

An explainable multi-parameter optimization approach for *de novo* drug design against proteins from central nervous system

Navneet Bung[§], Sowmya Ramaswamy Krishnan[§] and Arijit Roy*

TCS Research (Life Sciences Division), Tata Consultancy Services Limited, Hyderabad
500081, India.

*To whom correspondence should be addressed.

§Authors with equal contribution

Abstract

Motivation: Aim of a successful drug design and development is to produce a drug which can inhibit the target protein and has a balanced physicochemical and toxicity profile. Traditionally, this is a multi-step process where different parameters such as activity, physicochemical and pharmacokinetic properties are optimized sequentially, which often leads to high attrition rate during later stages of drug design and development.

Results: We have developed a deep learning-based *de novo* drug design method which can design novel small molecules by optimizing target-specificity as well as multiple parameters (including late stage parameters) in a single step. The model predictions were explained in two ways to remediate the black box nature of deep learning models: (1) the contribution of each parameter during multi-parameter optimization was computed using an adaptation of the SHAP algorithm and (2) an explainable predictive model was used to identify functional groups responsible for the property being optimized. The proposed method was validated against the human 5-hydroxy tryptamine receptor 1B (5-HT1B), a protein from the central nervous system (CNS). Various physicochemical properties specific to CNS drugs were considered along with the target specificity and blood-brain barrier (BBB), which acts as an additional challenge for CNS drug delivery. The contribution of each parameter towards molecule design was identified. The optimized generative model was able to design similar and better inhibitors compared to known inhibitors of 5-HT1B. In addition, the functional groups of the generated small molecules that guide the BBB permeability predictive model were identified through feature attribution techniques.

Contact: roy.arijit3@tcs.com

1 Introduction

Drug design and development is a long process with low success rate (Chen et al., 2018). It has been observed that due to undesirable biological profiles, majority of the drugs fail during various stages of the drug development. For example, during hit-identification, the activity of the drug-like molecules against the target protein remains the main focus (Hughes et al., 2011), while the other parameters are mostly optimized during later stages of drug development (Wager 2016). Optimization of multiple parameters during the initial

stages of drug design, can lead to better success rate and reduction in time can be achieved. Although multi-parameter optimization (MPO) in initial stages is desirable, choosing the optimal combination of parameters to be optimized is often challenging. Various parameters for drug-like molecules are often conflicting (Vallianatou et al., 2015) because improving the parameter of interest might adversely affect other related parameters (Segall et al., 2012). Consequently, appropriate selection of the parameters to be optimized can be a multi-parameter optimization problem in itself (Segall et al., 2012). The pharmacological properties to be optimized also depend on the target tissue of interest. One classic example is the drug candidates of the central nervous system (Waggoner et al., 2016; Waggoner et al., 2011), where multiple physicochemical properties influence ADME, binding efficiency and safety. Apart from target specificity, these drugs additionally require effective blood-brain barrier (BBB) permeability (Morofuji et al., 2020). Multiple properties such as octanol-water partition coefficient (logP), molecular weight (MW), polar surface area (PSA), and hydrogen bonding (Fong et al., 2015, Wager 2010, Wager 2016) are important factors governing the BBB permeability of molecules targeting proteins of the central nervous system.

Recent advances in the field of artificial intelligence and the success of reinforcement learning techniques in molecular optimization, have shown promising results (Segler et al. 2017; Olivecrona et al., 2017; Popova et al. 2018 Stahl et al., 2019; Winter et al., 2019; Born et al., 2020; Krishnan et al., 2021a; Bung et al., 2021, Krishnan et al., 2021b; He et al., 2021; Pereira et al., 2021). There have been various methods which attempted MPO to optimize different properties of generated molecules (Stahl et al., 2019; Winter et al., 2019; Deng et al., 2020; He et al., 2021; Pereira et al., 2021). Winter et al., applied the particle swarm optimization algorithm during drug design (Winter et al., 2019) while another study has used matched molecular pairs (MMP) to learn the chemical transformation involved in molecular optimization and have tested their model's ability to simultaneously optimize logD, solubility and clearance properties (He et al., 2021). Recently, two other studies have used MPO to optimize the BBB permeability along with other properties (Deng et al., 2020; Pereira et al., 2021). Although the study by Deng et al., does not consider the BBB permeability explicitly for optimization, it uses related basic properties to model the BBB permeability, which might not capture the complexity involved in BBB permeability (Deng et al., 2020). The study by Pereira et al., optimized a BBB permeability prediction model and binding affinity model to design molecules against a target protein (Pereira et al., 2021). Most of these methods used maximum two parameter optimization. Also, there has been no attempt to rank the contributions of individual parameter towards the optimization process, which would help understand the relative importance of each parameter during the optimization.

Most deep learning based *de novo* drug design methods are ligand based which require a target-specific ligand dataset for initial training. This restricts the application of ligand-based deep neural network models against novel drug targets or cases where limited experimental data is available. In our earlier work, we have proposed a method which can overcome the issue of insufficiency in target-specific ligand dataset and design small molecules specific to novel target proteins (Krishnan et al., 2021a). In this work, the *de novo* ligand-based drug design algorithm includes optimization of multiple physicochemical and late-stage pharmacological properties along with target specificity for CNS drug candidates. The reward function of the reinforcement learning framework used in our

previous study (Krishnan et al., 2021a) was modified to adapt the method to multi-parameter optimization. The method helps confine the design and optimization process to a specific region of the chemical space and desired property range.

As a proof of concept, the method was used to design novel small molecules against the human 5-hydroxy tryptamine receptor 1B (5-HT1B) protein, which acts as a major target protein for therapeutics in the central nervous system (CNS). 5-HT1B belongs to the G-protein coupled receptor family and is the target of serotonin (5-HT). It has been implicated in cancer proliferation (Gurbuz et al., 2014), and several CNS disorders including obsessive-compulsive disorder (OCD) (Pittenger et al., 2016), depression (Tiger et al., 2018), migraine, Parkinson's disease (Zhang et al., 2008). The 3D-structure of 5-HT1B protein has been crystallized in complex with the agonist ergotamine (PDB ID: 4IAR) (Wang et al., 2013). While designing novel small molecules, various parameters like binding affinity, physicochemical properties like logP and MW, and probability of crossing the BBB were also optimized. Further, the BBB permeability prediction model was interpreted using feature attribution methods to understand the key molecular features learned by the model, which were also cross validated against known molecular features governing BBB permeability reported in literature (Ghose et al., 2012). We propose that incorporation of late-stage pharmacological property such as BBB permeability in the early-stage drug design process, can improve the success rate of drug design and development. We also propose a novel method for quantifying the contribution of each parameter during the MPO, which influences the quality of the small molecules designed by the generative model.

2 Methods

1 Dataset curation and pre-processing

The dataset for pre-training the generative model was obtained from the ChEMBL database (Gaulton et al., 2012). The molecules were represented in the SMILES format (Weininger et al., 1989) to leverage the effectiveness of recurrent neural networks (RNNs) in handling sequential data through existing natural language processing algorithms. The RDKit library in Python was used for dataset pre-processing.

2 Pre-training the generative model and predictive models

The generative model was pre-trained on a dataset of ~1.6 million SMILES strings from the ChEMBL database as mentioned in our earlier studies (Bung et al. 2021; Krishnan et al. 2021), the aim of which is to learn the features of the existing small molecules and produce novel small molecules. The use of stack-augmented memory (Joulin and Mikolov, 2015) enabled the generation of chemically valid SMILES with high accuracy of 96.6% (Krishnan et al. 2021).

The predictive models learn a mapping between the small molecules and their corresponding property values (Olivecrona et al., 2017; Popova et al., 2018). In this study, three predictive models were trained to predict the blood-brain barrier (BBB) permeability, logP and docking score against the 5-HT1B protein. The corresponding datasets for training BBB, logP and docking score were obtained from MoleculeNet (Wu et al., 2017), an earlier work (Popova et al., 2018) and Ex-CAPE DB (Sun et al., 2017) respectively. The

predictive model uses gated recurrent units (GRU) as the internal memory followed by three dense layers (Bung et al., 2021). The models were trained using mini-batch gradient descent with the Adam optimizer (Kingma and Ba, 2014).

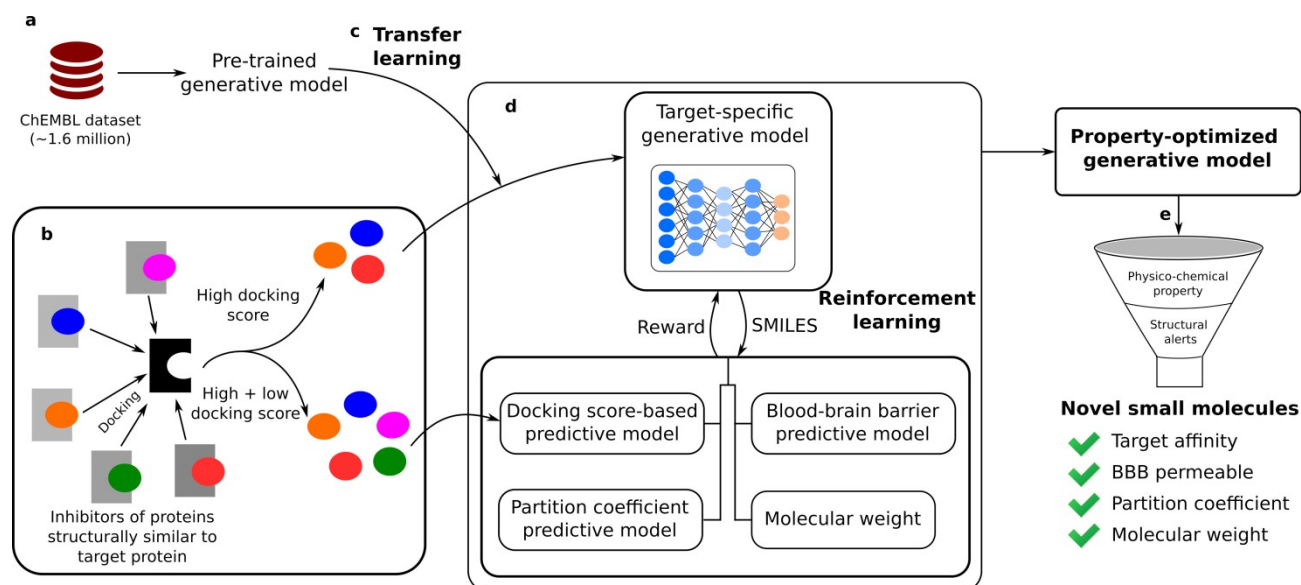


Fig. 1. Ligand-based *de novo* small molecule design. (a) Pre-trained generative model on the ChEMBL database; (b) A dataset curated from small molecules that modulate the activity of structurally related proteins; (c) Transfer learning with the curated dataset; (d) MPO using reinforcement learning; (e) Physicochemical properties and structural alerts (rule-based filters) were used to filter drug-like molecules specific to the target protein of interest.

3 Ligand-based drug design via transfer learning

A ligand-based drug design requires a target-specific ligand dataset to understand the chemical space and create more potent drug-like molecules. In most cases, there is limited or no knowledge about the small molecules that can bind to the target protein. In order to show that the proposed method can be applied against novel protein or against proteins where there is limited experimental target-specific ligand data, an initial target-specific small molecule dataset was curated from existing inhibitors of proteins whose active sites are similar to that of the target protein (step b, Fig. 1). In this case, the model was validated against the 5-hydroxy tryptamine receptor 1B (5-HT1B) protein. A dataset of 2,917 small molecules specific to other receptors in the 5-HT1 family, namely: 5-HT1A, 5-HT1D, 5-HT1E, and 5-HT1F were collected from ExCAPE-DB (Sun et al., 2017) along with their plC_{50} values. 185 known inhibitors of 5-HT1B were collected from ExCAPE-DB (Sun et al., 2017) and used only for comparing with the generated molecules. Redundancies within the training dataset (inhibitors of the homologs) and with the validation dataset (known inhibitors of 5-HT1B) were resolved by assigning a target protein to the small molecule based on the maximum plC_{50} value. The suitability of this dataset was further enhanced by docking these molecules in the active site of the target protein of interest. The molecules with high docking scores (≤ -7.0) were used to re-train the generative model to capture the molecular features specific to the target protein of interest through transfer learning (step c, Fig. 1). For this purpose, docking was performed at the 5-HT1B active site, and small molecules with a docking score less than -7.0 were chosen for training. Thus, the final training dataset for transfer learning consisted of 2,807 small molecules. During transfer learning, the weights of all the layers of the pre-trained generative model (prior network) were frozen except the last dense layer (Tan et al., 2018).

The model was trained until the inferred molecules showed an observable shift in similarity with respect to the training dataset, quantified using the Tanimoto coefficient (Lipkus, 1999).

4 Multi-parameter optimization with reinforcement learning

The generative model obtained after transfer learning was combined with the predictive model to bias the generative model towards the property space of interest using reinforcement learning (RL) (step d, Fig. 1) (Olivecrona et al., 2017). The method was modified to support simultaneous multi-parameter optimization. Four parameters were optimized including docking score, BBB permeability, logP, MW. All possible combinations of these four parameters were considered, for example single, double, triple, and maximum four combinations (table S1) were optimized to find an optimal combination. Since these parameters can conflict with each other, the percentage of newly designed molecules which satisfy all four parameters were checked after RL optimization of each combination. In addition, for each designed molecule, the physicochemical properties like topological polar surface area (TPSA), hydrogen bond donor (HBD), hydrogen bond acceptor (HBA) were computed. This helped to monitor the number of molecules that satisfy the last three physico-chemical properties, which were not considered during the RL step.

$$r(MPO) = \begin{cases} 11, & \text{if } BBB_{permeable} \\ 1, & \text{if } BBB_{impermeable} \end{cases} + 0.6 * e^{-(dockingscore)/3} + 11 * e^{-(logP-2)^2} + \begin{cases} 11, & \text{if } 0 < MW \leq 4 \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

This reward function (equation 1) helps in simultaneous optimization of all the four parameters of the generated molecules. To avoid the catastrophic forgetting behavior of the canonical policy gradient algorithm, regularization was required to keep the new policy anchored to the learned prior policy of the agent (Jaques et al., 2016; Olivecrona et al., 2017). The regularized policy gradient method was trained using mini-batch gradient descent with AMSGrad optimizer (Tran 2019). Variants of the above reward function (equation 1) with a combination of the appropriate parameters, were used to train all possible single, double and triple parameter optimized generative models after transfer learning.

The molecules obtained after RL step were further subjected to four empirical rule-based filters – PAINS, BRENK, NIH and ZINC, to remove molecules with potentially unwanted subgroups (step e, Fig. 1).

5 Evaluating the contribution of each parameter towards multi-parameter optimization

The physicochemical properties and molecular properties, including target affinity influences drug safety, activity and pharmacokinetic and pharmacodynamic properties, can

be often conflicting (Vallianatou et al., 2015; Oprea et al., 2000). To evaluate the contribution of each property during the molecule design, the SHapley Additive exPlanations (SHAP) method was used (Lundberg et al., 2017). The SHAP method is one of the most widely used explainability methods in machine learning, which calculates the contribution of different input features to the predicted output for a pre-trained model (Lundberg et al., 2017; Rodríguez-Pérez et al., 2019). Analogously, the four properties considered for the MPO were used as inputs, and the percentage of generated small molecules which passed all the property filters after the optimization was considered as the output.

Using the model before RL as the model with no input features (null model in SHAP), the SHAP contributions for each of the four parameters to the model decision were calculated (see Supplementary information, Section 1).

6 Explaining BBB permeability model predictions with gradient-based feature attribution

The Integrated gradient (IG) explainability method (Sundararajan et al., 2017) as implemented in PyTorch library, Captum (Kokhlikyan et al., 2020), was used to compute the feature attributions for the BBB permeability prediction model. The IG method computes the gradient of the model's prediction output with respect to its input features, to extract the importance attributed to every input feature by the pre-trained predictive model. The feature attributions were obtained for every character in the input SMILES string. Since the SMILES string contains both atomic symbols and tokens capturing molecular connectivity (including double bond, triple bond, branches and ring numbering), the feature attributions were extracted only for the atomic symbols. The resultant attributions were used to classify atoms in the molecule as either positively or negatively contributing to the model prediction, depending on the sign of the attribution. The results from the classification were visualized based on a coloring scheme (red atoms for positive attribution and cyan atoms for negative attribution) using RDKit.

3 Results and Discussion

The proposed method was validated by designing small molecules against the 5-hydroxy tryptamine receptor 1B (5-HT1B) receptor. The 5-HT1B receptor is responsible for several CNS disorders including obsessive-compulsive disorder (OCD) (Pittenger et al., 2016), depression (Tiger et al., 2018), migraine, Parkinson's disease (Zhang et al., 2008).

As mentioned in the Methods section, first a pre-trained generative model learnt the grammar of the SMILES. Next, transfer learning was performed on the pre-trained generative model with the training dataset of 2,807 small molecules so that the model focuses towards the chemical space specific to the target protein. Transfer learning was performed for 100 epochs by monitoring the distribution of the maximum Tanimoto coefficient of the generated molecules with respect to the validation dataset (Fig. 2a). The model obtained after transfer learning was subjected to MPO using reinforcement learning.

1 Performance of the predictive models

To optimize the generative model using reinforcement learning three predictive models, blood-brain barrier (BBB) permeability, logP and docking score specific to the 5-HT1B

protein were trained (see Methods). The task of the BBB permeability model was classification, whereas task of the docking score prediction model and the logP prediction model were regression. The ROC-AUC score was monitored for the classification task while, root mean square error (RMSE) and R^2 scores were monitored for the regression task. After extensive hyperparameter tuning, a ROC-AUC score of 0.90 was attained for the BBB permeability model. The docking score prediction model had a RMSE of 0.28 and an R^2 score of 0.82. The logP prediction model had a RMSE of 0.43 and an R^2 score of 0.91.

Table 1: Results from all MPO experiments where docking score was optimized along with other parameters. For level 1, the applied property filters were: 200 Da < MW < 600 Da, 0 < logP <= 5.0, Docking score <= -7.0, and BBB permeability = 1. And for level 2 additional property filters such as HBA <= 6, HBD <= 2 and TPSA <= 100 Å² were applied. The percentage of molecules passing each filter before the RL training phase is also provided.

Optimization experiment	Unique and valid molecules	Level 1 Filters (%)	Level 2 filters (%)
Before RL	8120	64.30	53.20
1MPO – Docking score	7856	61.12	50.9
2MPO – Docking score + MW	8930	70.7	59.2
2MPO – Docking score + BBB	8552	71.45	63.52
2MPO – Docking score + logP	8278	68.95	56.29
3MPO – Docking score + MW + BBB	8728	69.64	57.98
3MPO – Docking score + logP + BBB	8596	69.89	57.42
3MPO – Docking score + LogP + MW	8089	65.37	54.34
4MPO – Docking score + MW + logP + BBB	8646	75.06	67.63

2 Multi-parameter optimization using reinforcement learning

The choice of parameters for optimization should be such that the objectives are independent of one another. For the current study, all combination of activity, physicochemical and pharmacological properties such as docking score, logP, molecular weight and BBB permeability were considered for the optimization process (see Table S1 for all combinations). During training the deep learning model, the distributions of predicted docking score, logP, MW and BBB permeability were monitored. The generative model for the inference was selected based on the various factors such as 1) the rate of redundancy in the generated molecules, 2) optimization of the properties to the desired range and 3) the proportion of duplicates with respect to the training dataset. After reinforcement learning, 10,000 molecules were sampled from each of the 15 trained generative model (Table S1). The redundant and duplicate molecules with respect to the ChEMBL dataset were removed. For each combination of the RL optimization, the percentage of molecules that pass through all 4 parameters (predicted docking score <= -7, 200 Da < MW < 600 Da, 0 < logP <= 5.0 and BBB permeability = 1) were computed irrespective of the parameters being optimized in the particular combination. This is referred to as the level 1

filters in the Table 1 and Table S2. The same percentage was computed for the model before RL step and compared with the models with single (1-MPO) or multi-parameter (2-MPO, 3-MPO and 4-MPO) optimization. The percentage of molecules that passed other parameters ($\text{HBA} \leq 6$, $\text{HBD} \leq 2$ and $\text{TPSA} \leq 100 \text{ \AA}^2$) not considered for RL, were also computed and referred to as level 2 filters. To understand the effect of single or multi-parameter optimization on the generative model, the percentage of molecules that passed level 1 filter were considered.

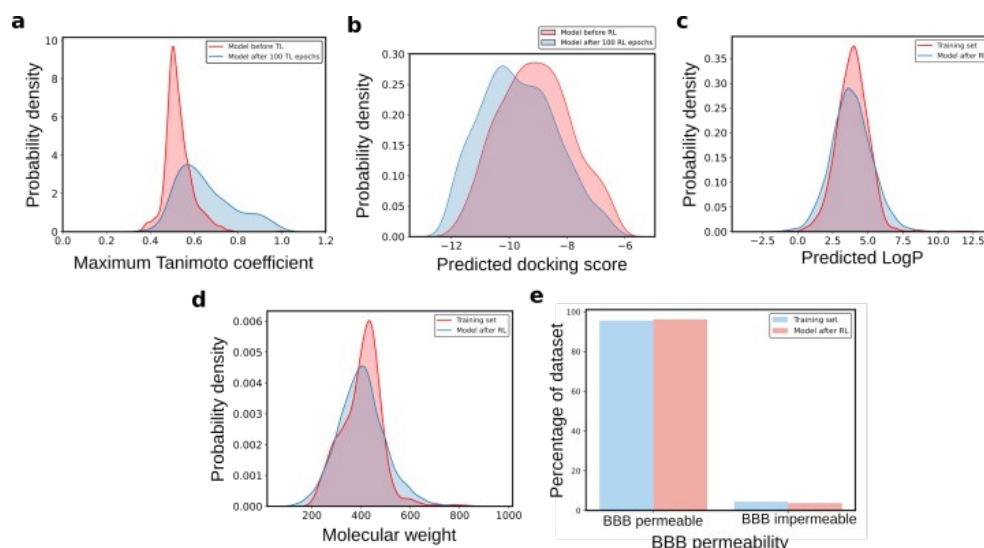


Fig. 2. Distributions monitored over the course of transfer learning (TL) and reinforcement learning (RL). (a) Maximum Tanimoto coefficient distribution before and after 100 epochs of transfer learning with respect to the validation dataset. Distribution of (b) predicted docking score, (c) predicted logP, (d) molecular weight, and (e) proportion of BBB permeable and impermeable molecules among generated small molecules and the training dataset.

Table 1 shows the percentage of molecules that passed level 1 and 2 filters with single and multiple parameter optimization involving docking score (see Supporting table S2 for all combination). While the percentage of molecules that passed level 1 filters was 64.30 before optimization (after transfer learning), the same was reduced (61.30) during docking score optimization alone. This was mainly due to the fact that in some cases, the molecules with better docking score tend to have higher MW. In all cases, single parameter (1-MPO) optimization reduced the efficiency of the model since they adversely affect other properties (see table S2). During two parameter optimization (2-MPO), three models (docking score + MW, docking score + BBB and docking score + log P) performed better compared to the single parameter optimization. All combinations (see Table S2) without docking score predictive model performed poorly. For 3-MPO, most models performed poorly or in a similar way compared to the 2-MPO models. Interestingly, from single to 3-MPO models, logP optimizations have decreased the performance of the model. Finally, it was observed that the 4-MPO model performed best.

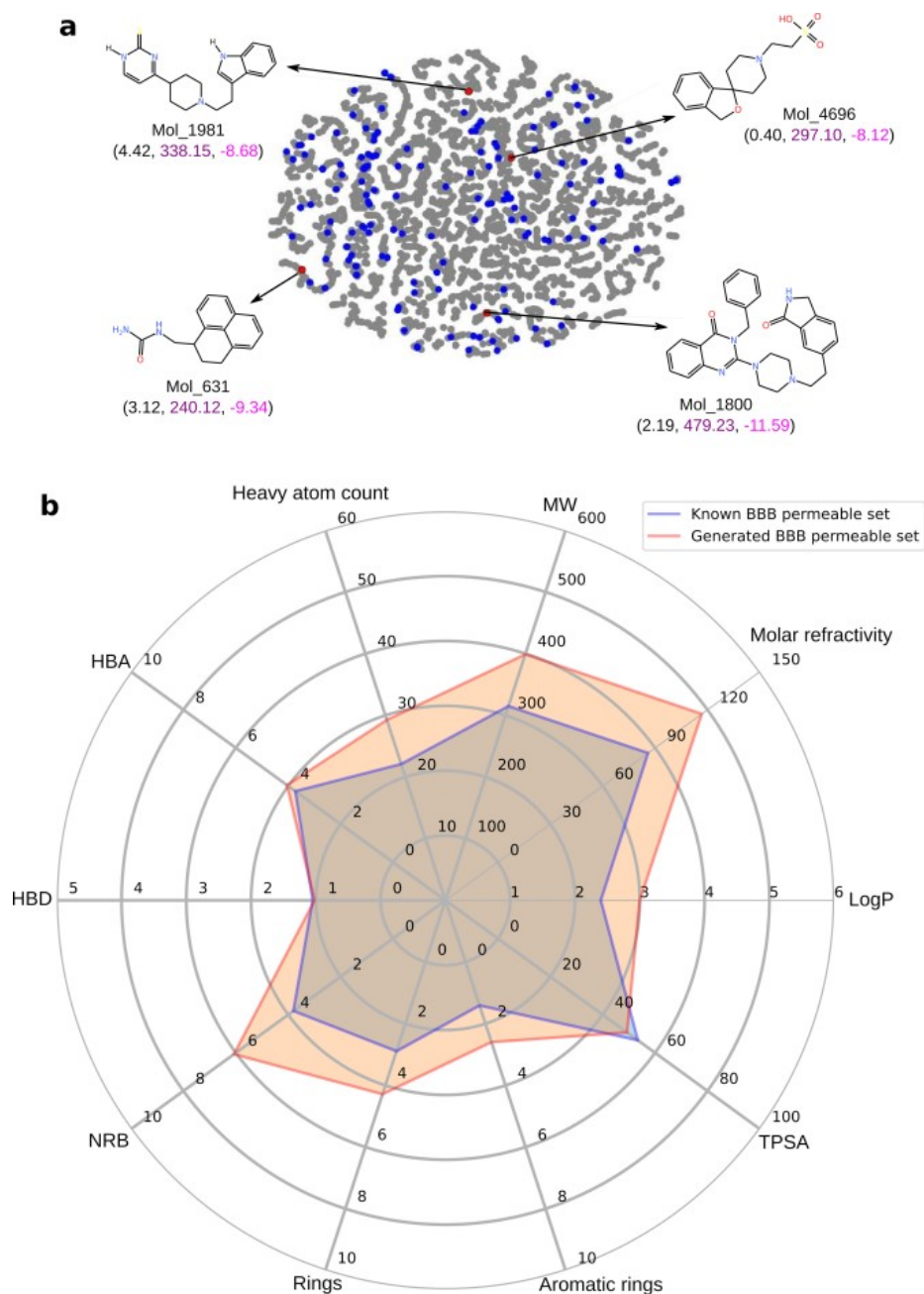


Fig. 3. Embedding of the validation dataset and the generated small molecules in the property space. The logP, MW and docking score values of the generated small molecules (grey) and the validation dataset (blue) were used to create the embedding. A selected set of small molecules occupying the distinct regions of the property space are shown in red. Their corresponding logP, MW and docking score values are shown within parentheses in black, mauve, and pink figures, respectively. b) A radial plot showing the property distribution of both known and generated BBB permeable molecules.

To understand the individual contribution of four parameters in the multi-parameter optimization process, a variant of the SHAP method was utilized (see Methods). The docking score optimization contributes the most to the optimization by increasing the filter-1 pass percentage by 7.7% (Table 2). Surprisingly, logP contributes negatively to the overall optimization. It decreases the percentage of generated small molecules passing the property filters by ~1.3%. As evident from Table 2, the contribution of four parameters can be ranked as: Docking score > BBB permeability > MW > logP (Table 2). The molecular weight and logP of the molecules generated after transfer learning was within the desired range, but the affinity of the molecules towards the target (docking score) was low. While maximizing the docking score, the molecular weight and logP deviates from the

required optimal range as the generative model samples molecules in the high range for both the properties. In such cases, including the conflicting parameter in the RL optimization step can improve the model performance as evident from 4-MPO optimization. The results show the usefulness of the SHAP method. This is the first attempt to repurpose the SHAP method for understanding the contribution of each drug design related parameters during multi-property optimization.

Table 2: Final parameter-level contributions obtained from the SHAP-based method for the multi-parameter optimization model.

Property	MPO attribution (%)
logP	-1.243
MW	1.945
Docking score	7.700
BBB permeability	2.358

As discussed above, four parameter optimization provides the highest improvement (75.06 %) in the percentage of molecules which pass the property filters (Table 1). Hence, the molecules generated after four parameter optimizations were used for further analysis in this work. The distribution of the optimized parameters with respect to the training dataset is shown in the plot (Fig. 2b-e). The docking score of the generated molecules is better when compared to the training set molecules (Fig. 2b). The generated molecules are also BBB permeable, and other physicochemical properties such as logP and MW follow a similar distribution as that of the training dataset (Fig. 2c, 2d). A lower dimensional embedding of the generated small molecule's property space is highlighted in Fig. 3a along with a few molecules from different property ranges.

The mean properties of the generated and existing BBB permeable set of molecules are shown using a radial plot (Fig. 3b). The generated small molecules not only satisfy the target specificity and BBB permeability, but also satisfy various drug-like properties. All relevant properties like heavy atom count, molecular weight, molar refractivity, logP, TPSA, aromatic rings, rings, number of rotatable bonds, hydrogen bond donor, hydrogen bond acceptor were considered. It is evident from the radial plot that the generated molecules stay within the desired property ranges although some of them were not explicitly considered during RL optimization. Possibly, in this case, some of the molecular properties remain correlated, and optimization of few properties (MW and LogP) further optimizes the related properties.

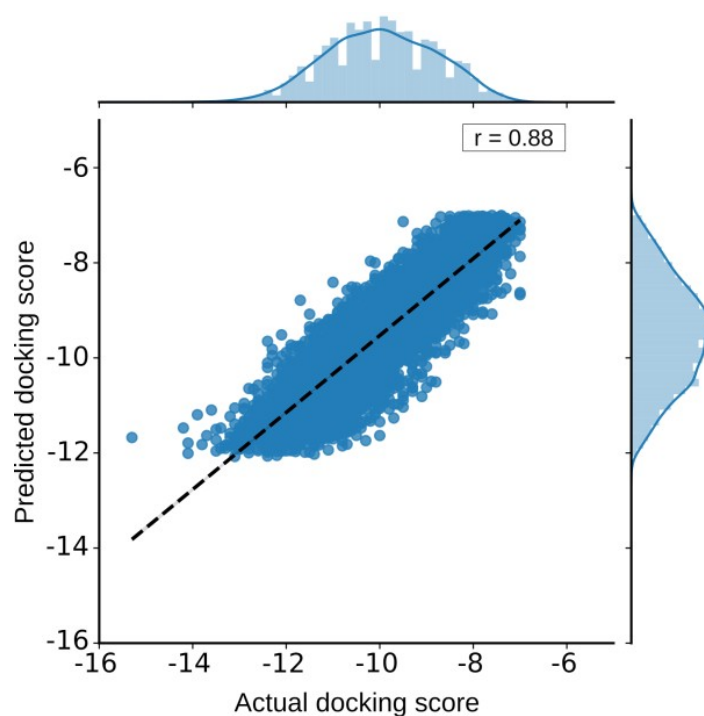


Fig. 4. Docking score comparison. The actual and predicted docking scores of the generated molecules after RL (4-MPO) are compared. The dashed black lines show the regression fit between the two datasets.

A dataset of 7125 molecules was obtained upon application of rule-based filters to remove molecules with potentially undesirable subgroups. These molecules were docked at the 5-HT1B active site using Autodock Vina (Trott et al., 2010) to obtain a final set of 7080 molecules with docking score less than -7.0. Only, 45 molecules were removed based on docking score indicating the effectiveness of reinforcement learning in optimizing the generative model. To further assess the performance of docking score predictive model, the docking scores obtained from Autodock Vina were compared with the scores obtained from predictive model. A high Pearson correlation coefficient (r) of 0.88 was observed between the actual and predicted docking scores (Fig. 4).

3 Model could reproduce molecules similar to known 5-HT1B inhibitors

The final set of 7080 molecules was compared against the validation dataset of 5-HT1B-specific molecules. It was observed that, 165 molecules from the generated set of molecules had Tanimoto coefficient above 0.75 (Lipkus, 1999) compared to the molecules from the validation dataset, indicating high similarity. Few of such highly similar molecules are shown in Supplementary figure S2. Based on the virtual screening results, some of the generated molecules with the desired property profile were also found to have better docking score than the existing inhibitors of 5-HT1B. The embedding (see Supplementary information, Fig. S3) highlights the molecule with the highest docking score and molecules which show high similarity compared to the training and validation datasets.

4 Analysis of the fragment distribution

The sub-structural fragment distribution was analyzed to understand whether the generated small molecules captured the features of the validation dataset. The generated molecules showed the presence of tertiary amines and secondary amines in accordance

with the natural ligand (serotonin, which contains an amine), of 5-HT1B receptor (Matzen et al., 2000). Also, bicyclic and tricyclic groups containing aromatic heterocyclics (anilines, piperazines, piperidines and indoles) were more frequently observed, which is also in alignment with their well-known capability to act as selective 5-HT1B inhibitors (Matzen et al., 2000; Halazy et al., 2005). These results indicate that the generative model was able to capture and generate molecules with features specific to inhibit the 5-HT1B protein and enhance the binding affinity using transfer learning and reinforcement learning.

5 Interactions with 5-HT1B binding site

The generated small molecules were docked in the active site of 5-HT1B receptor to study the interactions that stabilize the complex. It was observed that the molecule with highest binding affinity (Mol_7422) binds in the same pocket as ergotamine (Fig. 5a). Residues Q41 and D352 form hydrogen bonds with Mol_7422 molecule, while residues D129 and T134 which are proposed to be important for anchoring the small molecule in the pocket are in the vicinity of a nitrogen atom with a positive charge (Fig. 5b). The positively charged nitrogen atom is commonly observed in the small molecules that are targeted towards 5HT family of receptors (Wang et al., 2013; Peng et al., 2018). The Mol_7422 is in the pocket formed by several hydrophobic residues such as I180, A216, W327, F330, F331, L348 and F351.

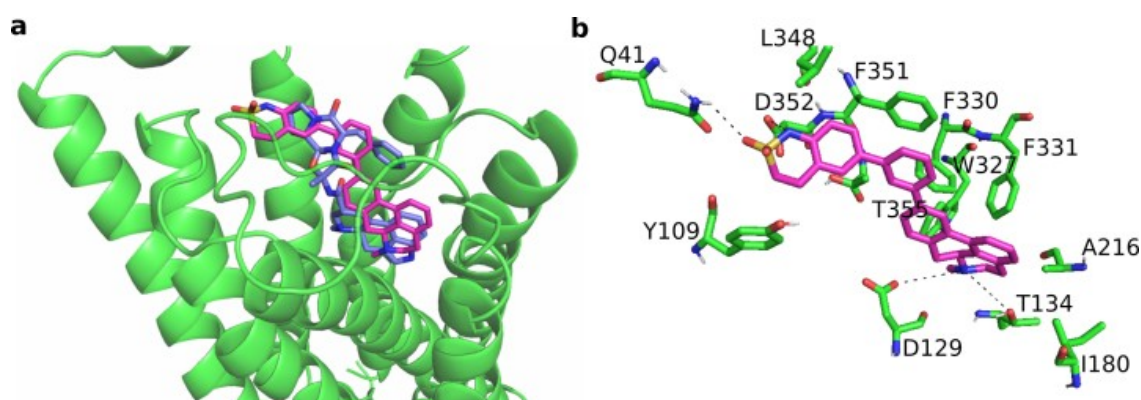


Fig. 5. Interactions of *de novo* generated molecule with 5-HT1B receptor. (a) Binding pocket of agonist ergotamine (blue sticks) and Mol_7422 (magenta sticks) molecules. The receptor is shown in green sticks. (b) Residues of 5-HT1B receptor interacting with the Mol_7422 molecule (magenta) are shown in green sticks.

6 Feature attribution method rationally explain the importance of enriched fragments in BBB permeable molecules

To understand the rationale behind the BBB permeability model predictions, the Integrated gradient (IG) feature attribution method (Sundararajan et al., 2017) was utilized. Ghose et al., have computed the fragments that are predominantly present in CNS drugs when compared to non-CNS drugs and vice versa (Ghose et al., 2012). The feature attribution obtained from the BBB predictive model were compared with the fragments that are predominant in the CNS drugs (Ghose et al., 2012). It was observed that while classifying the molecules as BBB permeable, the predictive model indeed considered the fragments which are preferred in the CNS drugs. Figure 6, shows some generated molecules that are classified as BBB permeable. It is known that the fragments like 1-ethylpiperidine, pyrrole, 1-ethylpiperazine, diphenylmethane contribute towards increasing the probability of crossing the BBB (Ghose et al., 2012). Similar trend was observed for the BBB

impermeable molecules where the fragments remain predominant with enriched fragments found in non-CNS drugs (e.g. acetamide, phenethylamine, phenetole and acetate groups). These fragments contribute towards decreasing the probability of BBB permeability (Fig. 6). A complete list of fragments along with the percentage of molecules with important fragments identified by the explainability method is provided in supporting information (Table S3 & S4).

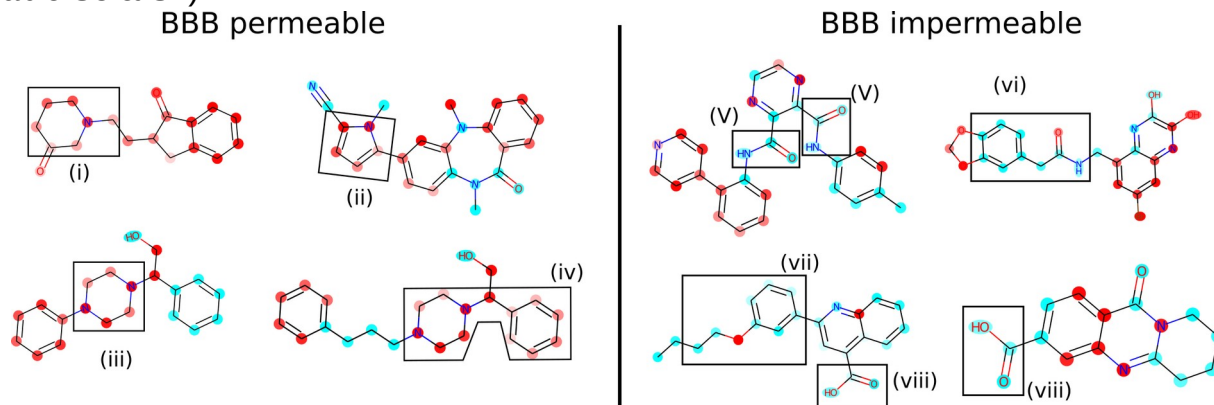


Fig. 6. Subset of small molecules showing the fragments that are enriched in BBB permeable molecules ((i) 1-ethylpiperidine, (ii) pyrrole, (iii) 1-ethylpiperazine, (iv) diphenylmethane) and in BBB impermeable molecules ((v) acetamide, (vi) Phenethylamine, (vii) Phenetole and (viii) acetate) as reported by Ghose et al. (Ghose et al., 2012). The red color indicates positive contribution, while the cyan color indicates negative contribution towards BBB permeability probability.

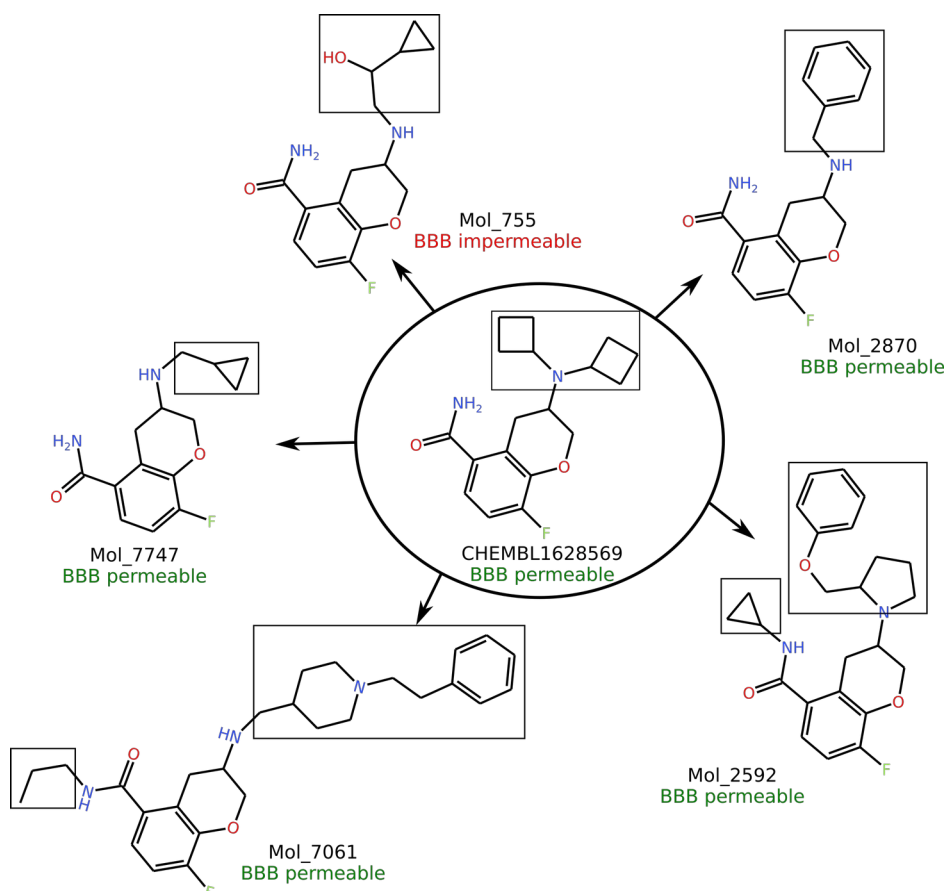


Fig. 7. Generated molecules that are similar to Robalzotan (CHEMBL1628569). The Robalzotan molecule is known to cross the BBB. Different fragment substitutions to the Robalzotan that make them BBB permeable or impermeable are highlighted using box.

Figure 7 shows that the generative model learnt to substitute fragments which enhance their predicted BBB permeability probability. For example, a compound from the training dataset (ChEMBL1628569) known as NAD-299 or Robalzotan was part of the training dataset. This is known to be a potent BBB permeable lead compound targeting the 5-HT_{1A} receptor (Johansson et al., 1997; Jerning et al., 1998). Some generated molecules were found to be similar to Robalzotan with some difference in their functional groups and varied predicted BBB permeability score (Fig. 7). The fragment substitutions made were found to correlate with the observation that, a balance between polar and non-polar substitutions is necessary to achieve optimal BBB permeability. The presence of the hydroxyl group in the designed molecule, Mol_755 (in Fig.7) makes it more polar. This was predicted to have a low BBB permeability probability of 0.39 (Fig. 7). On the other hand, the substitutions of aromatic, alkyl, and cyclic groups were found to increase the BBB permeability probability to over 0.95, although the overall similarity among the molecules was very high.

4 Conclusion

We have proposed a ligand-based *de novo* drug design algorithm for generating small molecules with multiple optimized parameters, crucial for improving the drug-like molecules. Traditionally, some of these parameters are optimized during later stages of drug design and development. We showed that they can be included during initial stage of drug design. For the current case study, transfer learning followed by reinforcement learning generated a focused library of molecules against 5-HT_{1b}, a target protein from central nervous system. It was also observed that four parameters (docking score, logP, MW and BBB permeability) during RL optimization gave best optimization result, where generated molecules satisfied most of the parameters important for a high-quality CNS drug. Around 168 generated molecules showed high similarity with known inhibitors of 5-HT_{1b}, although these known inhibitors were never used for training. Using a SHAP-based scoring method, the contribution of each parameter during multi-parameter optimization was explained. It was observed that the parameters can be conflicting, and favorable combination of parameters are required during multi-parameter optimization. The integrated gradients method for explainability identified fragments which were shown to be preferred for predicted BBB permeable molecules. This correlated well with the fragments enriched in FDA-approved BBB permeable drug molecules. This study showed that it is possible to consider appropriate combination of physicochemical and late-stage pharmacological properties during initial stages of drug design. Hopefully, this would significantly reduce the design cycles and improve the success rate of overall drug design and development.

Acknowledgements

We thank our colleagues Dr. Rajgopal Srinivasan and Dr. Gopalakrishnan Bulusu for helpful discussions.

Conflict of Interest: All the authors are employed at Tata Consultancy Services Ltd.

References

- Chen,J. et al. (2018) Drug discovery and drug marketing with the critical roles of modern administration. *Am. J. Transl. Res.*, 10(12), 4302–4312.
- Hughes,J.P. et al. (2011) Principles of early drug discovery. *Br. J. Pharmacol.*, 162(6):1239-49.
- Wager,T.T. et al. (2016) Central nervous system multiparameter optimization desirability: application in drug discovery. *ACS Chem. Neurosci.* 7(6):767-75.
- Vallianatou,T. et al. (2015) The Impact of Physicochemical and Molecular Properties in Drug Design: Navigation in the " Drug-Like" Chemical Space. *Adv. Exp. Med. Biol.*, 822:187-94.
- Segall,D.M. (2012) Multi-parameter optimization: identifying high quality compounds with a balance of properties. *Curr. Pharm. Des.*, 18(9):1292-310.
- Wager,T.T. et al. (2010) Defining desirable central nervous system drug space through the alignment of molecular properties, in vitro ADME, and safety attributes. *ACS Chem. Neurosci.*, 1(6):420-34.
- Morofuji,Y. and Nakagawa,S. (2020) Drug Development for Central Nervous System Diseases Using In vitro Blood-brain Barrier Models and Drug Repositioning. *Curr. Pharm. Des.*, 26(13), 1466-85.
- Fong,C.W. (2015) Permeability of the blood–brain barrier: molecular mechanism of transport of drugs and physiologically important compounds. *J. Membr. Biol.*, 248(4), 651-69.
- Segler,M.H.S. et al. (2017) Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent. Sci.*, 4(1), 120-131.
- Olivecrona,M. et al. (2017) Molecular *de-novo* design through deep reinforcement learning. *J. Cheminform.*, 9(1), 48.
- Popova,M. et al. (2018) Deep reinforcement learning for *de novo* drug design. *Sci. Adv.* 4(7), eaap7885.
- Ståhl,N. et al. (2019) Deep reinforcement learning for multiparameter optimization in *de novo* drug design. *J. Chem. Inf. Model.*, 59 (7), 3166-3176.
- Winter,R. et al. (2019) Efficient multi-objective molecular optimization in a continuous latent space. *Chem. Sci.*, 10(34):8016-24.
- Born,J. et al. (2020) Paccmann^{RL}: Designing anticancer drugs from transcriptomic data via reinforcement learning. *Presented At: International Conference on Research in Computational Molecular Biology.*, 231–233.
- Krishnan,S.R. et al. (2021) Accelerating De Novo Drug Design against Novel Proteins Using Deep Learning. *J. Chem. Inf. Model.*, 61(2), 621-630.
- Bung,N. et al. (2021) De novo design of new chemical entities for SARS-CoV-2 using artificial intelligence. *Fut. Med. Chem.*, 13(06):575-85.
- Krishnan,S.R. et al. (2021) De Novo Structure-Based Drug Design Using Deep Learning. *J. Chem. Inf. Model.* In press.
- He,J. et al. (2021) Molecular optimization by capturing chemist's intuition using deep neural networks. *J. Cheminform.*, 13(1):1-7.
- Pereira,T. et al. (2021) Optimizing blood–brain barrier permeation through deep reinforcement learning for *de novo* drug design. *Bioinformatics.*, 37:i84-92.

Deng,J. et al. (2020) Towards better opioid antagonists using deep reinforcement learning. arXiv preprint arXiv:2004.04768.

Gurbuz,N et al. (2014) Down-regulation of 5-HT_{1B} and 5-HT_{1D} receptors inhibits proliferation, clonogenicity and invasion of human pancreatic cancer cells. *PLoS One.*, 9(8), e105245.

Pittenger,C. et al. (2016) OCD is associated with an altered association between sensorimotor gating and cortical and subcortical 5-HT_{1b} receptor binding. *J. Affect. Disord.*, 196, 87-96.

Tiger,M. et al. (2018) The 5-HT_{1B} receptor – a potential target for antidepressant treatment. *Psychopharmacology.*, 235(5), 1317-1334.

Zhang,X. et al. (2008) Evidence for a role of the 5-HT 1B receptor and its adaptor protein, p11, in L -DOPA treatment of an animal model of Parkinsonism. *Proc. Natl. Acad. Sci. U S A.*, 105(6), 2163-2168.

Wang,C. et al. (2013) Structural basis for molecular recognition at serotonin receptors. *Science.*, 340(6132):610-4.

Ghose,A..K. et al. (2012) Knowledge-based, central nervous system (CNS) lead selection and lead optimization for CNS drug discovery. *ACS Chem. Neuro.*, 18;3(1):50-68.

Gaulton,A. et al. (2012) ChEMBL: a large- scale bioactivity database for drug discovery. *Nucleic Acid Res.*, 40(D1), D1100- D1107.

Weininger,D. et al. (1989) Smiles. 2. Algorithm for generation of unique smiles notation. *J. Chem. Inf. Comput. Sci.*, 29(2), 97-101.

RDKit: Open-source cheminformatics; <http://www.rdkit.org>

Joulin,A. and Mikolov,T. (2015) Inferring algorithmic patterns with stack-augmented recurrent nets. *Presented at: 29th Advances in Neural Information Processing Systems*. Montréal, Canada, 190–198.

Sun,J. et al. (2017) ExCAPE-DB: an integrated large scale dataset facilitating Big Data analysis in chemogenomics. *J. Cheminform.*, 9, 17.

Wu,Z. et al. (2017) MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.*, 9(2), 513-530.

Kingma,D.P. and Ba,J. (2014) Adam: a method for stochastic optimization. ArXiv doi:arXiv:1412.6980.

Tan,C. et al. (2018) A survey on deep transfer learning. *Presented At: International Conference on Artificial Neural Networks.*, Rhodes, Greece, 270-279.

Lipkus,A.H. (1999) A proof of the triangle inequality for the Tanimoto distance. *J. Math. Chem.*, 26(1–3), 263–265.

Jaques,N. et al. (2017) Tuning recurrent neural networks with reinforcement learning. *Presented At: International Conference on Learning Representations*. Toulon, France.

Tran,P.T. (2019) On the convergence proof of Amsgrad and a new version. *IEEE Access* 7, 61706–61716.

Vallianatou,T. et al. (2015) The impact of physicochemical and molecular properties in drug design: navigation in the “drug-like” chemical space. *InGeNeDis 2014*, 187-194). Springer, Cham.

Oprea,T.I. (2000) Current trends in lead discovery: are we looking for the appropriate properties?. *Mol. Divers.*, 5(4):199-208.

- Lundberg, S.M. and Lee, S.I. (2017) A unified approach to interpreting model predictions. In Proceedings of the 31st international conference on neural information processing systems 4768-4777.
- Rodríguez-Pérez, R. and Bajorath, J. (2019) Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. *J. Med. Chem.*, 63(16):8761-77.
- Sundararajan, M. et al. (2017) Axiomatic attribution for deep networks. In International Conference on Machine Learning, 3319-3328.
- Kokhlikyan, N. et al. (2020) Captum: A unified and generic model interpretability library for pytorch. arXiv preprint arXiv:2009.07896.
- Trott, O. and Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.*, 31(2), 455-461.
- Peng, Y. et al. (2018) 5-HT_{2C} receptor structures reveal the structural basis of GPCR polypharmacology. *Cell*, 172(4):719-30.
- Matzen, L. et al. (2000) 5-HT reuptake inhibitors with 5-HT_{1B/1D} antagonistic activity: A new approach toward efficient antidepressants. *J. Med. Chem.*, 43, 1149-1157.
- Halazy, S. et al. (2005) 5-HT_{1B/1D} antagonists and depression. *Expert Opin. Ther. Pat.*, 7(4), 339-352.
- Johansson, L. et al. (1997) The pharmacological characterization of a novel selective 5-hydroxytryptamine_{1A} receptor antagonist, NAD-299. *J. Pharmacol. Exp. Ther.*, 283(1), 216-25.
- Jerning, E. et al. (1998) Receptor binding characteristics of [3H]NAD-299, a new selective 5-HT_{1A} receptor antagonist. *Eur. J. Pharmacol.*, 360(2-3), 219-25.