworkers.²⁴. Briefly, the molecular structures were converted into SMILES strings using ChemAxon MolConverter⁵³ software. Then RDKit⁵⁴ software package and the Weisfeiler-Lehman algorithm was used to

extract r -radius subgraphs for molecules (**Figure 1A**). The following sections include details of the EGNN architecture.

Graph Neural Network for Molecular Graphs in EGNN

The following equations and notations with details for molecular GNN have been reproduced here from the original work²⁴ with minor modifications for clarification. The lowercase bold face letters (e.g. $\mathbf{v} \in \mathbb{R}^d$) indicate vectors, uppercase bold face letters (e.g. $\mathbf{M} \in \mathbb{R}^{m \times n}$) indicate matrices, and Italicized non-bold letters (e.g. S, G, v , and e) indicate scalars, sets, graphs, vertices, and edges. The GNN converts a molecular graph into a low dimensional real valued vector $\mathbf{y} \in \mathbb{R}^d$ with two neural network-based functions; transition and output.²¹ In a graph G , each vertex (v) is updated with considering the information of its neighboring vertices and edges by the transition function. These vertices have been mapped into a real valued vector $\mathbf{y} \in \mathbb{R}^d$ by the output function. Both functions are differentiable. All the input features and weights of the GNN model are updated using back propagation with the help of the cross-entropy loss function.

A graph can be defined as $G = (V, E)$, here; V and E are sets of vertices and edges respectively. When applied to chemistry, atoms can be defined as vertices and chemical bonds can be defined as edges. First, all the atoms and chemical bonds will be embedded as real valued vectors with d -dimensions based on their different types. Since the diversity of atoms (eg: C, N, O, etc.) and bonds (eg: single bonds, double bonds, triple bonds, etc.) in a small molecule is limited, the number of learning parameters are limited. Therefore, a strategy called r -radius sub-graphs⁵⁵ was used to avoid this limitation.

r -radius Sub-graphs

The set of all atoms within a defined r radius an atom i can be represented as $N(i, r)$. When the $r = 0$, $N(i, r) = \{i\}$, which is the set of all atoms in the molecule. The r -radius sub-graph of the i th vertex (v_i) is defined as follows;

$$v_i^{(r)} = (V_i^{(r)}, E_i^{(r)})$$

Here,

$$V_i^{(r)} = \{v_j \mid j \in N(i, r)\}$$

$$E_i^{(r)} = \{e_{mn} \in E \mid (m, n) \in N(i, r) \times N(i, r - 1)\}$$

The r -radius sub-graph for the edge between i th and j th atoms (e_{ij}) was defined as follows;

$$e_{ij}^{(r)} = (V_i^{(r-1)} \cup V_j^{(r-1)}, E_i^{(r-1)} \cap E_j^{(r-1)})$$

Randomly initialized embeddings (**Figure 1**) are assigned to each r -radius edge $e_{ij}^{(r)}$ and vertex ($v_i^{(r)}$) based on the type. Backpropagation has been used to train these random embeddings.

Vertex Transition Function

Say $\mathbf{v}_i^{(t)} \in \mathbb{R}^d$ is the embedded vector for the i th vertex of a given molecular graph G at time step t . Then the updated $\mathbf{v}_i^{(t+1)} \in \mathbb{R}^d$ vector can be written as follows;

$$\mathbf{v}_i^{(t+1)} = \sigma \left(\mathbf{v}_i^{(t)} + \sum_{j \in N(i)} \mathbf{h}_{ij}^{(t)} \right)$$

Here, $N(i)$ is denoting the set of neighboring atoms, σ is the sigmoid function which is defined as $\sigma(x) = \frac{1}{1+e^x}$, and $\mathbf{h}_{ij}^{(t)} \in \mathbb{R}^d$ is the hidden vector which defines the neighborhood and can be calculated as follows;

$$\mathbf{h}_{ij}^{(t)} = f \left(\mathbf{W}_{neighbor} \begin{bmatrix} \mathbf{v}_j^{(t)} \\ \mathbf{e}_{ij}^{(t)} \end{bmatrix} + \mathbf{b}_{neighbor} \right)$$

Here, f is the Rectified Linear Units (ReLU), a non-linear activation function such $f(x) = \max(0, x)$. $\mathbf{W}_{neighbor} \in \mathbb{R}^{d \times 2d}$ and $\mathbf{b}_{neighbor} \in \mathbb{R}^{d \times 1}$ are the weight matrix and the bias vector respectively. The vector between the i th and j th atoms (vertices) of the molecular graph after the time step t is defined as $\mathbf{e}_{ij}^{(t)}$.

Edge Transition Function

As mentioned before, edge transition function is used to update each embedded edge vector $\mathbf{e}_{ij}^{(t)}$ during the training process.

$$\mathbf{e}_{ij}^{(t+1)} = \sigma \left(\mathbf{e}_{ij}^{(t)} + \mathbf{g}_{ij}^{(t)} \right)$$

$$\mathbf{g}_{ij}^{(t)} = f \left(\mathbf{W}_{edge} \left(\mathbf{v}_i^{(t)} + \mathbf{v}_j^{(t)} \right) + \mathbf{b}_{edge} \right)$$

Here, $\mathbf{W}_{edge} \in \mathbb{R}^{d \times d}$, $\mathbf{b}_{edge} \in \mathbb{R}^{d \times 1}$ are the weight matrix and the bias vector respectively.

Moreover, $\mathbf{v}_i^{(t)}$ and $\mathbf{v}_j^{(t)}$ are added, because there is no direction for edges in molecular graphs.

Molecular Vector Output of Molecular GNN

The transition function generates an updated set of atom (vertex) vectors ($V = \{\mathbf{v}_1^{(t)}, \mathbf{v}_2^{(t)}, \dots, \mathbf{v}_{|V|}^{(t)}\}$). Then the output function uses this set of atom vectors to obtain an unique molecular vector $\mathbf{y}_{molecule} \in \mathbb{R}^d$ (**Figure 1A**), which is defined as follows;

$$\mathbf{y}_{molecule} = \frac{1}{|V|} \sum_{i=1}^{|V|} \mathbf{v}_i^{(t)}$$

Here, the total number of vertices in the full molecular graph is denoted by the $|V|$.

Generation of Energy Features with Docking and Energy Vector (E) in EGNN

First, all the reported molecules were carefully drawn using MarvinSketch⁵⁶ software. Then, all the drawn molecules were cleaned in 3D and converted into a sybyl.mol2 file, which was used for docking with our in-house CANDOCK³¹ software package (version 0.6.0) using default parameters with 20,000 max num possible and 20% top seed percent (**Figure 1B**). CANDOCK source code is available on GitHub at <https://github.com/chopralab/candock/releases/tag/v0.6.0>. The docking was done with a PD-L1 homodimer crystal structure (PDB ID: 5N2F). We selected the binding site based on the coordinates of the crystal ligand in the protein structure (ligand ID:8HW). Then, radial-mean-reduced-6 (RMR6)³¹ was used as “Selector” parameters for docking to select the top pose³¹. Next, the top pose of each docked compound was selected, and its docking score was recalculated using all the available 96 different potential energy functions in CANDOCK³¹ software. All 96 CANDOCK docking energy scores of each molecule were normalized for each potential energy function to use as a vector in the EGNN model;

$$\hat{S}_{i,j} = \left(\frac{S_{i,j} - \min(S_j)}{\max(S_j) - \min(S_j)} \right)$$

Where, $i: 1 \rightarrow n$ and $j: 1 \rightarrow m$. Here, n is the number of potential energy scoring functions and m is the number of molecules in the dataset. $\hat{S}_{i,j}$ is the normalized docking energy value for the energy score with i th potential energy function for the j th docked molecule. Similarly, $S_{i,j}$ is the docking energy score before normalization. Also, $\max(S_j)$ and $\min(S_j)$ are the maximum and minimum energy values within the j th scoring function for all docked molecules. Then Cohen's Kappa scores were calculated for each scoring function for all the training set data using `Cohen_kappa_score` tool in `scikit-learn` package⁴⁶. All the scoring functions which gave a positive Cohen's kappa score were selected and top in each class was selected for the EGNN model. Thus, the normalized docking score vector for each molecule in the EGNN model is represented using RCR15 and RCC15 normalized potential energy scoring functions as $\mathbf{y}_{energy} \in \mathbb{R}^2$ (**Figure 1B**).

Output of EGNN

As represented in **Figure 1C**, the normalized docking energy score vector (\mathbf{y}_{energy}) is concatenated with the molecular vector output of the GNN ($\mathbf{y}_{molecule}$). Then, the concatenated long vector ($\mathbf{y}_{molecule} \oplus \mathbf{y}_{energy}$) $\in \mathbb{R}^{(d+2)}$ was used for the training as follows to obtain an output vector $\mathbf{x}_{output} \in \mathbb{R}^2$;

$$\mathbf{x}_{output} = \mathbf{W}_{output}(\mathbf{y}_{molecule} \oplus \mathbf{y}_{energy}) + \mathbf{b}_{output}$$

Here \oplus denotes concatenation, $\mathbf{W}_{output} \in \mathbb{R}^{2 \times (d+2)}$ denotes the weight matrix and the $\mathbf{b}_{output} \in \mathbb{R}^2$ denotes the bias vector. Then, a SoftMax classifier (**Figure 1D**) is added on to the top of the $\mathbf{x}_{output} = [y_0, y_1]$ vector to get the high or low potency probabilities.

$$p_t = \frac{e^{(y_t)}}{\sum_i e^{(y_i)}}$$

Here, $t \in \{0,1\}$; 0 indicates low potency and 1 indicates high potency, and the p_t is the probability of the given y_t .

Bootstrapping the EGNN model

Bootstrapping which carried out random sampling with replacement was used with the final model to get predictions. One hundred different models with different sampled training data (sample size of 60) were trained and predictions were obtained for an in-house molecular designs test set. Averaged SoftMax scores were used as the final prediction results of the bootstrapped model. Finally, molecules in the synthetic test set were classified as high Potency or low potency based on the averaged SoftMax score. If it is greater or equal to 0.5, it was considered as high Potency, else low potency (**Figure 1**). Thus, the EGNN model was trained with back propagation with given SMILES strings, the vectors of RCR15 and RCC15 scores generated by CANDOCK³¹ and their high potency or low potency status with the PD-L1 protein. The trained model can be used to predict the probability of a given molecule to be a high or low potent molecule towards the PD-L1 protein.

EGNN Training and Hyperparameter Optimization

The model takes a SMILES string and a docking energy score string for a given molecule as inputs. Hyperparameters of the model were optimized before using it for predictions. Dimension of the GNN hidden vector (dim), number of hidden layers of the GNN, and sub-graph radius were optimized by considering the five-fold cross validated F1 score. Three values were used for the

dimension of the GNN hidden molecular vector output (i.e. dim = 5, 10 and 15). Numbers 1, 2, and 3 were used to check for the optimum number of hidden layers in the GNN. Finally, the optimum sub-graph radius for the model was selected out of radius = 1, 2 and 3.

Calculation of the F1 Score and the Cohen's Kappa

Following terms were used to calculate the Cohen's Kappa and the F1 score. The number of compounds predicted to be high potency that experimentally reported to be high potency was considered as true positives (TP). When the number of compounds predicted to be high potency but experimentally reported as low potency in patents were taken as false positives (FP). True negatives (TN) are defined as the number of compounds predicted to be low potency and experimentally reported as low potency as well. Then false negatives (FN) are defined as the number of compounds predicted to be low potency but experimentally reported as high potency. F1 score is defined as follows.

$$F1\ Score = \frac{2 \cdot precision \cdot recall}{precision + recall}$$

Here the precision and recall are defined as follows.

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

Cohen's Kappa:

$$Cohen's\ Kappa = \frac{P_o - P_e}{1 - P_e}$$

P_o = Relative observed agreement among raters

$$P_o = \frac{TP + TN}{TP + TN + FP + FN}$$

P_e = Probability of random agreement

$$P_e = \frac{[TP + FN \quad FP + TN] \begin{bmatrix} TP + FP \\ FN + TN \end{bmatrix}}{(TP + TN + FP + FN)^2}$$

Synthesis

Unless noted otherwise, all reagents and solvents were purchased from commercial sources and used as received. All reactions were performed in a screw-capped vial. The proton (^1H) and carbon (^{13}C) NMR spectra were obtained using a 500 MHz using Me_4Si as an internal standard and are reported in δ units. Coupling constants (J values) are reported in Hz. Column chromatography was performed on silica gel using flash chromatography (Teledyne ISCO EZprep). High-resolution mass spectra (HRMS) were obtained using the electron spray ionization (ESI) technique and as TOF mass analyzer. Organic solvents and starting materials were used as received. The BMS compound **4a** (BMS-1 or KPGC01S94)⁷ as well as compounds **4b-c** were synthesized according to the reported procedures starting from compound **1**, **2a-b**, **3a-b** and spectral data were in accordance with reported data.⁶⁻⁸

Compound 4a (BMS-1 or KPGC01S94), (S)-1-(2,6-dimethoxy-4-((2-methyl-[1,1'-biphenyl]-3-yl)methoxy)benzyl)piperidine-2-carboxylic acid: Compound **3b** from scheme 1 (45 mg, 0.125 mmol), (S)-piperidine-2-carboxylic acid (64.5 mg, 4 equiv, 0.5 mmol), sodium cyanoborohydride

(40.8 mg, 5.2 equiv, 0.65 mmol), were dissolved in DMF (1 mL) and then added acetic acid (2 drops). The reaction mixture was allowed to stir at 80 °C for 1 hour. The reaction was monitored by TLC. The crude was purified by 0-20% MeOH:DCM to afford desire product as an off-white solid (31.5 mg, 53% yield). ¹H NMR (500 MHz, DMSO-*d*₆): δ 7.49 – 7.41 (m, 3H), 7.39 – 7.34 (m, 1H), 7.32 – 7.25 (m, 3H), 7.19 (dd, *J* = 7.7, 1.5 Hz, 1H), 6.41 (s, 2H), 5.17 (s, 2H), 4.08 (s, 2H), 3.78 (s, 7H), 3.11 (t, *J* = 5.5, 5.5 Hz, 1H), 3.08 – 2.99 (m, 1H), 2.60 (dd, *J* = 13.5, 6.7 Hz, 1H), 2.20 (s, 3H), 1.80 (q, *J* = 6.0, 5.9, 5.9 Hz, 2H), 1.55 (q, *J* = 6.7, 6.1, 6.1 Hz, 2H), 1.37 (ddt, *J* = 18.4, 12.8, 6.5, 6.5 Hz, 3H); ¹³C NMR (126 MHz, DMSO-*d*₆): δ 172.0, 161.47, 160.32, 142.70, 141.85, 135.83, 134.59, 130.25, 129.63, 128.86, 128.72, 127.44, 126.04, 92.05, 69.23, 64.35, 56.42, 48.70, 46.21, 31.16, 26.11, 22.13, 21.27, 16.41.

Compound 4b (KPGC01S32), (4-((3-(2,3-dihydrobenzo[*b*][1,4]dioxin-6-yl)-2-methylbenzyl)oxy)-2,6-dimethoxybenzyl)-D-serine: Compound **3a** from scheme 1 (35.6 mg, 0.104 mmol), D-serine (32.8 mg, 3 equiv), sodium cyanoborohydride (19.6 mg, 3 equiv), were dissolved in DMF (1 mL) and then added acetic acid (0.104 mmol, 1 equiv, 2 drops). The reaction mixture was allowed to stir overnight at room temperature. The reaction was monitored by TLC. The crude was purified by 0-20% DCM: MeOH to afford desire product as an off-white solid (42% yield). ¹H NMR (500 MHz, DMSO-*d*₆): δ 7.42 (dd, *J* = 7.6, 1.5 Hz, 1H), 7.22 (t, *J* = 7.6, 7.6 Hz, 1H), 7.15 (dd, *J* = 7.6, 1.5 Hz, 1H), 6.90 (d, *J* = 8.2 Hz, 1H), 6.76 (d, *J* = 2.1 Hz, 1H), 6.73 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.37 (s, 2H), 5.13 (s, 2H), 4.26 (s, 4H), 3.86 (s, 2H), 3.77 (s, 6H), 3.58 (dt, *J* = 8.5, 3.4, 3.4 Hz, 2H), 2.94 (t, *J* = 6.0, 6.0 Hz, 1H, NH), 2.20 (s, 3H); ¹³C NMR (126 MHz, DMSO-*d*₆): δ 172.85, 160.75, 159.69, 159.53, 143.42, 142.96, 142.12, 135.85, 134.96, 134.65, 130.20, 128.56, 125.93, 122.59, 118.17, 117.26,

104.65, 92.06, 69.18, 64.57, 62.51, 61.34, 56.43, 56.32, 16.41; HRMS (ESI): for $C_{28}H_{32}NO_8$ $[M + H]^+$ found, 510.2132 m/z; calculated mass, 510.2128.

Compound 4c (KPGC01S138), *N*-(2,6-dimethoxy-4-((2-methyl-[1,1'-biphenyl]-3-yl)methoxy)benzyl)-3,3,3-trifluoro-1-phenylpropan-1-amine: Compound **3b** from scheme 1 (8 mg, 0.022 mmol), 3,3,3-trifluoro-1-phenylpropan-1-amine (16.7 mg, 0.088 mmol, 4 equiv), sodium cyanoborohydride (7.2 mg, 0.114 mmol, 5.2 equiv), were dissolved in DMF (0.5 mL) and then added acetic acid (1 drop). The reaction mixture was allowed to stir at 80 °C for 3 hours. The reaction was monitored by TLC. The crude was purified by 0-20% MeOH:DCM to afford desired product as oily product (68% yield). 1H NMR (500 MHz, $CDCl_3$): δ 7.45 – 7.38 (m, 5H), 7.38 – 7.30 (m, 6H), 7.30 – 7.27 (m, 1H), 7.26 (d, J = 5.4 Hz, 1H), 6.22 (s, 2H), 5.08 (s, 2H), 3.96 (t, J = 6.5, 6.5 Hz, 1H), 3.75 (s, 6H), 3.73 (d, J = 6.5 Hz, 2H), 2.46 (s, 1H), 2.27 (s, 3H), 1.29 – 1.24 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 160.23, 159.41, 143.06, 141.94, 135.1, 134.49, 132.81, 130.34, 129.40, 128.49, 128.32, 128.11, 127.91, 127.62, 127.18, 126.90, 125.65, 91.17, 69.34, 56.21, 55.53, 39.11, 16.23; HRMS (ESI): for $C_{32}H_{33}F_3NO_3$ $[M + H]^+$ found, 536.2419 m/z; calcd mass, 536.2413.

Compound 4d (GCL.2), (7*R*,8*R*,9*S*,13*S*,14*S*,17*R*)-17-ethynyl-17-hydroxy-7,13-dimethyl-1,2,6,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-3*H*-cyclopenta[*a*]phenanthren-3-one:

Tibolone (156 mg, 0.5 mmol) was taken in a round bottom flask containing 10 mL of THF and 100 μ L of water was added. Next, *p*-toluene sulfonic acid (85 mg, 0.5 mmol) was added to it and the mixture was refluxed at 80 °C for 48 hours and the progress of the reaction was monitored by TLC. The organic solvent was then evaporated to dryness to get the crude product, which was purified by flash column chromatography using 20% ethyl acetate in pet-ether solvent mixture as eluent to give off-white solid pure compound GCL2 (53% yield). 1H NMR (500 MHz, MeOD): δ

5.80 (t, $J = 2.1, 2.1$ Hz, 1H), 2.88 (s, 1H), 2.56 (ddt, $J = 14.1, 5.4, 1.6, 1.6$ Hz, 1H), 2.42 – 2.28 (m, 4H), 2.27 – 2.19 (m, 1H), 2.18 – 2.13 (m, 1H), 2.06 – 1.90 (m, 3H), 1.77 – 1.53 (m, 6H), 1.44 – 1.25 (m, 2H), 1.14 (qd, $J = 11.0, 11.0, 10.9, 4.2$ Hz, 1H), 0.91 (d, $J = 0.7$ Hz, 3H), 0.79 (d, $J = 7.1$ Hz, 3H); ^{13}C NMR (126 MHz, MeOD): δ 201.06, 167.84, 125.32, 87.28, 78.82, 73.45, 48.12, 47.95, 47.78, 47.61, 47.44, 47.27, 47.10, 46.67, 45.82, 43.05, 42.76, 42.18, 38.28, 36.02, 32.25, 30.66, 26.44, 26.41, 21.71, 11.79, 11.77.

Compound 4e (KPGC01S42), (S)-1-(2,6-dimethoxy-4-((2-methyl-[1,1'-biphenyl]-3-yl)methoxy)benzyl)piperidine-3-carboxylic acid: 3b from scheme 1 (24 mg, 0.0066 mmol), (D)-Nipecotic acid (34.2 mg, 0.0265 mmol, 4 equiv), sodium cyanoborohydride (5.2 mg, 0.0343 mmol, 5.2 equiv), were dissolved in DMF (1 mL) and then added acetic acid (1 drop). The reaction mixture was allowed to stir at room temperature for 14 hours. The reaction was monitored by TLC (Silica, 5% DCM:MeOH. The crude was purified by flash chromatography using 0-20% DCM:MeOH to afford desire product as oily product (43% yield). ^1H NMR (500 MHz, DMSO): δ 7.49 – 7.41 (m, 3H), 7.39 – 7.34 (m, 1H), 7.32 – 7.25 (m, 3H), 7.19 (d, $J = 7.7$ Hz, 1H), 6.37 (s, 2H), 5.15 (s, 2H), 3.75 (s, 6H), 3.53 (s, 2H), 2.38 – 2.30 (m, 2H), 2.20 (s, 3H), 1.97 (s, 4H), 1.16 (s, 2H); ^{13}C NMR (126 MHz, DMSO): δ 170.81, 160.50, 160.06, 142.67, 141.87, 135.99, 134.56, 130.18, 129.63, 128.81, 128.71, 127.42, 126.01, 92.00, 69.09, 60.22, 56.27, 55.37, 55.04, 52.65, 49.05, 31.14, 26.68, 21.21, 16.40, 14.54; LCMS/MS (ESI): for $\text{C}_{29}\text{H}_{33}\text{NO}_5$ $[\text{M} + \text{H}]^+$ found, 476.3 m/z; calcd mass, 476.24

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: XXX-XXX-XXX.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gchopra@purdue.edu

ORCID

Gaurav Chopra: 0000-0003-0942-7898

Author Contributions

P.R.W., J.A.F. and K.P.J. contributed equally to this work. P.R.W. and K.P.J. synthesized and characterized all the compounds. P.R.W. performed all docking calculations. P.R.W. and J.A.F. implemented the EGNN machine learning model and analyzed model performance. P.R.W. performed all IC₅₀ experiments and analyzed the data. G.C. conceived the idea and supervised the study. All authors discussed the models, experiments, wrote, and edited the manuscript and have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by Startup Funds from Department of Chemistry at Purdue University and Integrative Data Science Initiative award to G.C. Additional support, in part, by a

NCATS Clinical and Translational Sciences Award from the Indiana Clinical and Translational Sciences Institute (UL1TR002529). The authors acknowledge the support from the Purdue University Center for Cancer Research, NIH grant P30 CA023168. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ABBREVIATIONS USED

PD-1 programmed cell death-1; PD-L1 programmed death-ligand-1; GNN graph neural network; EGNN energy graph neural network; SVM support vector machines; RF random forest; FDA, U.S. Food and Drug Administration; HTRF, homogeneous time-resolved fluorescence; PPh_3 , triphenylphosphine; DIAD, Diisopropyl azodicarboxylate; THF, tetrahydrofuran; Me_4Si Tetramethylsilane; ML, machine learning; AUROC, Area Under the Receiver Operator Characteristic.

REFERENCES

- (1) Pardoll, D. M. The Blockade of Immune Checkpoints in Cancer Immunotherapy. *Nat. Rev. Cancer* **2012**, 12 (4), 252–264. <https://doi.org/10.1038/nrc3239>.
- (2) Buchbinder, E. I.; Desai, A. CTLA-4 and PD-1 Pathways Similarities, Differences, and Implications of Their Inhibition. *American Journal of Clinical Oncology: Cancer Clinical Trials*. Wolters Kluwer Health 2016, pp 98–106.
<https://doi.org/10.1097/COC.0000000000000239>.
- (3) Sasikumar, P. G. N.; Ramachandra, M.; Naremaddepalli, S. S. S. 1, 2, 4-Oxadiazole Derivatives as Immunomodulators. Google Patents September 2017.
- (4) Sunshine, J.; Taube, J. M. Pd-1/Pd-L1 Inhibitors. *Curr. Opin. Pharmacol.* **2015**, 23, 32–38.
- (5) Imai, K.; Takaoka, A. Comparing Antibody and Small-Molecule Therapies for Cancer. *Nat. Rev. Cancer* **2006**, 6 (9), 714–727. <https://doi.org/10.1038/nrc1913>.
- (6) Chupak, L.; Ding, M.; Martin, S.; Zheng, X.; Hewawasam, P.; Connolly, T.; Xu, N.; Yeung, K.; Zhu, J.; Langley, D.; et al. Compounds Useful as Immunomodulators. WO/2015/160641, October 22, 2015.
- (7) Chupak, L. S.; Zheng, X.; Mingo, P. A. Compounds Useful as Immunomodulators. WO/2015/034820, December 3, 2015.
- (8) Guzik, K.; Zak, K. M.; Grudnik, P.; Magiera, K.; Musielak, B.; Törner, R.; Skalniak, L.; Dömling, A.; Dubin, G.; Holak, T. A. Small-Molecule Inhibitors of the Programmed Cell Death-1/Programmed Death-Ligand 1 (PD-1/PD-L1) Interaction via Transiently Induced

- Protein States and Dimerization of PD-L1. *J. Med. Chem.* **2017**, *60* (13), 5857–5867.
- (9) Zak, K. M.; Grudnik, P.; Guzik, K.; Zieba, B. J.; Musielak, B.; Dömling, A.; Dubin, G.; Holak, T. A. Structural Basis for Small Molecule Targeting of the Programmed Death Ligand 1 (PD-L1). *Oncotarget* **2016**, *7* (21), 30323–30335.
<https://doi.org/10.18632/oncotarget.8730>.
- (10) Broach, J. R.; Thorner, J. High-Throughput Screening for Drug Discovery. *Nature*. [London: Macmillan Journals], 1869- 1996, pp 14–16. <https://doi.org/10.1038/384014a0>.
- (11) Artis, D. R.; Lin, J. J.; Zhang, C.; Wang, W.; Mehra, U.; Perreault, M.; Erbe, D.; Krupka, H. I.; England, B. P.; Arnold, J.; et al. Scaffold-Based Discovery of Indeglistazar, a PPAR Pan-Active Anti-Diabetic Agent. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (1), 262–267.
<https://doi.org/10.1073/pnas.0811325106>.
- (12) Böhm, H.-J.; Flohr, A.; Stahl, M. Scaffold Hopping. *Drug Discov. today Technol.* **2004**, *1* (3), 217–224.
- (13) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* **2004**, *3* (11), 935–949. <https://doi.org/10.1038/nrd1549>.
- (14) Joachims, T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **1998**, *1398*, 137–142. <https://doi.org/10.1007/s13928716>.
- (15) Burbidge, R.; Trotter, M.; Buxton, B.; Holden, S. Drug Design by Machine Learning:

- Support Vector Machines for Pharmaceutical Data Analysis. *Comput. Chem.* **2001**, 26 (1), 5–14. [https://doi.org/10.1016/S0097-8485\(01\)00094-8](https://doi.org/10.1016/S0097-8485(01)00094-8).
- (16) Heikamp, K.; Bajorath, J. Support Vector Machines for Drug Discovery. *Expert Opin. Drug Discov.* **2014**. <https://doi.org/10.1517/17460441.2014.866943>.
- (17) Breiman, L. Random Forests. *Mach. Learn.* **2001**, 9 (1), 5–32. <https://doi.org/10.1186/1478-7954-9-29>.
- (18) Li, H.; Leung, K. S.; Wong, M. H.; Ballester, P. J. Improving Autodock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Mol. Inform.* **2015**, 34 (2–3), 115–126. <https://doi.org/10.1002/minf.201400132>.
- (19) Zhang, Q. Y.; Aires-de-Sousa, J. Random Forest Prediction of Mutagenicity from Empirical Physicochemical Descriptors. *J. Chem. Inf. Model.* **2007**, 47 (1), 1–8. <https://doi.org/10.1021/ci050520j>.
- (20) Torng, W.; Altman, R. B. Graph Convolutional Neural Networks for Predicting Drug-Target Interactions. *J. Chem. Inf. Model.* **2019**. <https://doi.org/10.1021/acs.jcim.9b00628>.
- (21) Scarselli, F.; Gori, M.; Tsoi, A. C.; Hagenbuchner, M.; Monfardini, G. The Graph Neural Network Model. *IEEE Trans. Neural Networks* **2009**, 20 (1), 61–80. <https://doi.org/10.1109/TNN.2008.2005605>.
- (22) 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 (9), 3981–3988.
<https://doi.org/10.1021/acs.jcim.9b00387>.
- (23) Ma, T.; Zhang, A. *AffinityNet: Semi-Supervised Few-Shot Learning for Disease Type Prediction*.
- (24) Tsubaki, M.; Tomii, K.; Sese, J. Compound-Protein Interaction Prediction with End-to-End Learning of Neural Networks for Graphs and Sequences. *Bioinformatics* **2019**, 35 (2), 309–318. <https://doi.org/10.1093/bioinformatics/bty535>.
- (25) Gupta, A.; Müller, A. T.; Huisman, B. J. H.; Fuchs, J. A.; Schneider, P.; Schneider, G. Generative Recurrent Networks for De Novo Drug Design. *Mol. Inform.* **2018**, 37 (1). <https://doi.org/10.1002/minf.201700111>.
- (26) Ståhl, N.; Falkman, G.; Karlsson, A.; Mathiason, G.; Boström, J. Deep Reinforcement Learning for Multiparameter Optimization in de Novo Drug Design. *J. Chem. Inf. Model.* **2019**.
- (27) Li, Y.; Zhang, L.; Liu, Z. Multi-Objective de Novo Drug Design with Conditional Graph Generative Model. *J. Cheminform.* **2018**, 10 (1). <https://doi.org/10.1186/s13321-018-0287-6>.
- (28) Esaki, T.; Ohashi, R.; Watanabe, R.; Natsume-Kitatani, Y.; Kawashima, H.; Nagao, C.; Mizuguchi, K. Computational Model to Predict the Fraction of Unbound Drug in the Brain. *J. Chem. Inf. Model.* **2019**.
- (29) Fine, J.; Chopra, G. Lemon: A Framework for Rapidly Mining Structural Information from

the Protein Data Bank. *Bioinformatics* **2019**.

<https://doi.org/10.1093/bioinformatics/btz178>.

- (30) Chopra, G.; Samudrala, R. Exploring Polypharmacology in Drug Discovery and Repurposing Using the CANDO Platform. *Curr. Pharm. Des.* **2016**, 22 (21), 3109–3123.
- (31) Fine, J.; Konc, J.; Samudrala, R.; Chopra, G. CANDOCK: Chemical Atomic Network Based Hierarchical Flexible Docking Algorithm Using Generalized Statistical Potentials. *bioRxiv* **2018**. <https://doi.org/10.26434/chemrxiv.7187540.v1>.
- (32) Kischuk, E.; Majumder, J.; Fine, J. A.; Lantz, T. C.; Dhawan, D.; Knapp, D. W.; Ratliff, T. L.; Chopra, G. Abstract 4709: Cell-Specific Gene Program-Based Small-Molecule Immunomodulators Targeting Solid-Tumor Microenvironments; 2018.
<https://doi.org/10.1158/1538-7445.am2018-4709>.
- (33) Esensten, J. H.; Helou, Y. A.; Chopra, G.; Weiss, A.; Bluestone, J. A. CD28 Costimulation: From Mechanism to Therapy. *Immunity*. 2016, pp 973–988.
<https://doi.org/10.1016/j.immuni.2016.04.020>.
- (34) Hernandez-Perez, M.; Chopra, G.; Fine, J.; Conteh, A. M.; Anderson, R. M.; Linnemann, A. K.; Benjamin, C.; Nelson, J. B.; Benninger, K. S.; Nadler, J. L.; et al. Inhibition of 12/15-Lipoxygenase Protects Against β -Cell Oxidative Stress and Glycemic Deterioration in Mouse Models of Type 1 Diabetes. *Diabetes* **2017**, 66 (11), 2875–2887.
<https://doi.org/10.2337/db17-0215>.
- (35) Ma, X.; Zhou, J.; Wang, C.; Carter-Cooper, B.; Yang, F.; Larocque, E.; Fine, J.; Tsuji, G.;

- Chopra, G.; Lapidus, R. G.; et al. Identification of New FLT3 Inhibitors That Potently Inhibit AML Cell Lines via an Azo Click-It/Staple-It Approach. *ACS Med. Chem. Lett.* **2017**, *8* (5).
<https://doi.org/10.1021/acsmchemlett.6b00468>.
- (36) Fine, J.; Lackner, R.; Samudrala, R.; Chopra, G. Computational Chemoproteomics to Understand the Role of Selected Psychoactives in Treating Mental Health Indications. *Sci. Rep.* **2019**. <https://doi.org/10.1038/s41598-019-49515-0>.
- (37) CHUPAK, L. S.; DING, M.; MARTIN, S. W.; ZHENG, X.; HEWAWASAM, P.; CONNOLLY, T. P.; XU, N.; YEUNG, K.-S.; ZHU, J.; LANGLEY, D. R.; et al. COMPOUNDS USEFUL AS IMMUNOMODULATORS. WO/2015/160641, 2015.
- (38) Lu, L.; Zhang, F.; Li, J.; Wang, H.; Xiao, K.; Wu, L.; Qian, D.-Q.; Lajkiewicz, N.; Konkol, L.; Li, Z.; et al. Heterocyclic Compounds Derivatives as PD-L1 Internalization Inducers. WO/2018/119263, June 28, 2018.
- (39) Xiao, K.; Zhang, F.; Wu, L.; Yao, W. Heterocyclic Compounds as Immunomodulators. US 2018/0273519 A1, September 27, 2018.
- (40) Willett, P. Similarity-Based Virtual Screening Using 2D Fingerprints. *Drug Discov. Today* **2006**, *11* (23–24), 1046–1053.
- (41) DeLano, W. L. *Pymol: An Open-Source Molecular Graphics Tool*; CCP4, 2002; Vol. 40.
- (42) McHugh, M. L. Interrater Reliability: The Kappa Statistic. *Biochem. Medica* **2012**, *22* (3), 276–282. <https://doi.org/10.11613/bm.2012.031>.
- (43) Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.;

- Pande, V. MoleculeNet: A Benchmark for Molecular Machine Learning. *Chem. Sci.* **2018**, *9* (2), 513–530.
- (44) Jeni, L. A.; Cohn, J. F.; De La Torre, F. Facing Imbalanced Data--Recommendations for the Use of Performance Metrics. In *2013 Humaine association conference on affective computing and intelligent interaction*; IEEE, 2013; pp 245–251.
- (45) Evdokimov, S.; Karpinski, M.; Ponomarenko, I. On a New High Dimensional Weisfeiler-Lehman Algorithm. *J. Algebr. Comb.* **1999**, *10* (1), 29–45.
- (46) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (47) Saito, T.; Rehmsmeier, M. Precrec: Fast and Accurate Precision–Recall and ROC Curve Calculations in R. *Bioinformatics* **2017**, *33* (1), 145–147.
- (48) Guzik, K.; Tomala, M.; Muszak, D.; Konieczny, M.; Hec, A.; Błaszkieicz, U.; Pustuła, M.; Butera, R.; Dömling, A.; Holak, T. A. Development of the Inhibitors That Target the PD-1/PD-L1 Interaction—A Brief Look at Progress on Small Molecules, Peptides and Macrocycles. *Molecules* **2019**, *24* (11), 2071.
<https://doi.org/10.3390/molecules24112071>.
- (49) Zak, K. M.; Kitel, R.; Przetocka, S.; Golik, P.; Guzik, K.; Musielak, B.; Dömling, A.; Dubin, G.; Holak, T. A. Structure of the Complex of Human Programmed Death 1, PD-1, and Its Ligand PD-L1. *Structure* **2015**. <https://doi.org/10.1016/j.str.2015.09.010>.

- (50) Lee, H. T.; Lee, S. H.; Heo, Y. S. Molecular Interactions of Antibody Drugs Targeting PD-1, PD-L1, and CTLA-4 in Immuno-Oncology. *Molecules*. 2019.
<https://doi.org/10.3390/molecules24061190>.
- (51) Zak, K. M.; Grudnik, P.; Magiera, K.; Dömling, A.; Dubin, G.; Holak, T. A. Structural Biology of the Immune Checkpoint Receptor PD-1 and Its Ligands PD-L1/PD-L2. *Structure*. 2017.
<https://doi.org/10.1016/j.str.2017.06.011>.
- (52) PyTorch Community. Tensors and Dynamic Neural Networks in Python with Strong GPU Acceleration. *Github* **2016**, 6.
- (53) Chemaxon, M. Molecule File Converter, Version 5.10. 1,(C) 1999– 2012 ChemAxon Ltd.
- (54) Landrum, G. A. RDKit: Open-Source Cheminformatics. 2006.
- (55) Costa, F.; De Grave, K. Fast Neighborhood Subgraph Pairwise Distance Kernel. In *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*; Omnipress; Madison, WI, USA, 2010; pp 255–262.
- (56) Csizmadia, P. MarvinSketch and MarvinView: Molecule Applets for the World Wide Web. In *Proceedings of ECSOC-3, the third international electronic conference on synthetic organic chemistry, September 1q30*; 2019; p 1775. <https://doi.org/10.3390/ecsoc-3-01775>.

Table of Contents graphic

