JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

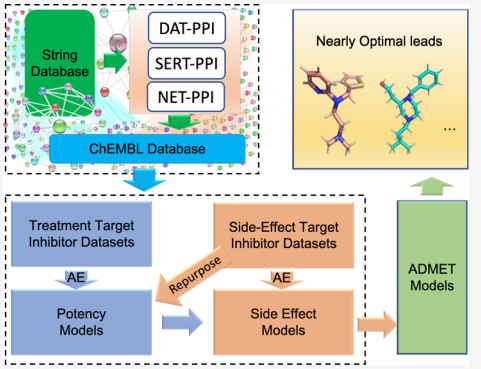
Machine Learning Analysis of Cocaine Addiction Informed by DAT, SERT, and NET-Based Interactome Networks

Hongsong Feng, Kaifu Gao, Dong Chen, Li Shen, Alfred J Robison, Edmund Ellsworth, and Guo-Wei Wei*

Cite This: *J. Chem. Theory Comput.* 2022, 18, 2703−2719 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Cocaine addiction is a psychosocial disorder induced by the chronic use of cocaine and causes a large number of deaths around the world. Despite decades of effort, no drugs have been approved by the Food and Drug Administration (FDA) for the treatment of cocaine dependence. Cocaine dependence is neurological and involves many interacting proteins in the interactome. Among them, the dopamine (DAT), serotonin (SERT), and norepinephrine (NET) transporters are three major targets. Each of these targets has a large protein−protein interaction (PPI) network, which must be considered in the anticocaine addiction drug discovery. This work presents DAT, SERT, and NET interactome network-informed machine learning/deep learning (ML/DL) studies of cocaine addiction. We collected and analyzed 61 protein targets out of 460 proteins in the DAT, SERT, and NET PPI networks that have sufficiently large existing inhibitor datasets. Utilizing autoencoder (AE) and other ML/DL algorithms, including gradient boosting decision tree (GBDT) and multitask deep neural network (MT-DNN), we built predictive models for these targets with 115 407 inhibitors to predict drug repurposing potential and possible side effects. We further screened their absorption, distribution, metabolism, and excretion, and toxicity (ADMET) properties to search for leads having potential for developing treatments for cocaine addiction. Our approach offers a new systematic protocol for artificial intelligence (AI)-based anticocaine addiction lead discovery.



1. INTRODUCTION

Cocaine abuse is a serious public health concern in the United States (US) and around the world. It is associated with a series of medical complications, including increased risk of HIV (human immunodeficiency virus), hepatitis B, and heart disease. In addition, it is also associated with rising rates of crime and violence.[1−3] Despite significant attention to discover effective pharmacotherapies for the treatment of cocaine dependence, no effective medication has been approved by the US Food and Drug Administration (FDA).

Cocaine is a tropane alkaloid and a stimulant drug with significant addictive potential. It is a nonselective inhibitor of monoamine transporters including dopamine (DAT), serotonin (SERT), and norepinephrine (NET) transporters.[4−6] By binding to these transporters, cocaine blocks reuptake of dopamine, serotonin, and norepinephrine, leading to higher synaptic and extracellular concentrations of these critical neurotransmitters. Cocaine elicits psychostimulant activities through increased activation of the monoamine receptors on post-synaptic neurons and can cause enhanced euphoric experiences.

The rewarding and addictive effects of psychostimulants are directly associated with the increased levels of dopamine in the nucleus accumbens (NAc),[7] which is a critical component of the mesolimbic and mesocortical dopamine pathways. This pathway originates from the ventral tegmental area of the midbrain and terminates with dopamine release in NAc[8] contributing to stimulant reward.[7] DAT is considered to play a primary role in the addictive effect of cocaine. Because of the critical role of DAT in cocaine addiction, many experimental medications have targeted the dopamine system.[9,10]

In addition to DAT, SERT also plays an important role in cocaine pharmacology. In vivo studies in rats demonstrated that enhanced dopamine transmission by acute cocaine intoxication in the nucleus accumbens is accompanied by elevated release of serotonin.[11] Moreover, cocaine withdrawal is associated with decreased serotonin in nucleus accumbens in microdialysis studies.[12,13] Mice with genetic deletion of DAT still show the rewarding effects of cocaine and cocaine conditioned place preference,[14] which suggests non-DAT targets contribute to psychostimulant effects. However,

Received: January 1, 2022
Published: March 16, 2022

combined dopamine and serotonin transporter knockouts eliminate cocaine place preference in mice, indicating that SERT makes a key contribution.[15,16]

Serotonin neurons originate in the raphe nuclei of the midbrain and are found in various regions of the brain, including a dense innervation of terminals to ventral tegumental area (VTA) and nucleus accumbens (NAc).[17] Cocaine-induced elevated extracellular levels of serotonin hyperactivate serotonin receptors in these and other brain regions. The actions of serotonin are mediated by at least 16 receptor subtypes that are grouped into seven families.[18] Different receptors likely serve different modulatory effects due to various neurochemical mechanisms with serotonin. Serotonin 1A receptor (5-HT_{1A}) is one of the most important receptors, existing pre- and post-synaptically in many brain areas,[19] and is involved in nearly all serotonin-mediated effects. Another family of serotonin receptors, including 5-HT_{2A} and 5-HT_{2C}, is associated with impulsivity and cue reactivity to cocaine. Either selective 5-HT_{2A} antagonist or 5-HT_{2C} lessen impulsivity and cocaine-seeking in animal models, and the synergism of pharmacotherapeutics targeting these receptors was reported to attenuate a variety of aspects of cocaine relapse.[20] Some studies show that 5-HT_3 receptor antagonists also have potential therapeutic efficacy in curbing systems of cocaine consumption[21] or are effective in abolishing restatement of cocaine self-administration.[21] Preclinical studies in rats show that serotonin-enhancing medications could help decrease self-administration of cocaine,[22,23] although selective serotonin reuptake inhibitors gave mixed results for treating cocaine addiction in clinical trials.[24]

Many noradrenergic neurons are localized in brainstem nuclei, while noradrenergic axons project virtually everywhere in the brain.[25] NET is in the plasma membrane of noradrenergic neurons and plays a primary role in the inactivation of noradrenergic signaling by reuptake of synaptically released norepinephrine (NE). Cocaine causes elevated synaptic concentration of NE by its competitive binding to NET and subsequently increases activation of post-synaptic NE receptors.[26] NE is a crucial neurochemical messenger in central noradrenergic and peripheral sympathetic pathways, and its effects are mediated by three families of adrenergic receptors: α_1, α_2, and β.[27] Stimulation of α_1-adrenergic receptors on VTA dopaminergic neurons[28] or those in the prefrontal cortex[29] promotes activity of dopaminergic neurons in the VTA. Preclinical studies found that the noradrenergic system plays a role in mediating stress-induced reinstatement of cocaine seeking. Experiments on rats found that both α_2 receptors agonists and β_1- and β_2-adrenergic receptor antagonists can reduce stress-induced cocaine-seeking behavior.[30,31] Some clinical studies suggest that adrenergic blockers are effective for cocaine dependence treatment in patients with severe cocaine withdrawal symptoms[32] and are useful in reducing cocaine self-administration.[33] Serotonin and norepinephrine reuptake inhibitor (SNRI) administration was proved to be effective in attenuating cue-induced relapse to cocaine seeking after abstinence in rodents.[34] Clinical studies on a small group of subjects indicated that SNRIs may have therapeutic potential for cocaine dependence treatment.[35] These studies indicate the promise of mediating noradrenergic signaling for the treatment for cocaine dependence.

DAT, SERT, and NET are targeted by cocaine. However, cocaine addiction involves many more proteins in their interactome with complicated molecular and functional interactions, as well as a significant number of proteins upstream and downstream. The three transporters or related receptors in their protein−protein interaction networks are of frequent focus for the therapeutic treatment of stimulant dependence with the goals of an initial period of abstinence and reduced incidence of relapse.[36] On the one hand, antagonists of these proteins could be potential therapies or assist with abstinence. The side effects and addictive liability from such antagonists in the interactome network should be a concern. This motivated us to carry out an interactome-informed systematic investigation of the potency and off-target side effects of antagonists pertaining to specific protein targets. A priority concern is potential effects on the human ether-a-go-go (*hERG*) potassium channel, and the FDA included *hERG* side effects in the most recent regulations.[37]

Protein−protein interaction (PPI) networks involve both direct (physical/chemical) and indirect (functional) interactions and associations,[38] where a connection represents two proteins jointly contributing to a specific biological function even without direct physical/chemical interaction. Interactome provides a large number of proteins associated with a specific disease, facilitates the understanding of pathogenic mechanisms underlying the cause and progression of diseases, and promotes development of novel disease treatments. The analysis of PPIs on the proteome scale is a promising approach for the design of novel treatments for cocaine addiction as well as potential side effects. Many interacting proteins in the interactome are collected in the String database,[38] which makes it possible for interactome-informed systematic analysis of compounds targeting pathways related to cocaine addiction.

In our previous work,[39] a machine learning/deep learning (ML/DL) approach was built around a PPI network extracted from the String dataset and centered on DAT, providing cross-target binding prediction and searching for potential inhibitors to treat cocaine addiction. ML/DL models are especially useful to predict binding affinities to targets in large-scale PPI network analysis, in contrast to traditional in vivo or in vitro experiments, which are time-consuming, expensive, and ethically constrained due to animal testing. With autoencoder (AE)-generated representative features for inhibitor molecules, the gradient boosting decision tree method can be used to build ML models for the protein targets with sufficient inhibitor datasets in the PPI network. However, it is very common that machine learning models may suffer from poor predictive performance when the training dataset is very small. Fortunately, multitask algorithms offer a promising option to resolve this issue by taking advantage of other similar tasks with a large dataset.[39,40] The philosophy of multitask learning relies on the explorations of sharing relatedness and transferring knowledge between tasks, resulting in improved prediction of models with a small dataset. In this study, multitask learning becomes a very useful tool to enhance the model's predictive power since many protein targets may belong to the same family, and the underlying molecular mechanism of protein−ligand binding shares commonalities, which can be learned by joint training.

In the present work, we extend our earlier effort on DAT[39] to the SERT and NET networks to perform systematic analysis on potential therapeutic compounds, side effects, and potential drug repurposing. Models constructed by gradient boosting decision tree (GBDT) and multitask deep neural network (MT-DNN) are adopted to accomplish these predictions and analyses. More extensive and comprehensive investigations are

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

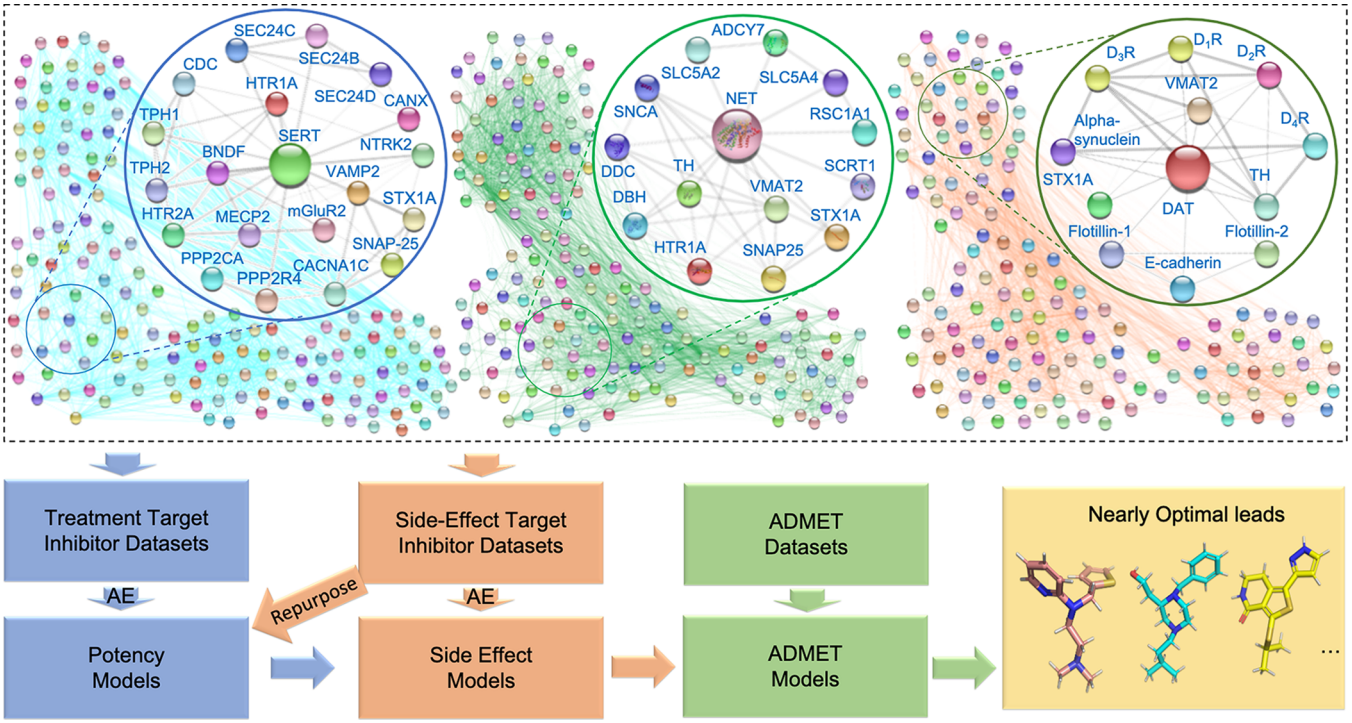


Figure 1. A core and global network centered on DAT, SERT, and NET, as well as the proteome-informed ML workflow of anticocaine addiction drug discovery. An autoencoder (AE)-based machine-learning (ML) approach is used to encode inhibitors or antagonists of proteins in the networks, and ML models are built to predict binding affinities to each protein. Screening of DAT, SERT, or NET inhibitor datasets and repurposing of inhibitors or antagonists from other protein targets are two key processes for drug discovery. ADMET screening is performed following the screening or repurposing process, resulting in potentially nearly optimal leads. Abbreviations for the core SERT network: SERT (serotonin transporter), HTR1A (5-hydroxytryptamine receptor 2A), HTR2A (5-hydroxytryptamine receptor 2A), BDNF (brain-derived neurotrophic factor), CANX (calnexin), PPP2R4 (serine/threonine-protein phosphatase 2A activator), PPP2CA (serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform), SEC24B (protein transport protein Sec24B), SEC24C (protein transport protein Sec24C), STX1A(syntaxin-1A), TPH1 (tryptophan 5-hydroxylase 1), TPH2 (tryptophan 5-hydroxylase 2), DDC (dopa decarboxylase), mGluR2 (metabotropic glutamate receptor 2) and SNAP-25 (synaptosomal-associated protein, 25 kDa). Abbreviations for the core NET network: NET (sodium-dependent noradrenaline transporter), VMAT2 (synaptic vesicular amine transporter), ADCY7 (adenylate cyclase type 7), RSC1A1 (regulatory solute carrier protein family 1 member 1), SCRT1 (transcriptional repressor scratch 1), SLC5A2 (sodium/glucose cotransporter 2), SLC5A4 (solute carrier family 5 member 4), SNCA (alpha-synuclein), TH (tyrosine 3-monooxygenase), DBH (dopamine beta-hydroxylase). Abbreviations for the core DAT network: DAT (dopamine transporter), D_1R (dopamine receptor 1), D_2R (dopamine receptor 2), and D_3R (dopamine receptor 3).

implemented by combining the SERT, NET, and previously studied DAT networks. Special attention to the repurposing potentials of compounds targeting DAT, SERT and NET, as well as off-target side effects of inhibitors of these three transporters in the combined network is considered. Potency, side effects, pharmacokinetic properties, and synthetic accessibility through ML/DL predictions that form a series of filters for screening are evaluated for potential lead compounds.

2. METHODS

2.1. The Cocaine Addiction PPI Networks. In addition to DAT, both SERT and NET play modulatory roles in behavioral responses to cocaine, and their functions may be pivotal to addiction. They have been frequently investigated to understand the neurobiology of cocaine addiction, providing insights for therapeutical intervention.[16,41] PPI networks help to reveal molecular interactions and cellular mechanisms. We extracted the PPI networks of SERT and NET by respectively inputting the protein names "serotonin transporter" and "norepinephrine transporter" in the String Web site (https://string-db.org/).

Two global networks centered on SERT and NET were extracted from the String database[38] for proteins that have direct or indirect interactions with SERT or NET. The global network for SERT is formed of 151 nodes and 1720 edges, among which a core network with 20 protein nodes that have direct interaction with SERT protein and total 66 edges are considered. The global network for NET is composed of 158 nodes and 1791 edges and contains a core network of 14 nodes and 37 edges.

There are 20 proteins in the core of the SERT network comprised of SERT itself and 19 other proteins that have direct interactions with SERT. SERT transports serotonin molecules from synaptic cleft back into the pre-synaptic neuron for repackaging and rerelease. It plays a critical role in mediating the availability of serotonin for other receptors in the serotonergic system and terminates the effects of serotonin by removing it from the synaptic cleft. It is one of the three direct targets of cocaine, and its inhibition by cocaine may contribute to cocaine dependence. 5-HT1A receptor (5-hydroxytryptamine receptor 1A) and 5-HT2A receptor (5-hydroxytryptamine receptor 2A) are inhibitory G-protein coupled receptors for serotonin. Through binding to serotonin, they mediate hyperpolarization and reduction of the firing rate

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

of the post-synaptic neuron. Antagonists targeting 5-HT1A receptor and 5-HT2A receptor were found to diminish the motivational effects of cocaine.[42,43] Brain-derived neurotrophic factor (BDNF) promotes the survival and differentiation of selected neuronal populations of the central and peripheral nervous systems. CANX (calnexin) assists in protein assembly and quality control of the endoplasmic reticulum. PPP2R4 (serine/threonine-protein phosphatase 2A activator) accelerates the folding of proteins and acts as a regulatory subunit for serine/threonine-protein phosphatase 2A, modulating its activity or substrate specificity. PPP2CA (serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform) is an enzyme providing negative control of cell growth and division. Proteins SEC24B (protein transport protein Sec24B), SEC24C (protein transport protein Sec24C), and SEC24D (protein transport protein Sec24D) are involved in vesicle trafficking and promoting the formation of transport vesicles from the endoplasmic reticulum. STX1A (syntaxin-1A) is critical for hormone and neurotransmitter exocytosis and is implicated in the docking of synaptic vesicles with the pre-synaptic plasma membrane. TPH1 (tryptophan 5-hydroxylase 1) and TPH2 (tryptophan 5-hydroxylase 2) belong to the biopterin-dependent aromatic amino acid hydroxylase family and are rate-limiting steps in serotonin synthesis. DDC (dopa decarboxylase) catalyzes the decarboxylation of L-3,4-dihydroxyphenylalanine (DOPA) to dopamine, L-5-hydroxytryptophan to serotonin, and L-tryptophan to tryptamine, hence modulating the amount of serotonin and catecholamine in the human body. MECP2 is widely found in neurons in the brain and promotes the maturation of the central nervous system. Metabotropic glutamate receptor 2 (mGluR2) inhibits the emptying of vesicular contents at the pre-synaptic terminal of glutamatergic neurons. SNAP-25 (synaptosomal-associated protein, 25 kDa) is critical for synaptic membrane fusion of vesicles containing neurotransmitters, as is VAMP2, and both are critical for neurotransmitter release. The functions of serotonin receptors and STX1A can be seen to have direct interactions with SERT, since they are directly related the transportation of serotonin across SERT. Other proteins are critical in modulating the protein function of SERT to transport serotonin or related to the serotonin transportation. The edges in the SERT core network in Figure 1 represent the direct interactions between SERT and other 19 proteins.

The core of the NET network contains 14 proteins. NET is responsible for the reuptake of extracellular norepinephrine, extracellular dopamine, and regulates the concentration of these two neurotransmitters in the synaptic cleft. VMAT2 (synaptic vesicular amine transporter) transports monoamines, especially neurotransmitters such as dopamine, norepinephrine, serotonin, and histamine, from the cytosol into synaptic vesicles. ADCY7 (adenylate cyclase type 7) is an enzyme that can catalyze the formation of cyclic AMP from ATP. RSC1A1 (regulatory solute carrier protein family 1 member 1) mediates transcriptional and post-transcriptional regulation of gene SLC5A1 and is also involved in transcriptional regulation of SLC22A2. SCRT1 (transcriptional repressor scratch 1) modulates the function of multiple transcription factors to regulate neuronal differentiation. SLC5A2 (sodium/glucose cotransporter 2) is the sodium-dependent glucose transporter and is responsible for the reabsorption of 80%−90% of the glucose filtered by the kidney glomerulus. SLC5A4 (solute carrier family 5 member 4) may function as a glucose sensor. SNCA (alpha-synuclein) is a neuronal protein that regulates synaptic vesicle trafficking and subsequent neurotransmitter release including dopamine release and transport. TH (tyrosine 3-monooxygenase) is the enzyme responsible for catalyzing the conversion of the amino acid L-tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA), a precursor for dopamine, and its activity may be increased by cocaine exposure. DBH (dopamine beta-hydroxylase) catalyzes the conversion of dopamine to norepinephrine. SNAP-25, DDC, HTR1A, and STX1A are also in the core of the NET network and are described above.

DAT, SERT, and NET constitute important components in cocaine addiction networks as they are directly inhibited by cocaine. Networks centered at the three proteins have intricate intranetwork and internetwork interactions. As shown in Figure 1, the core networks formed around the three proteins serve as the center of their global networks, respectively. The SERT global network consists of four clusters of proteins, each cluster potentially serving different functions in the serotonin system and modulatory effects of serotonin in cocaine responses. SERT and most proteins in its core network are in one of the four clusters. Proteins in the four clusters have close interactions with SERT or proteins in its core network. Both global networks for NET and DAT consist of three clusters of proteins, with most of the proteins in the core network sitting in one of the clusters. The three global networks are not independent of each other. There are many common proteins between networks. For instance, syntaxin-1A is found in all three core networks. As discussed, SERT and NET core networks share up to four proteins. Common proteins can also be found by other pairwise network comparisons. Because of the critical roles of DAT, SERT, and NET in the global protein network, and their essential modulatory effects in cocaine addiction, medications targeting them could achieve profound pharmacological effects in reducing cocaine addiction.

2.2. Datasets. We collected inhibitor datasets for the proteins in SERT and NET networks from ChEMBL.[44] We then built models for the sets with sufficient data (i.e., over 250 data points), and 37 distinct target proteins from SERT and NET networks. The details of these inhibitor datasets are listed in the Supporting Information. Because of the role of protein *hERG* in causing serious side effects in drug design, we also built a model for the *hERG* blocker dataset.

These models allow us to perform cross-target binding affinity (BA) prediction for other datasets. In addition to the aforementioned 37 datasets, we also included 30 protein datasets and corresponding reliable models from the previous work[39] for the DAT network, as they may be involved in cocaine addiction. The DAT network shares some common proteins with the current SERT-NET network. In total, we built 61 different datasets with models. A Venn diagram summarizing the 60 proteins from three networks can be found in the Supporting Information. In addition, we performed a brief categorization analysis for the 61 proteins and include the description in Table S4 in the Supporting Information.

2.3. Molecular Representations. Molecular fingerprints are used to represent molecules, usually in the form of vectors with each vector element indicating the existence, degree, or frequency of each structural characteristic or property. These fingerprints have a variety of applications in ML/DL analysis, virtual screening, similarity-based compound searches, target molecule ranking, drug ADMET prediction, and other drug discovery processes.[45]

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

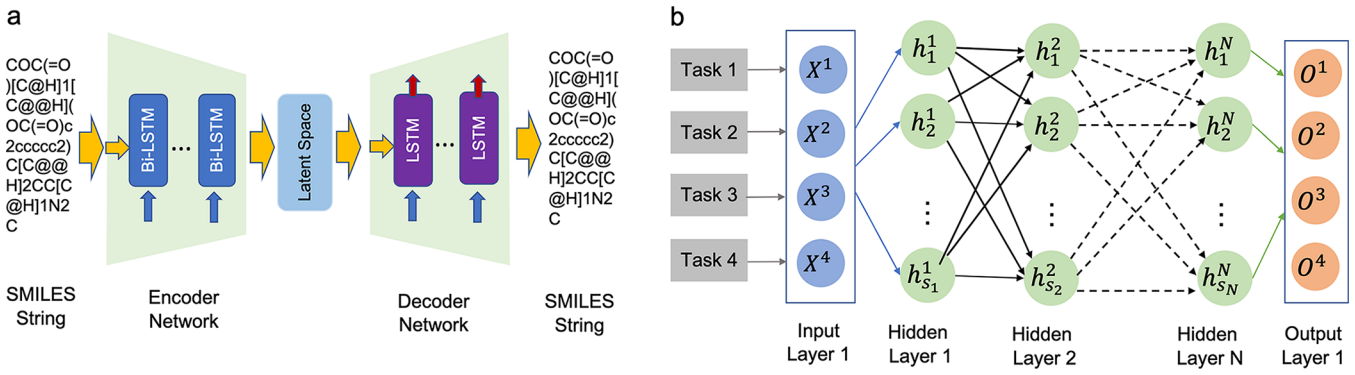


Figure 2. Illustration of a seq2seq AE model and a MT-DNN model. (a) seq2seq AE model used for LV-FP generation. Bi-LSTM and LSTM are used in encoder and decoder networks, respectively. (b) A simple illustration of MT-DNN for four tasks. The input layer consists of the feature vectors X^t for each task. The hidden layers constitute the shared neuron networks for the four tasks, and the output layer represents the predictions for four tasks.

2D fingerprints, mathematical representations, and latent-vector fingerprints (LV-FPs) constitute three types of molecular fingerprints frequently used in ML/DL models. Molecular feature extraction by means of state-of-the-art 2D fingerprints can be classified into four categories, including substructure key-based fingerprints, topological or path-based fingerprints, circular fingerprints, and pharmacophore fingerprints.[45] MACCS, FP2, ECFP and 2D-pharmacophores are all popular 2D fingerprints and can be generated by the open-source cheminformatics software RDKit. In addition, models based on deep neural network (DNN) draw interest in extracting mathematical features. Many differential geometry, algebraic topology, and spectral graph-based approaches have been investigated in this regard.[46] LV-FPs are the molecular representations in the neural layers of DNN architectures and gained wide popularity in a drug discovery[47,48] and molecular analysis.[49] Among those DNN architecture models, the seq2seq model draws particular attention. In this work, we adopted the LV-FP of molecules generated by our in-house seq2seq models.

2.4. Deep seq2seq Autoencoder Model. As shown in Figure 2a, the seq2seq model is a DL autoencoder architecture originated from natural language processing. It has achieved breakthrough success in English−French translation and conversational modeling. The seq2seq autoencoder model consists of two neural networks, an encoder, and a decoder. The encoder translates an input sequence of variable length with discrete values to a fixed-sized continuous representation (latent representation) using a gated recurrent unit (GRU)[50] or a long short-term memory (LSTM) network,[51] and the decoder maps the latent representation to a target sequence with another GRU or LSTM network. These intermediate continuous latent representations are called latent vectors and are often used to characterize the source or target sequence.

In our study, input and output sequences are both SMILES strings−a one-dimensional (1D) "language" of chemical structures. Our autoencoder model was trained to have a high reconstruction ratio between input and output SMILES strings while latent vectors carry faithful information on the chemical structures. The latent vectors can be used to represent compounds. The seq2seq model and LV-FPs were realized by our in-house source code. We applied bidirectional LSTM as the encoder and LSTM as the decoder. The generated LV-FPs have a dimension of 512. Our seq2seq autoencoder network is illustrated in Figure 2a. More details about our seq2seq autoencoder model can be found in the Supporting Information.

2.5. Gradient Boosting Decision Tree. The gradient boosting decision tree (GBDT) is a popular ensemble method and is robust against overfitting, insensitiveness to hyperparameters, and ease of implementation. Particularly, in terms of efficiency, it is faster than DNN. When training small datasets, it can have better performance than DNN and a variety of other deep learning algorithms. It has been adopted in a wide range of quantitative structure−activity relationship (QSAR) prediction problems,[40,52] and achieved some state-of-the-art results. Other ensemble methods including random forests (RF) and support vector machines (SVM) also share a variety of merits as GBDT, but their performances are generally not as good as GBDT in QSAR prediction problems.[40,45,53] In the Supporting Information, we gave a performance comparison between GBDT models and RF models with our LV-FPs on the prediction for DAT, SERT, and NET datasets. It was observed that GBDT models outperformed RF models. The superiority of GBDT is largely due to the fact that the random forest model builds trees in parallel, but the GBDT model builds trees in a sequential manner, such that existing trees in GBDT pass information on to subsequently produced trees. In this study, a GBDT regressor in scikit-learn (version 0.20.1) was adopted to build models for the 61 total protein targets. The hyperparameters were tuned by grid search method, and those hyperparameters, namely, "n_estimators, max_depth, min_samples_split, subsample, max_features" are the tuning ones. The hyperparameters leading to the highest average Pearson correlation coefficient via 10-fold cross-validation on each dataset are chosen as our optimal parameters and are summarized in the Supporting Information.

2.6. Multitask Deep Neural Network. Through exploiting the relatedness among different tasks, multitask learning transfers learned information and knowledge between similar tasks in the process of simultaneously training several tasks. It is particularly beneficial in enhancing the predictive performance of models with small training datasets by joint learning with those equipped with larger datasets. In this way, poor predictive performance due to the difficulty of extracting enough representative information from a small dataset is overcome by the shared information/distribution from coupled similar tasks with large datasets. Multitask deep neural network (MT-DNN) has gained wide popularity in drug discovery[45]

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

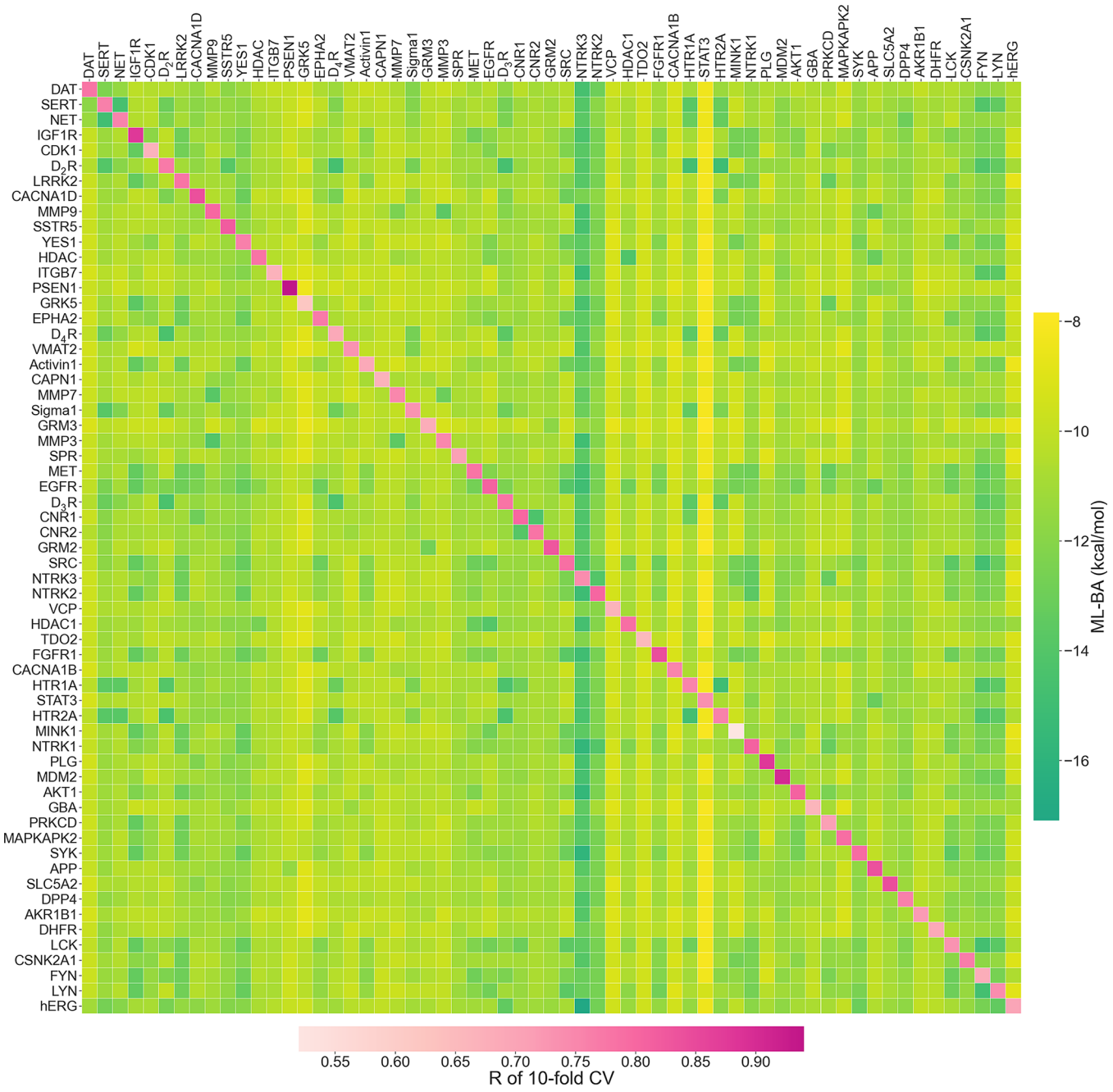


Figure 3. Heatmap of cross-target BA prediction indicating the inhibitor specificity of each dataset. In each row, the diagonal element shows the Pearson correlation coefficients of 10-fold cross-validation (R of 10-fold CV) on the machine-learning predicted BAs (ML-BAs) of each dataset. Off-diagonal elements represent the highest ML-BAs of the inhibitors in each dataset to other targets.

and molecular bioactivity predictions.[40,52] MT-DNN builds a wide neural network with multiple hidden layers and hundreds, or thousands of neurons exist in each layer. In such a large network, latent representation features can be formed and complex nonlinear relationships can be formulated between data and target labels. MT-DNN performs simultaneous training on the shared hidden layers for several tasks and produces predicted labels for each task. Figure 2 is a simple illustration of a MT-DNN model for four tasks. There are N hidden layers, and each layer consists of s_k neurons for k = 1, 2, ..., N. The input layer represents the feature vectors of each task, and the output stands for the predictions for the four tasks.

To further illustrate, if we have T different tasks with the training information given as $(X_i^t, y_i^t)_{i=1}^{N_t}$ where X_i^t is the autoencoder feature representation for the ith sample molecule in task t with the corresponding label y_i^t, and N_t denotes the number of molecules for task t. Indices t = 1, 2, ..., T indicate the tasks to work on.

The goal in multitask learning is to minimize the loss function of each task simultaneously:

$$\arg\min L(\boldsymbol{y}^t, f^t(\boldsymbol{X}^t; \{\boldsymbol{W}^t, \boldsymbol{b}^t\})) \quad (1)$$

where $\boldsymbol{y}^t = \{y_i^t\}$ is the label vector for task t, f^t is the predictor for task t, $\boldsymbol{X}^t = \{X_i^t\}$ is the feature collection for dataset of task t, and $\boldsymbol{W}^t$,$\boldsymbol{b}^t$ are weight vector and bias term. It is noteworthy that

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

the feature length of data in all X^t for $t = 1, 2, ..., T$ must be the same in order to be fed into the same neural network. In such a regression problem, the mean square loss function is frequently used as a loss function, with which we have the following loss function formulation for task t:

$$L(\boldsymbol{y}^t, f^t(\boldsymbol{X}^t; \{\boldsymbol{W}^t, \boldsymbol{b}^t\})) = \frac{1}{2}\sum_{i=1}^{N_t}(y_i^t - f^t(X_i^t; \{\boldsymbol{W}^t, \boldsymbol{b}^t\}))^2$$

MT-DNN on the shared network has a tendency to significantly attenuate the risk of overfitting since simultaneous learning promotes representations that can capture all the tasks and reduce the chances of fitting noise for a specific task. A popular improved loss function is introduced by adding regularization term with weight vector, i.e., the loss function for task t is reformulated as

$$L(\boldsymbol{y}^t, f^t(\boldsymbol{X}^t; \{\boldsymbol{W}^t, \boldsymbol{b}^t\})) = \frac{1}{2}\sum_{i=1}^{N_t}(y_i^t - f^t(X_i^t; \{\boldsymbol{W}^t, \boldsymbol{b}^t\}))^2 + \beta\|\boldsymbol{W}^t\|_2^2$$

where $\|\cdot\|_2$ represents the L_2 norm, and β indicates a penalty weight.

In our study, we are interested in the binding affinity predictions of inhibitor molecules to specific protein targets. It is reasonable to anticipate that a collection of proteins could share similar protein structures, especially similar binding sites, around which analogous molecular mechanisms occur. Joint training for similar protein targets promotes sharing of a more-complete feature statistical distribution between those tasks, which is beneficial to those with only a small training dataset available.

Network hyperparameters. The MT-DNN in our implementation consists of four hidden layers with the neuron number in each layer chosen as 1024, 1536, 1536, and 1024. The optimizer is set to be stochastic gradient descent (SGD) with momentum of 0.5. 1000 epoch runs were performed for the network. The mini-batch is 4 and the learning rate is 0.001 for the first 900 epochs and 0.0005 for the last 100 epochs. No dropout or L_2 decay is used in the implementation. The multitask training was performed with Pytorch (1.10).

2.7. Binding Affinity Predictions. The features encoded with LV-FPs provides us with a valuable tool to describe the inhibitor molecules. With LV-FPs, we make use of GBDT and MT-DNN to build machine learning models. A GBDT regressor in scikit-learn was adopted for the GBDT models, and Pytorch was used to perform multitask training. Most of the models were built with GBDT, and we mainly used MT-DNN to improve the prediction for protein targets with small datasets. Detailed discussion will be presented in the following section. We then used these models to make BA predictions. In our study of cocaine addiction, we built models for the 58 datasets with GBDT, and the other three came from MT-DNN. We primarily seek to improve the prediction performance for models of small datasets using multitask transfer learning. Consequently, we are not concerned if the performance for models of large datasets is not improved. We can continue to use the obtained GBDT model for the protein target of a large dataset if its multitask prediction becomes worse. The validation for effectiveness and robustness of our GBDT models by comparisons with the literature was reported in our previous work,[39] and some of the comparison results are included in the Supporting Information.

3. RESULTS

3.1. Reliability Tests. In this section, we adopt ML/DL models to systematically predict inhibitor BAs to analyze side effects and repurposing potential. These ML models show promising predictions with Pearson correlation coefficient (R) in 10-fold cross-validation (CV) tests for each inhibitor dataset, as indicated in the heatmap of Figure 3. Eleven (11) of the 38 new models for inhibitor datasets have $R > 0.8$, and 30 of the 38 models have $R > 0.7$. Only 1 of the 38 have $R < 0.6$. The details of the 38 new models in addition to previously built DAT models can be found in the Supporting Information. We started with GBDT model constructions for all 61 protein targets. However, some models have low prediction performance mainly because of their small inhibitor datasets. Thus, we employed MT-DNN models to enhance the predictions for such protein targets. FYN, LYN, and NTRK3 datasets have very small sizes of 470, 468, and 355, respectively, and the R values for their GBDT models are 0.56, 0.54, and 0.68, respectively. As mentioned, multitask transfer learning could assist with such cases by joint training with large datasets, assuming that such large datasets are related to the small one and can transfer useful statistical distributions. It is critical to determine when to use multitask transfer learning. We choose proteins with large datasets, having high protein similarities, and high associated dataset similarities. It is also necessary to consider the sequence similarity, because high sequence similarity indicates similar protein structures and potentially similar binding sites. Figures S2 and S3 in the Supporting Information display the similarities of pairwise inhibitor datasets and protein sequences for the 61 targets. Once we determined a few candidates with high sequence similarity scores, we select a large dataset with the highest similarity to the small target dataset. In this way, modeling LCK dataset was found to be one of the candidate tasks with a modeling LYN dataset, because of their high similarities. Specifically, LYN and LCK have a sequence similarity of 0.848 and LYN dataset has a similarity of 0.867 with LCK dataset. The modeling of the FYN dataset also has the modeling of the LCK dataset as its most related task, because FYN has a sequence similarity of 0.876 with LCK and dataset similarity of 0.784 with LCK data. Importantly, LCK has a relatively large inhibitor dataset of 1855 compounds, which allows its dataset to be used for multitask training with the FYN dataset or the LYN dataset. In fact, FYN and LYN share sequence and dataset similarities with each other, which makes sense as these three proteins are in the tyrosine−protein kinase family. Thus, it is reasonable to put the three tasks in our MT-DNN. The R values show that the multitask training indeed boosted the performances of predictive models for FYN and LYN. The R values for FYN and LYN models are increased from 0.56 to 0.68, and from 0.54 to 0.74 respectively. The comparison results for LCK, FYN, and LYN modeling are detailed in the multitask 1 column in Table 1. In addition, the modeling of the NTRK1 dataset is found to be most suitable for multitask training with the modeling of the NTRK3 dataset. NTRK3 has a sequence similarity of 0.754 and a dataset similarity of 0.823 with NTRK1. The NTRK1 dataset has 2783 molecules, which makes it a good candidate for multitask training with NTRK3 dataset. The high sequence similarity between NTRK1 and NTRK3 is attributable to the fact that they are both in the

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

Table 1. Improvement on Model Performance via MT-DNN[a]

	Multitask 1			Multitask 2	
	LCK	LYN	FYN	**NTRK1**	NTRK3
dataset size	1855	468	470	2783	355
R-GBDT	0.73	0.55	0.56	0.81	0.68
R-MT-DNN	0.72	0.74	0.68	0.71	0.74
sequence similarity	1.000	0.848	0.876	1.000	0.754
dataset similarity	1.000	0.867	0.784	1.000	0.823
improvement	−1.38%	+34.6%	+20.4%	−12.4%	+8.8
	Multitask 3		Multitask 4		
	LCK	MINK1	**SRC**	MINK1	
dataset size	1855	364	3054	364	
R-GBDT	0.73	0.52	0.86	0.52	
R-MT-DNN	0.32	0.36	0.39	0.41	
sequence similarity	1.000	0.677	1.000	0.695	
dataset similarity	1.000	0.674	1.000	0.568	
improvement	−56.2%	−30.8%	−54.7%	−21.2%	

[a]R-GBDT and R-MT-DNN stand for the Pearson correlation coefficient of GBDT and MT-DNN models, respectively. LCK, NTRK1, and SRC are the three large datasets used for MT-DNN. The + or − signs indicate how much improvement on the *R* value the MT-DNN has, compared to GBDT models. Our sequence and dataset similarity analysis sheds light on the performance of MT-DNN models.

factor receptor family. Through this multitasking approach, the *R* value for the NTRK3 model is increased from 0.69 to 0.74. All these results show the effectiveness of using MT-DNN to boost prediction performance. The improvement in prediction performance via MT-DNN is listed in Table 1. Importantly, we also encountered a case when multitask fails due to low task similarity. The GBDT model for MINK1 has an *R* value of 0.52. Modeling the LCK dataset is found to be a similar task to modeling the MINK1 dataset as MINK1 has a sequence similarity of 0.677 and dataset similarity of 0.674 with LCK. The multitask training did not help improve the prediction performance of MINK1 model due to relatively low similarities. The modeling of the SRC dataset was also explored for multitask training with MINK1 task considering MINK1's sequence similarity 0.695 and dataset similarity of 0.568 with SRC. Both MT-DNN efforts were not successful probably because MINK1 is a misshapen-like kinase, but LCK and SRC are in the family of tyrosine−protein kinases. This explains why LCK and SRC have relatively low sequence similarities to MINK1. In addition, relatively low dataset similarities also makes it hard to transfer useful information to the modeling of the MINK1 dataset. In this case, we continued to use the GBDT model for MINK1. The results of four multitask learning are shown in Table 1.

3.2. Cross-Target Binding Affinity Predictions. The highest BA of cross-target prediction represents the side-effect strength of inhibitors to other targets. The 3660 cross-target predictions in Figure 3 exhibit 3336 potential side effects determined by BAs higher than −9.54 kcal/mol (K_i = 0.1 μM). On the other hand, the remaining 324 predictions with all BA values greater than −9.54 kcal/mol suggest weak side effects. Figure 3 depicts the results of the cross-target BA prediction of the 61 inhibitor datasets. Along the diagonal line are the *R* values of 10-fold CV tests for the corresponding protein datasets, while the off-diagonal elements indicate the predicted minimal BA of the datasets to a specific target protein. For example, the *j*th element in the *i*th row shows the predicted

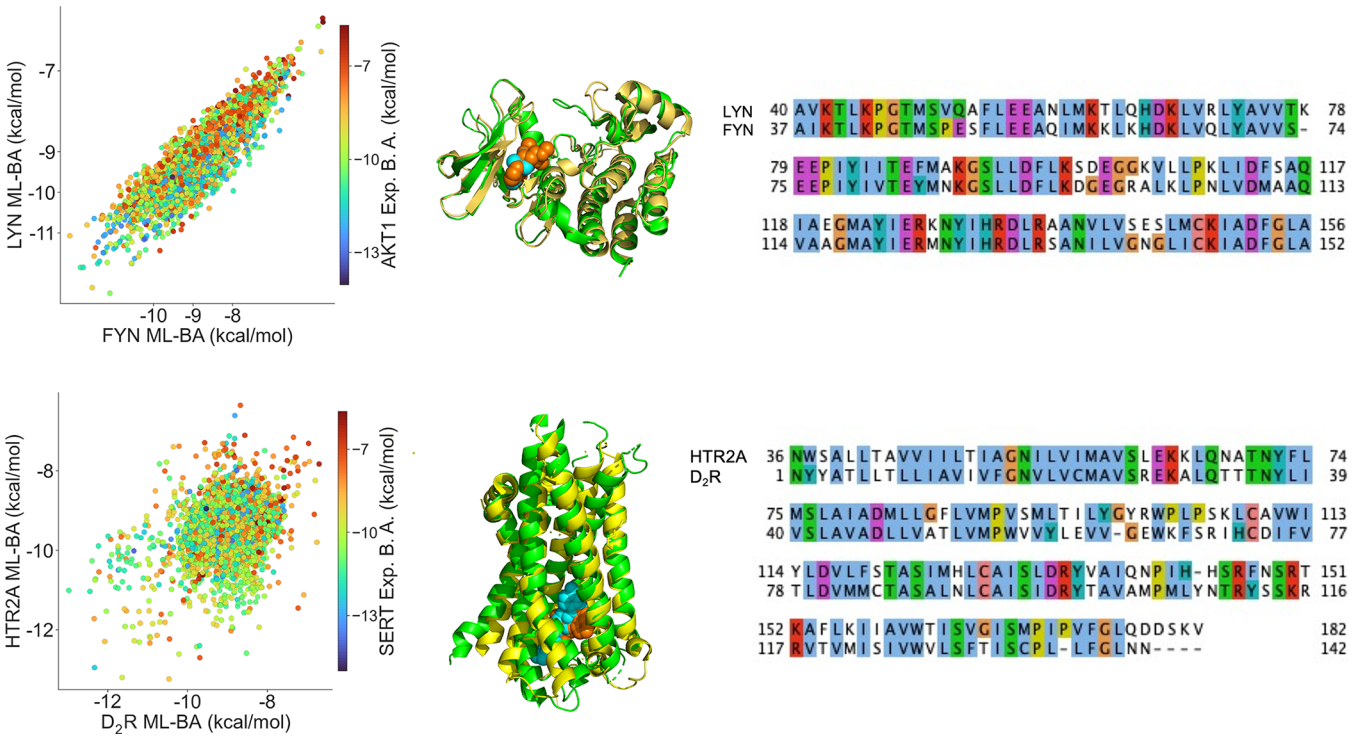


Figure 4. Examples of cross-target predicted BA correlations detecting binding site similarities of different proteins. The first column exhibits the predicted BAs of the inhibitor dataset to different targets, the second column corresponds to the 3D structure alignment of two proteins, and the third column displays the sequence alignment of the binding site. The PDB IDs are 6CM4, 6WGT, 2DQ7, and 2ZV9 for D_2R, HTR2A, FYN, and LYN, respectively.

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

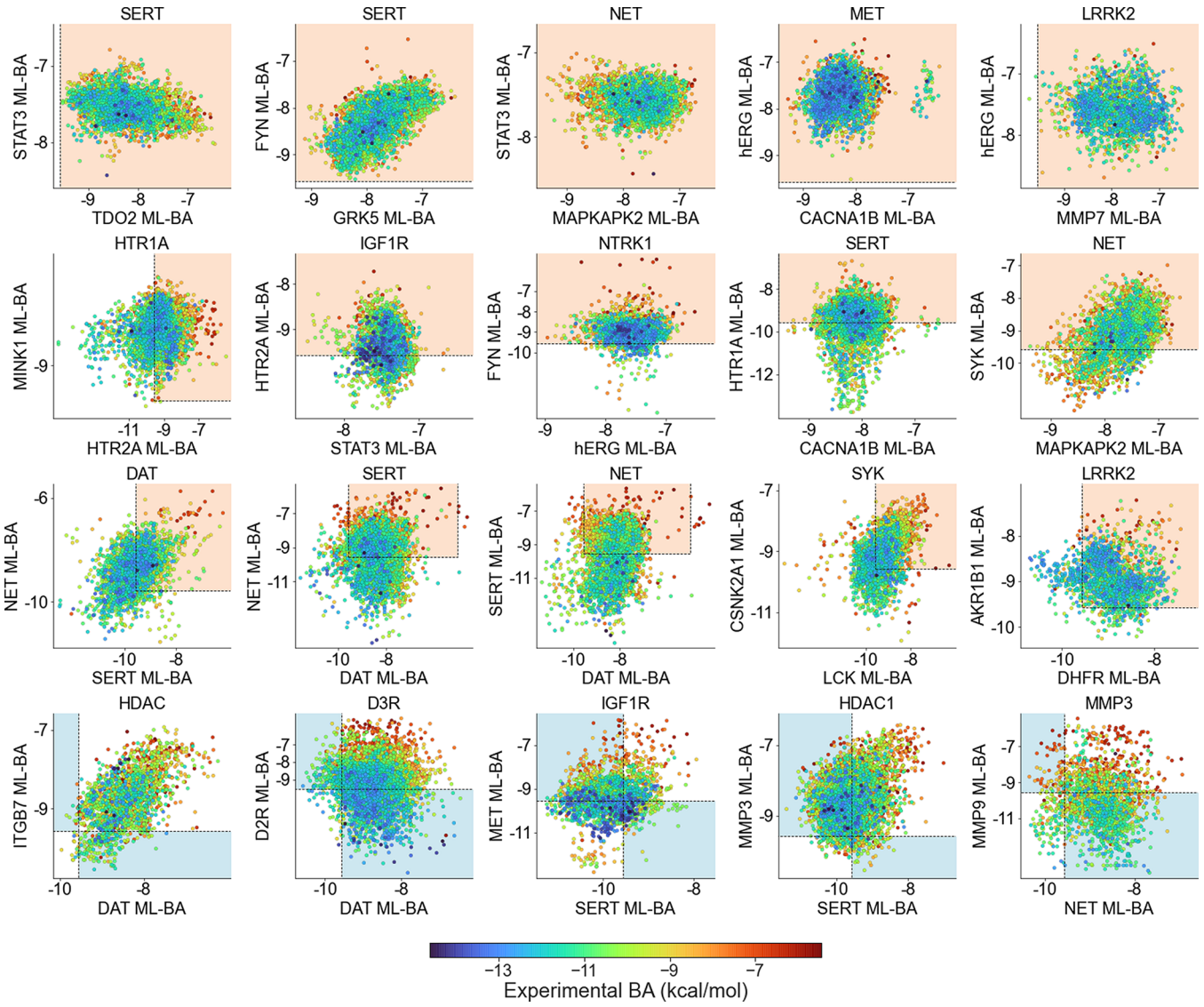


Figure 5. Examples of inhibitors' possible side effects or repurposing potentials. The first three rows list some inhibitor datasets that have side effects to 0, 1, or 2 of two off-target proteins. The orange rectangular frames outline the ranges where no side effects of inhibitors are caused to either off-target protein. The colors of points represent the experimental BAs to these targets. The last row displays some inhibitor datasets that have repurposing potential to other proteins. The two blue rectangles highlight the domains where inhibitors can have repurposing potential to one protein but have no side effect to the other one.

highest BA of the *i*th inhibitor dataset listed to the left of the heatmap by the *j*th target model listed on top of the heatmap.

3.2.1. Cross-Target BA Correlation Revealing Binding Site Similarities. Binding site similarities can yield cross-target BA correlation. On the other hand, high BA correlation can help identify binding site similarities. According to our cross-target prediction, there are some examples of high BA correlation to similar binding sites. For instance, FYN (tyrosine−protein kinase FYN) and LYN (tyrosine-protein kinase LYN) are both kinases. The AKT1 inhibitor dataset has BAs correlation R of 0.54 to these two proteins. Their 3D protein structure alignment displayed on the right shows their highly similar structure. Their sequence alignment also has a sequence identity as high as 75% around the binding site.

We also found that the proteins from different families can also have similar binding sites, leading to cross-target BA correlations, and an example is shown in the second row of Figure 4. The predicted BAs of the SERT inhibitor dataset to D_2R and HTR2A have a Pearson correlation coefficient (R) of 0.31. D_2R and HTR2A are receptors for dopamine and serotonin, respectively, but they are not in the same family. The highly similar 3D structure is shown on the right, and the sequence alignments of the binding site have a sequence identity of 42.11%. Some additional examples of similar binding sites can be found in the Supporting Information.

3.2.2. Predictions of Side Effects and Repurposing Potentials. It is desirable for a drug candidate to be highly specific, i.e., having a high BA to its target and very low BAs to all other human proteins, reducing likelihood of side effects. In addition, if a drug candidate interacts weakly with its designated target but is potent to another unintended protein, it has repurposing potential. Our ML model is a useful tool to study possible side effects and repurposing potentials by systematically performing cross-target predictions. We adopted the 61 available models to predict the BAs of compounds in other datasets to specific protein targets. The 61 inhibitor

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

datasets and related models allowed us to evaluate possible side effects of drug candidates as well as their repurposing potentials for other protein targets.

Figure 5 plots some examples of inhibitors' side effects and repurposing potentials by prediction. As mentioned previously, the selected proteins come from DAT, SERT, and NET networks. Proteins in the three networks have internetwork interaction. When considering cross-target predictions by the 61 models in our study, we do not limit our discussion within each network. All possible cross-target predictions among 61 models on the 61 datasets are performed. Each chart in Figure 5 involves one inhibitor dataset and two target proteins. The x and y values represent the predicted BAs of a designated inhibitor dataset to two other proteins, the dot color represents the experimental binding affinity of molecules in the designated dataset. The first three rows depict cases of predicted side effects, and the last row depicts cases of repurposing potential. The orange frames in the first three rows highlight the domains where inhibitors for a designated protein would not have side effects on the other two targets, i.e., BAs > −9.54 kcal/mol (K_i = 0.1 μM). The two blue frames in each chart of the last row indicate the ranges where inhibitors have repurposing potential to one target, i.e., predicted < −9.54 kcal/mol (K_i = 0.1 μM)) and no side effects are caused on the other ones, i.e., predicted BA > −9.54 kcal/mol.

The first row of Figure 5 shows cases of inhibitor dataset causing no side effects to two other targets. All the active inhibitors of desired targets are predicted to have low BAs to other targets. For example, the first and second charts indicate that all active inhibitors of SERT are predicted to cause no side effect to TDO2, STAT3, GRK5, and FYN proteins. It is the same case in the third and fourth charts that NET inhibitors have no predicted side effects on MAPKAPK2, STAT3, CACNA1B, and *hERG* proteins. The second and third rows show inhibitors of the designated protein having side effects on either or both of two other targets. The first panel in the second row shows that the HTR1A dataset has more than half of its inhibitors predicted to bind with affinity < −9.54 kcal/mol to HTR2A. This can be due to a high structural similarity, since they are both serotonin receptors. The first three charts in the third row show that a large number of inhibitors of DAT, SERT, and NET can cause side effects to two other proteins.

Some examples of inhibitors with repurposing potential are provided in the last row of Figure 5. The third panel in the fourth row shows that many inhibitors of IGFR1 may have repurposing potentials for SERT and have no side effect on MET. In the fifth panel of the fourth row, we can observe that few inhibitors of MMP3 may have repurposing potentials for NET and MMP9.

3.2.3. Repurposing Potential for SERT, NET, and DAT and Side Effects on hERG. DAT is a well-known mediator of cocaine's behavioral effects. SERT and NET also play significant roles in cocaine responses and inhibitors of each have been considered as potential medications for cocaine dependence. In this vein, we performed BA cross predictions of inhibitors from other datasets in the PPI networks and the *hERG* inhibitor dataset to find compounds with repurposing potentials for either SERT or NET. As discussed above, compounds with repurposing potential should be inactive to their designated target with an experimental BA value of > −9.54 kcal/mol, yet having BA values of < −9.54 kcal/mol for SERT or NET. Since *hERG* is a priority side effect concern for novel medications, the side effect threshold to *hERG* was considered to be −8.18 kcal/mol (K_i = 1 μM).

Figures S12 and S13 in the Supporting Information show the predicted repurposing potentials to SERT and potential *hERG* effects of inhibitors from 59 other datasets. We show that a larger portion of the datasets may qualify for further screening. In some datasets, nearly half of all compounds may have repurposing potential and low *hERG* effects. In the CDK1 dataset, 534 out of 1253 inhibitor compounds possessed repurposing potential, and 507 still remained after considering *hERG* side effects. Among the 392 DHFR inhibitors with repurposing potentials for SERT, 349 had no predicted *hERG* side effect. In addition, DAT and NET datasets provided 685 and 451 such compounds, respectively.

Predictions of repurposing potentials and *hERG* side effects of 59 datasets for NET are provided in Figures S14 and S15 in the Supporting Information. A fairly large number of compounds were obtained from the 59 datasets, according to the prediction. For example, 151, 155, 229, and 382 compounds with repurposing potential to NET and low predicted *hERG* side effects were found in the DHFR, LCK, DAT, and SERT datasets, respectively.

The predicted repurposing potential for DAT and the *hERG* side effect are shown in Figures S16 and S17 in the Supporting Information. Here, we only consider inhibitor datasets in the SERT-NET networks but not in the DAT network. The predictions for the datasets in the DAT network were already reported in our previous work.[39] Most of the datasets have a large portion of the inhibitors whose predicted *hERG* BA values are greater than −8.18 kcal/mol, suggesting no serious potential *hERG* side effect. According to our predictions, several datasets including DPP4, FGFR1, HDAC1, HTR1A, HTR2A, HDM2, NET, SERT, and SRC have approximately half of their inhibitors with high *hERG* side effects. On the other hand, in search for compounds with repurposing potential to DAT, most datasets have a limited number of compounds satisfying repurposing requirements. However, a few datasets contain several compounds with repurposing potential and low predicted *hERG* side effects. For instance, the LCK dataset of 1855 inhibitors contains 17 compounds with repurposing potential to DAT. Fortunately, all 17 compounds are predicted to have no *hERG* side effects, with predicted BA values to *hERG* > −8.18 kcal/mol. The SERT dataset has 88 compounds that are predicted to have repurposing potential to DAT. Forty-one (41) of the 88 compounds are predicted to have no *hERG* risk. In the NET inhibitor dataset, 43 out of the 75 compounds with repurposing potential have low predicted *hERG* potential. In summary, an encouraging number of inhibitor compounds with repurposing potential to DAT and low predicted *hERG* side effects are available from the various datasets in the three networks. According to predictions, 41, 43, 11, 35, and 19 compounds can be obtained respectively from SERT, NET, DPP4, HTR1A, and HTR2A inhibitor datasets.

3.2.4. Possible Side Effects of SERT, NET, and DAT Inhibitors to Other Proteins. Herein, we investigate the possible side effects of inhibitors of SERT, NET, and DAT. Figures S9, S10, and S11 in the Supporting Information show the predicted side effects of SERT, NET, and DAT inhibitor datasets, respectively.

Figures S9 and S10 show the BA predictions of inhibitors of SERT and NET to other proteins. It can be seen that nearly

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

half of the SERT inhibitors may have side effects on D_3R and YES1 proteins in the DAT network. This convinced us to perform complete investigations across networks. Many NET inhibitors may pose side effects on from several targets including LRRK2, Sigma1, SERT, HTR2A, D_4R, and others.

Figure S11 shows that DAT inhibitors could cause side effects on several proteins including GRK5 and STAT3 from BA predictions. In addition, many DAT inhibitors may cause serious side effects through other proteins, including HTR1A, HTR2A, D_2R, D_3R, SERT, NET, and others. HTR1A and HTR2A serve as serotonin receptors, while D_2R, D_3R play roles as dopamine receptors. Such side effects of DAT inhibitors on the receptors may interfere with dopamine or serotonin transmission. HRT1A and HTR2A do not belong to the network of DAT, demonstrating the likelihood that a DAT inhibitor with no side effects in its own network can pose serious side effects on proteins in other networks. It is consequently reasonable to consider the side effects of DAT inhibitors on a larger scale involving more proteins.

3.3. Druggable Property Screening. We performed systematic screenings of ADMET properties, synthetic accessibility (SAS), and *hERG* risk of all inhibitor datasets. Accurate predictions of pharmacokinetic properties are vital for drug design. ADMET (absorption, distribution, metabolism, excretion, and toxicity) includes a diversity of attributes associated with the pharmacokinetic attributes of a compound and is an important factor in drug discovery.[54] In this work, we restrict our attention to six indexes of ADMET, i.e., FDAMDD, $T_{1/2}$ and $F_{20\%}$, log P, log S, and Caco-2, and SAS,[39] as well as a *hERG* risk assessment. The optimal ranges of ADMET properties and SAS are provided in Table 2, while

Table 2. Optimal Ranges of Six Selected ADMET Characteristics and Synthesizability (SAS) Considered in This Work

property	optimal range
FDAMDD	excellent, 0−0.3; medium, 0.3−0.7; poor, 0.7−1.0
$F_{20\%}$	excellent, 0−0.3; medium, 0.3−0.7; poor, 0.7−1.0
log *P*	the proper range: 0−3 log mol/L
log *S*	the proper range: −4−0.5 log mol/L
$T_{1/2}$	excellent, 0−0.3; medium, 0.3−0.7; poor, 0.7−1.0
Caco-2	proper range: > −5.15
SAS	proper range: <6

the BA value > −8.18 kcal/mol is applied as the required range for exempting *hERG* side effects. The available ML models for ADMET[55] and SAS, as well as our models for cross-target prediction, enable us to systematically search for promising compound leads with desired ADMET properties.

Figure 6 illustrates an example of screening datasets of five proteins: SERT, NET, HTR1A, HTR2A, and DPP4. SERT and NET are two of the direct targets of cocaine, and efforts are often dedicated to pharmacological effects on these two proteins in cocaine treatment. Serotonin receptors including HTR1A and HTR2A also draw attention since some pharmacological manipulations were found to reduce cocaine use in preclinical studies.[16] Each column records the ADMET predictions of the given inhibitor dataset, while each row in Figure 6 represents a pair of ADMET characteristics. The orange frames indicate the optimal domains of the specified two characteristics for the inhibitor dataset. Finally, the dot color portrays the experimental binding affinities.

The first row corresponds to the FDA maximum recommended daily dose (FDAMDDs) and the BA to *hERG* (*hERG* BA), which reflect toxicity of potential drug candidates to the human body. A small fraction of SERT and NET datasets lie in the optimal domains outlined by this pair of properties. The restriction by FDAMDD filters out more than two-thirds of inhibitors from all the datasets. The second row stands for the screening of absorption properties $T_{1/2}$ (half-life) and $F_{20\%}$ (human oral bioavailability 20%). The half-life is the amount of time it takes for a drug's active substance to reduce by half in human body, and $T_{1/2}$ indicates the probability of half-life less than 3 h. $F_{20\%}$ represents the probability of an oral drug reaching systemic circulation with <20% of the initial dose remaining. This pair of properties together enforces strict thresholds for drug screening, as we can see that the orange frames only cover a small portion of a given dataset. Moreover, the third row of Figure 6 denotes the screening on log *P* and log *S*, which are the logarithm of the *n*-octanol/water distribution coefficient and aqueous solubility value, respectively. These two screening properties fail many inhibitors, especially from the NET dataset. The last row depicts caco-2 and SAS screening. Caco-2 is commonly used to estimate in vivo permeability of oral drugs, while SAS is designed to estimate the ease of synthesis of druglike molecules. Based on the predictions, most inhibitors of the five datasets are not hard to synthesize. As a result, these two properties allow a large portion of inhibitors to pass the screening.

The ADMET and other properties are important indexes of eligible candidate drug compounds. We anticipate that several of these properties may pose significant challenges when searching for desired drugs to treat cocaine addiction. Reliable ML-based models are in need to accomplish the prediction of these properties. We took advantage of the ADMETlab 2.0 solver[55] to obtain these screening results.

4. DISCUSSION

4.1. Side-Effect Predictions of Existing Experimental Medications. Because of the importance of noradrenergic and serotonergic systems in mediating cocaine effects, some experimental medications targeting these two systems have been investigated. In this study, we utilized ML models to predict the side effects of these medications.

4.1.1. Medications Targeting the Serotonin System. Cocaine dependence is closely linked to deficits in the serotonin system. Some preclinical studies have already shown that self-administration of cocaine can be reduced through serotonin-enhancing medications.[22,23] On the other hand, pharmacologic manipulation of serotonin is associated with the dopamine system[16] and can indirectly modulate the dopamine circuits relevant to dependence. Serotonin-enhancing medications also benefit the dopamine system by indirectly increasing extracellular dopamine levels. Among those serotonergic medications, ibogaine is investigated most frequently. Structures and experimental or predicated BA to various targets of many other experimental medications are shown in the Supporting Information.

4.1.1.1. Ibogaine and Its Derivatives. Ibogaine is a hallucinogenic alkaloid found in the root and bark of the African shrub, Tabernanthe iboga. It may have effects in the treatment of not only cocaine dependence but also for alcohol, opiate, and methamphetamine dependence. The pharmacology of Ibogaine is complex, and it has affinities for *k*-opioid receptors, *N*-methyl-D-aspartate receptors, and $\sigma 1$ and $\sigma 2$

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

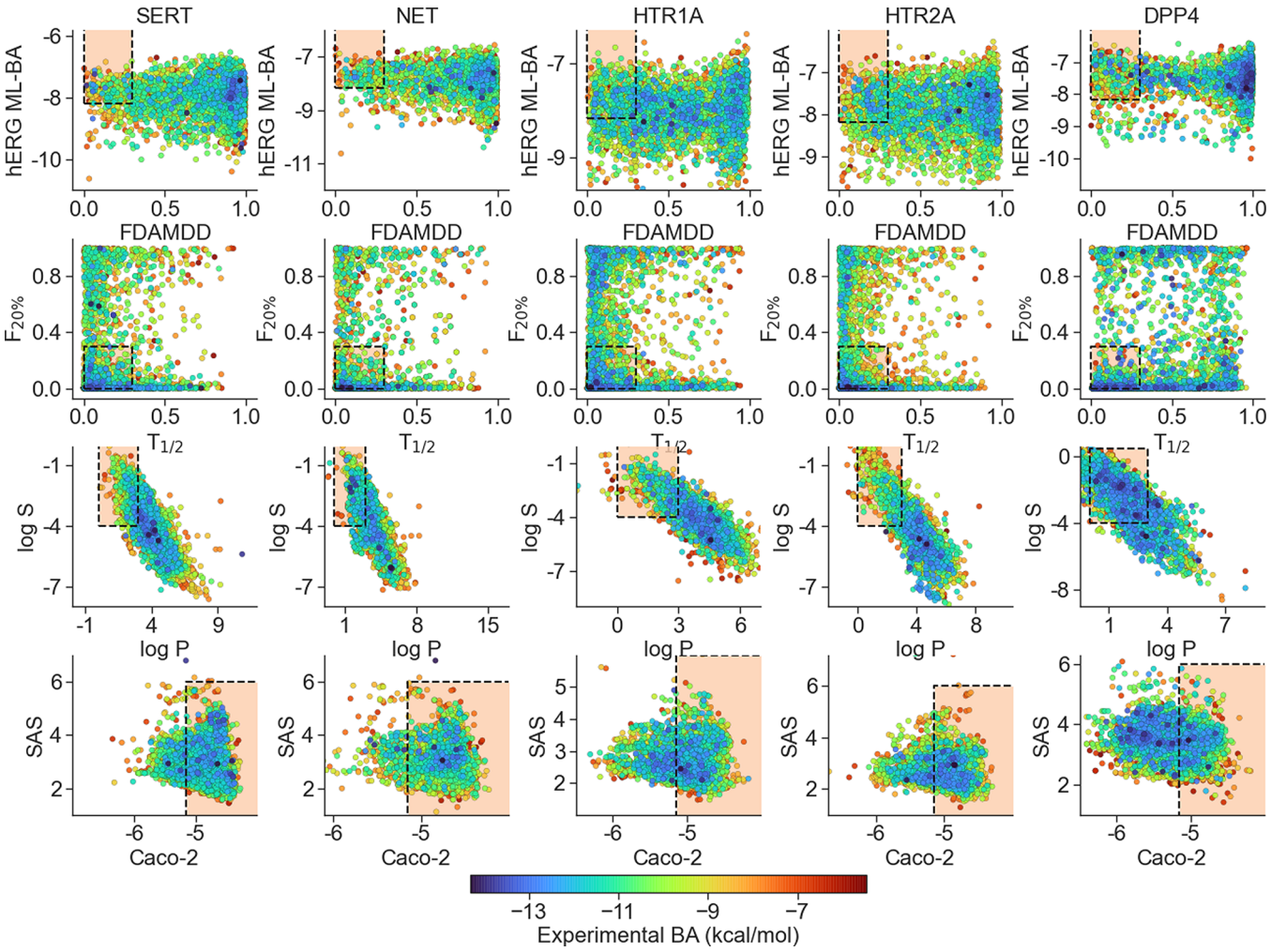


Figure 6. Druggable property screening based on ADMET properties, synthesizability, and *hERG* side effects to compounds from five critical protein datasets: SERT, NET, HTR1A, HTR2A, DPP4. The colors of points represent the experimental BAs to these targets. The *x*- and *y*-axes show predicted ADMET properties, synthesizability, or *hERG* side effects. Orange frames outline the optimal ranges of these properties and side effects.

receptors, as well as DAT and SERT.[56,57] The psychoactive effects of Ibogaine are associated with *k*-opioid receptors,[58] while the agonistic action with serotonin 5-HT_{2A} receptor (HTR2A) may contribute to the hallucinogenic effects.[59] Its actions are complex in mediating different neurotransmitters at the same time. In particular, its interaction with DAT and SERT may underlie its antiaddiction properties for cocaine and it has been promoted as a treatment for addiction in Europe and North America.

Ibogaine inhibits both DAT and SERT with IC_{50} values of 4.0 μM and 0.59 μM, respectively.[60] The mechanism of Ibogaine inhibition of SERT is different from other known inhibitors in the sense that it is not competitive with substrates and it stabilizes the transporter in an inward-open conformation. Ibogaine binds to a site accessible from the cell exterior that does not overlap with the substrate-binding site on SERT[57] and DAT.[61] Its molecular mechanism of inhibiting SERT either by binding to the outward-open conformation or the inward-open conformation has been discussed in detail.[57] Despite its promising effects for the treatment of cocaine addiction, its associated side effects, including death, are a serious concern, and have led to its prohibition in some countries. From 1990 to 2008, 19 fatalities associated with the ingestion of Ibogaine were reported, and 6 of these fatalities were caused by acute heart failure or cardiopulmonary arrest.[62]

One side effect of Ibogaine is that it may cause long QT syndrome at higher doses, perhaps by blocking the *hERG* potassium channel in the heart.[62] The predicted BA value of Ibogaine to *hERG* using our model is −8.43 kcal/mol, and this modest prediction can indicate the potential of cardiac risk. In addition, our models anticipated other high-risk side effects. The predicted BA value to YES1 is −9.71 kcal/mol, and YES1 inhibition is associated with sarcoma and acute myeloid leukemia. Ibogaine is also predicted to have high BAs of −10.69, −10.47, −10.46, and −10.70 kcal/mol to NTRK1, NTRK2, NTRK3, and SYK, respectively.

In recent years, there is increased interest in 18-methoxycoronaridine (18MC), which is a derivative of Ibogaine. It has shown its effectiveness in reducing the self-administration of cocaine, morphine, methamphetamine, nicotine, and sucrose in preclinical models.[63] It has similar pharmacological effects to those of Ibogaine, but it does not cause tremors, Purkinje cell dysfunction, or toxicity in the brain.[64] Moreover, 18MC has no affinity for SERT, in contrast to Ibogaine, and so it provides an enhanced safety profile to

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

humans compared to ibogaine and is under clinical trials. The predicted *hERG* BA of 18MC by our model is −7.67 kcal/mol, which reflects moderate potential in incurring heart issues. However, it is predicted to have BA values of −10.42, 10.15, −10.20, 10.15, and −9.91 kcal/mol, respectively, for SSTR5, YES1, LRRK2, VMAT2, and CNR2. Diseases associated with SSTR5 include acromegaly, pituitary adenoma, and prolactin-pecreting. YES1 is associated with sarcoma. These strong off-target binding affinities may indicate the potential risk of such diseases or related side effects when 18MC is tested in humans.

4.1.1.2. Selective Serotonin Reuptake Inhibitors. Selective serotonin reuptake inhibitors (SSRIs) are typically used as antidepressants. They increase the extracellular level of serotonin by limiting its reuptake into the presynaptic cell. Generally, SSRIs have a stronger affinity to SERT than to DAT or NET. Because of their enhancement of extracellular serotonin levels, attempts have been made to use SSRIs to treat cocaine addiction, and some clinical trials using SSRIs have shown some promise in treating cocaine addiction.

Fluoxetine is an SSRI approved by FDA to treat several psychiatric disorders including depressive depression, bulimia nervosa, and others. Some preclinical studies have shown the effectiveness of fluoxetine in the treatment of cocaine addictions, but clinical trials have yielded mixed results. Some studies showed the efficacy of fluoxetine in significantly reducing cocaine use,[65,66] while others showed that fluoxetine is not effective in altering cocaine effects.[24] However, it is encouraging that fluoxetine has a tendency to be more effective if higher doses are used in the cocaine treatment.[24,65] Our ML models show that fluoxetine has low binding affinities to most of the 61 targets in our network except for SERT, with predicted BA lower than −10 kcal/mol. The high binding affinity to SERT is reasonable since fluoxetine is an SSRI. Fluoxetine is predicted to have binding affinity higher than −9.0 kcal/mol to 55 out of the 61 targets, which reflects its low side effects on the targets in the network. Since fluoxetine is already an FDA approved medication, such low side effects are anticipated. It is predicted to have binding affinity of −8.25 kcal/mol to *hERG*, which indicates relatively low potential to cause prolongation of the QT interval.

Sertraline is also an SSRI used as antidepressant for various psychiatric conditions. It is given as a generic medication, and it was the most prescribed psychiatric medication in the USA in 2016. It has also been tested in clinical trials for the treatment of cocaine dependence. Sertraline is effective in reducing cocaine cravings and can produce delays in relapse in recently abstinent cocaine abusers with depressive symptoms. According to our ML model predictions, sertraline has low binding affinities to most of the 61 targets in the networks. It is predicted to have relatively high BAs of −9.15 and −9.51 kcal/mol for D_4R and SYK, respectively. The predicted BA to *hERG* is −7.75 kcal/mol, which suggests a low potential for causing heart issues.

Citalopram is another SSRI antidepressant. Recently, some very encouraging results were obtained in reducing cocaine use when given in combination with contingency management.[67] Side effects were observed to be mild. These studies provided support that citalopram combined with behavioral therapy can be a promising treatment for cocaine dependence. According to our ML model predictions, citalopram has low BAs to most of the 61 protein targets. It is predicted to have values of −9.30, −9.34, −9.72, and −10.11 kcal/mol to SSTR5, D_3R, CNR1, and HTR2A, respectively. SSTR5 is associated with diseases including acromegaly and prolactin-secreting pituitary adenoma. The binding affinity of citalopram indicates potential risk for the aforementioned diseases. Its strong binding affinity to 5-HT_{2A} and D_3R indicate its effect in the transmission of neurotransmitters serotonin and dopamine. 5-HT_{2A} and D_3R have been investigated as pharmacological targets in the treatment of cocaine dependence, and citalopram may cause unexpected effects due to binding these two targets. The predicted BA to *hERG* is −8.16 kcal/mol, consistent with the announcement that "causes dose-dependent QT interval prolongation" by FDA.[16]

4.1.1.3. 5-HT3 Receptor Antagonists. Efforts to target the 5-HT_{1B}, 5-HT_{2A}, and 5-HT_3 receptors have been made due to their important roles in potential cocaine addiction mechanisms. Ondansetron, which is a 5-HT_3 receptor antagonist, has been investigated in clinical trials for cocaine addiction treatment. Preclinical studies showed its efficacy in abolishing the reinstatement of cocaine administration,[21] and further studies showed that ondansetron can be particularly effective in reducing oral cocaine self-administration when given during the acute cocaine withdrawal period.[68] Our ML models show that it has binding affinity values of −9.82, −9.64, −9.86, and −9.4 kcal/mol for D_4R, Sigma1, HTR1A, and DPP4, respectively. The high binding affinities to these targets indicate the potential for side effects mediated by them.

4.1.2. Medications Targeting the Noradrenergic System. Some preclinical and clinical studies showed that pharmacological manipulations of the noradrenergic systems could be a potential treatment for cocaine addiction.[69] In noradrenergic systems, norepinephrine (NE) is the main chemical messenger and plays a contributing role in mediating the rewarding effects of cocaine. NET is regarded as a potential target for treatment of cocaine addiction. Atomoxetine and reboxetine are two selective noradrenaline reuptake inhibitors with NET-blocking effects.

Reboxetine, which is an antidepressant medication, has been tested in the treatment of cocaine addiction[35] and reported as an effective and safe therapeutic option. However, more rigorous double-blind studies of reboxetine need to be performed before its efficacy in the treatment of cocaine dependence can be fully confirmed. Another promising outcome was reported for the combination of reboxetine with the SSRI escitalopram.[70] Predictions from our ML models shows side effects to D_4R, Sigma1, D_3R, and GRM2, based on the corresponding predicted BA values of −9.61, −9.55, −9.55, and −9.37 kcal/mol. In addition, the BA value to *hERG* is predicted to be −7.80 kcal/mol.

Atomoxetine is a selective NET inhibitor and has been approved for the treatment of ADHD, and recently found to prevent relapse to cocaine use. Other preclinical studies have shown that atomoxetine can significantly attenuate cue-induced relapse to cocaine seeking after abstinence, which reflects the potential of atomoxetine as an effective treatment in preventing relapse in cocaine addiction.[34] Safety studies were performed on atomoxetine when used with intravenous cocaine on cocaine-experienced participants and found that atomoxetine can be safely tolerated.[71] Using our ML models, it is predicted to have BAs of −9.27 and −9.21 kcal/mol, respectively, for SPR and APP. Side effects to these targets, discussed above, may be anticipated. Atomoxetine is already an FDA-approved medication with low *hERG* side effects.

4.2. Nearly Optimal Leads from Our Systematic Screening and Repurposing. We are dedicated to finding

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

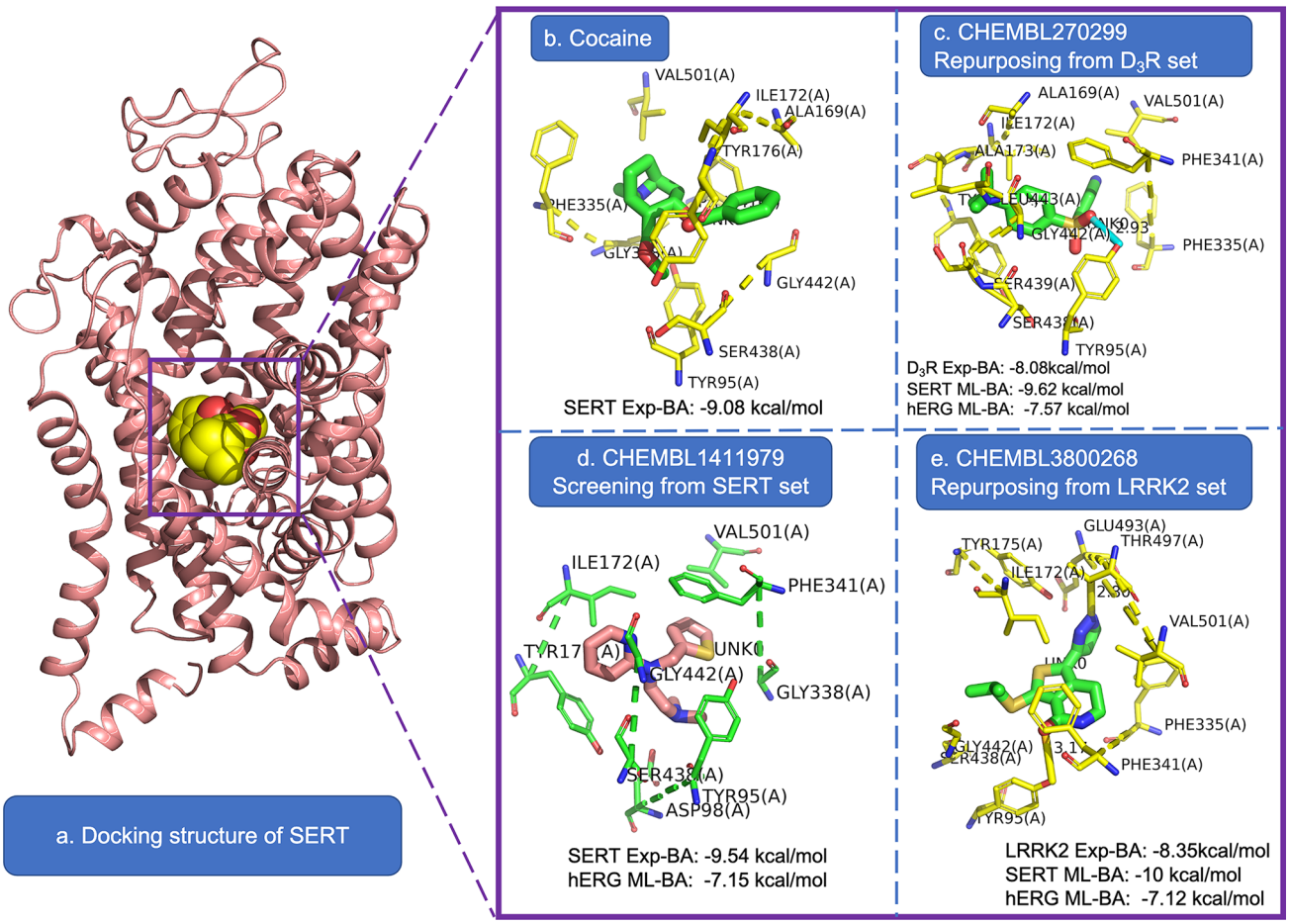


Figure 7. Docking structure of SERT with cocaine, ChEMBL270299, ChEMBL14111979, and ChEMBL3800268, using AutoDock Vina software. The experimental or predicted binding affinities of cocaine, and three compounds from screening and repurposing, are also presented.

promising lead compounds targeting SERT and NET. Screening and repurposing are two key processes for us to filter molecules from 61 available inhibitor datasets, as summarized in Figure 3. In addition to drug potency and side effects, we also consider the six ADMET properties in Table 2, as well as synthetic accessibility. In total, we required 68 criteria to be satisfied for a drug to be considered as a potential reliable agent. In screening, we started with potent inhibitors (experimental BA values of < -9.54 kcal/mol) from the inhibitor dataset of the target protein. Those potent inhibitors then were examined for their side effects on the other 59 proteins with a uniform BA requirement, i.e., predicted BA values higher than -9.54 kcal/mol. A stricter BA threshold is applied for *hERG*, with predicted BA values higher than -8.18 kcal/mol required. These predicted side effect BA values were obtained from our proteme models of the 61 proteins. For repurposing, we started with inhibitors of low BA value (experimental BA > -9.54 kcal/mol) to their designated target protein yet having high predicted BA (< -9.54 kcal/mol) to SERT or NET. Following this, side-effect examinations were performed on the other 59 proteins. Finally, excellent ADMET properties and synthesizability specified in Table 2 had to be satisfied for both processes to SERT or NET.

Several compounds through screening or repurposing were found as potential agents for SERT. They all have excellent ADMET properties and are readily prepared. Compound ChEMBL1411979 was obtained via screening from the SERT inhibitor dataset. It has an experimental BA of -9.54 kcal/mol to SERT, and its potential side effects on *hERG* are low with a predicted BA of -7.15 kcal/mol. Moreover, its predicted BA is stronger than cocaine (experimental BA value = -9.08 kcal/mol). This observation deserves further investigation. Seven additional compounds were predicted to have low side effects against the other 59 targets. Among those, ChEMBL3800268, from the LRRK2 inhibitor dataset, has a low BA of -8.35 kcal/mol to its designated target, but it is predicted to have a BA of -10.01 kcal/mol to SERT. Its potential side effects on other proteins are weak. For example, its predicted BA to *hERG* is as low as -7.12 kcal/mol. Another compound, ChEMBL270299, from the D_3R dataset is predicted to have a BA of -9.62 kcal/mol for SERT. It has a weak potential *hERG* side effect with a predicted BA of -7.57 kcal/mol, while it has a low experimental BA of -8.08 kcal/mol to its designated target D_3R. The information on other compounds and their predicted BAs can be found in the Supporting Information. All these compounds are predicted to have more potent BAs to SERT than that of cocaine.

In searching for effective inhibitors for NET, compound ChEMBL454675 was found to satisfy all requirements. It has an experimental BA of -11 kcal/mol for NET while its predicted BA to *hERG* is just -7.52 kcal/mol. It may provide good pharmacological effects, as its BA for NET is much higher than that of cocaine (-9.33 kcal/mol). No compounds satisfying our criteria were found for NET by repurposing from the other 60 inhibitor datasets.

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

Figure 7 provides the docking information on cocaine and three nearly optimal lead compounds at the central site of SERT. The residuals at active sites near these compounds are also labeled. More docking information about other nearly optimal lead compounds can be found in the Supporting Information. These docking predictions were implemented by software AutoDock Vina.[72]

5. CONCLUSION

Substance use disorder (SUD) is associated with a variety of mental/emotional, physical, and behavioral problems, including chronic guilt, the seeking and taking of drugs despite adverse consequences, driving or making important decisions while intoxicated, and physiological withdrawal symptoms. As a specific example of SUD, millions of people are addicted to cocaine. However, there currently is no therapeutic approved by the U.S. Food and Drug Administration to address it, in part because cocaine addiction involves intricate molecular mechanisms. DAT, SERT, and NET are each associated with a complex interactome network, and one cannot develop anticocaine addiction medications without considering all interactome networks.

We proposed proteome-informed machine learning studies of cocaine addiction as the first interactome-based machine learning/deep learning (ML/DL) protocol for anticocaine addiction lead discovery,[39] but only the DAT interactome network was considered. The present work extends these previous studies to SERT and NET, enabling us to perform a comprehensive evaluation of existing potential cocaine addiction inhibitors. With molecular features generated by autoencoder (AE), gradient boosted decision tree (GBDT) is used to build robust predictive models, while multitask deep neural network (MT-DNN) is also leveraged to enhance predictive performance for models with small datasets. Using such ML/DL tools, we have considered repurposing existing inhibitors and screened for side effects as well as ADMET properties. After this rigorous screening, we have identified a small group of promising compounds.

The knowledge and understanding obtained from the present work will be employed for the automated generation and screening of anticocaine addiction candidates using our generative network complex.[48] The next step is to test the resulting leads in in vitro and animal assays. It will be critical to examine toxicity and blood-brain barrier permeability characteristics of candidate compounds using cell-based assays to refine our lists and prioritize compounds for animal models. This may include iterative medicinal efforts to optimize critical qualities to identify compounds with the greatest therapeutic potential. These compounds must then be tested in rodent models, since in silico and cell-based assays cannot definitively determine the behavioral effects of a drug. Cocaine locomotor sensitization in mice can uncover the ability of a compound to block the physiological and psychomotor effects of the drug while cocaine-conditioned place preference can determine whether a compound can prevent a drug reward or drug environment.[73−75] Compounds with effects in either of these assays would then be good candidates for testing in the more time-consuming but more rigorous cocaine self-administration model. Examining characteristics like acquisition, breakpoint, extinction, and context- or cue-reinstatement can reveal whether a compound might be useful to block key aspects of cocaine addiction like drug seeking, craving, and relapse.[74,76] Of course, compounds that produce promising results in animal testing would then move into further development.

Finally, our work establishes a new protocol for artificial intelligence (AI)-based nearly optimal lead discovery that can be applied to any disease for which some portion of the molecular etiology has been studied. We hope this technology can be applied to many other neuropsychiatric diseases going forward to uncover new classes of therapeutic agents to improve disease outcomes.

DATA AND SOFTWARE AVAILABILITY

The 61 cocaine-addiction related datasets studied in this work are publicly available at: https://weilab.math.msu.edu/DataLibrary/2D/. Our source code and trained autoencoder model for LV-FP generation can be found at https://github.com/WeilabMSU/antoencoder-v01.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00002.

Section S1, datasets and performance summary; section S2, validation comparisons of LV-GBDT models; section S3, more results on cross-target predictions; and section S4, cross-target predictions of SERT, NET, DAT, and hERG (PDF)

AUTHOR INFORMATION

Corresponding Author

Guo-Wei Wei − *Department of Mathematics, Department of Biochemistry and Molecular Biology, and Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan 48824, United States;* orcid.org/0000-0002-5781-2937; Email: weig@msu.edu

Authors

Hongsong Feng − *Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, United States*

Kaifu Gao − *Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, United States;* orcid.org/0000-0001-7574-4870

Dong Chen − *Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, United States*

Li Shen − *Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, United States*

Alfred J Robison − *Department of Physiology, Michigan State University, East Lansing, Michigan 48824, United States;* orcid.org/0000-0002-4150-9570

Edmund Ellsworth − *Department of Pharmacology & Toxicology, Michigan State University, East Lansing, Michigan 48824, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00002

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by NIH Grant Nos. GM126189 and DA040621, NSF Grant Nos. DMS-2052983, DMS-1761320, and IIS-1900473, NASA 80NSSC21M0023,

MSU Foundation, MSU Institute for Cyber-Enabled Research, Michigan Economic Development Corporation, George Mason University Award No. PD45722, Bristol-Myers Squibb 65109, and Pfizer.

REFERENCES

(1) Morris, D. C. Cocaine heart disease. *Hospital Pract.* **1991**, *26*, 83−92.
(2) Chaisson, R. E.; Bacchetti, P.; Osmond, D.; Brodie, B.; Sande, M. A.; Moss, A. R. Cocaine use and HIV infection in intravenous drug users in San Francisco. *JAMA* **1989**, *261*, 561−565.
(3) de Carvalho, H. B.; Seibel, S. D. Crack cocaine use and its relationship with violence and HIV. *Clinics* **2009**, *64*, 857−866.
(4) Beuming, T.; Kniazeff, J.; Bergmann, M. L.; Shi, L.; Gracia, L.; Raniszewska, K.; Newman, A. H.; Javitch, J. A.; Weinstein, H.; Gether, U.; Loland, C. J. The binding sites for cocaine and dopamine in the dopamine transporter overlap. *Nat. Neurosci.* **2008**, *11*, 780−789.
(5) Kristensen, A. S.; Andersen, J.; Jørgensen, T. N.; Sørensen, L.; Eriksen, J.; Loland, C. J.; Strømgaard, K.; Gether, U. SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacol. Rev.* **2011**, *63*, 585−640.
(6) Elliott, J. M.; Beveridge, T. J. Psychostimulants and monoamine transporters: upsetting the balance. *Curr. Opin. Pharmacol.* **2005**, *5*, 94−100.
(7) Tzschentke, T. M.; Schmidt, W. J. Functional relationship among medial prefrontal cortex, nucleus accumbens, and ventral tegmental area in locomotion and reward. *Crit. Rev. Neurobiol.* **2000**, *14*, 12.
(8) Koob, G. F. Neurobiology of addiction: toward the development of new therapies. *Ann. N.Y. Acad. Sci.* **2000**, *909*, 170−185.
(9) Newman, A. H.; Ku, T.; Jordan, C. J.; Bonifazi, A.; Xi, Z.-X. New drugs, old targets: tweaking the dopamine system to treat psychostimulant use disorders. *Ann. Rev. Pharmacol. Toxicol.* **2021**, *61*, 609−628.
(10) Gorelick, D. A.; Gardner, E. L.; Xi, Z.-X. Agents in development for the management of cocaine abuse. *Drugs* **2004**, *64*, 1547−1573.
(11) Broderick, P. A.; Hope, O.; Okonji, C.; Rahni, D. N.; Zhou, Y. Clozapine and cocaine effects on dopamine and serotonin release in nucleus accumbens during psychostimulant behavior and withdrawal. *Progress Neuro-Psychopharmacol. Biol. Psych.* **2004**, *28*, 157−171.
(12) Dworkin, S. I.; Co, C.; Smith, J. E. Rat brain neurotransmitter turnover rates altered during withdrawal from chronic cocaine administration. *Brain Res.* **1995**, *682*, 116−126.
(13) Parsons, L. H.; Koob, G. F.; Weiss, F. Extracellular serotonin is decreased in the nucleus accumbens during withdrawal from cocaine self-administration. *Behav. Brain Res.* **1995**, *73*, 225−228.
(14) Rocha, B. A.; Fumagalli, F.; Gainetdinov, R. R.; Jones, S. R.; Ator, R.; Giros, B.; Miller, G. W.; Caron, M. G. Cocaine self-administration in dopamine-transporter knockout mice. *Nat. Neurosci.* **1998**, *1*, 132−137.
(15) Sora, I.; Hall, F. S.; Andrews, A. M.; Itokawa, M.; Li, X.-F.; Wei, H.-B.; Wichems, C.; Lesch, K.-P.; Murphy, D. L.; Uhl, G. R. Molecular mechanisms of cocaine reward: combined dopamine and serotonin transporter knockouts eliminate cocaine place preference. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 5300−5305.
(16) Filip, M.; Frankowska, M.; Zaniewska, M.; Golda, A.; Przegaliński, E. The serotonergic system and its role in cocaine addiction. *Pharmacol. Rep.* **2005**, *57*, 685−700.
(17) Halliday, G. M.; Törk, I. Serotonin-like immunoreactive cells and fibres in the rat ventromedial mesencephalic tegmentum. *Brain Res. Bull.* **1989**, *22*, 725−735.
(18) Hoyer, D.; Hannon, J. P.; Martin, G. R. Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacol., Biochem. Behav.* **2002**, *71*, 533−554.
(19) Vanderhoek, G.; Cooper, S. Evidence That Ondansetron, A Selective 5-HT3 Antagonist, Reduces Cocaine Psychomotor Stimulant Effects in the Rat. *Psychopharmacology.* **1990**, *101*, S59−S59.
(20) Cunningham, K. A.; Anastasio, N. C.; Fox, R. G.; Stutz, S. J.; Bubar, M. J.; Swinford, S. E.; Watson, C. S.; Gilbertson, S. R.; Rice, K. C.; Rosenzweig-Lipson, S.; Moeller, F. G. Synergism between a serotonin 5-HT2A receptor (5-HT2AR) antagonist and 5-HT2CR agonist suggests new pharmacotherapeutics for cocaine addiction. *ACS Chem. Neurosci.* **2013**, *4*, 110−121.
(21) Davidson, C.; Lee, T. H.; Xiong, Z.; Ellinwood, E. H. Ondansetron given in the acute withdrawal from a repeated cocaine sensitization dosing regimen reverses the expression of sensitization and inhibits self-administration. *Neuropsychopharmacology* **2002**, *27*, 542−553.
(22) Carroll, M. E.; Lac, S. T.; Asencio, M.; Kragh, R. Fluoxetine reduces intravenous cocaine self-administration in rats. *Pharmacol., Biochem. Behav.* **1990**, *35*, 237−244.
(23) Peltier, R.; Schenk, S. Effects of serotonergic manipulations on cocaine self-administration in rats. *Psychopharmacology* **1993**, *110*, 390−394.
(24) Batki, S. L.; Washburn, A. M.; Delucchi, K.; Jones, R. T. A controlled trial of fluoxetine in crack cocaine dependence. *Drug Alcohol Dependence* **1996**, *41*, 137−142.
(25) Smythies, J. Section III. The norepinephrine system. *Int. Rev. Neurobiol.* **2005**, *64*, 173−211.
(26) Tellioglu, T.; Robertson, D. Genetic or acquired deficits in the norepinephrine transporter: current understanding of clinical implications. *Expert Rev. Mol. Med.* **2001**, *3*, 1−10.
(27) Bylund, D. B. International Union of Pharmacology nomenclature of adrenoceptors. *Pharmacol. Rev.* **1994**, *46*, 121−136.
(28) Paladini, C. A.; Williams, J. T. Noradrenergic inhibition of midbrain dopamine neurons. *J. Neurosci.* **2004**, *24*, 4568−4575.
(29) Blanc, G.; Trovero, F.; Vezina, P.; Herve, D.; Godeheu, A.-M.; Glowinski, J.; Tassin, J.-P. Blockade of prefronto-cortical α1-adrenergic receptors prevents locomotor hyperactivity induced by subcortical D-amphetamine injection. *Eur. J. Neurosci.* **1994**, *6*, 293−298.
(30) Erb, S.; Hitchcott, P. K.; Rajabi, H.; Mueller, D.; Shaham, Y.; Stewart, J. Alpha-2 adrenergic receptor agonists block stress-induced reinstatement of cocaine seeking. *Neuropsychopharmacology* **2000**, *23*, 138−150.
(31) Leri, F.; Flores, J.; Rodaros, D.; Stewart, J. Blockade of stress-induced but not cocaine-induced reinstatement by infusion of noradrenergic antagonists into the bed nucleus of the stria terminalis or the central nucleus of the amygdala. *J. Neurosci.* **2002**, *22*, 5713−5718.
(32) Kampman, K. M.; Dackis, C.; Lynch, K. G.; Pettinati, H.; Tirado, C.; Gariti, P.; Sparkman, T.; Atzram, M.; O'Brien, C. P. A double-blind, placebo-controlled trial of amantadine, propranolol, and their combination for the treatment of cocaine dependence in patients with severe cocaine withdrawal symptoms. *Drug Alcohol Dependence* **2006**, *85*, 129−137.
(33) Sofuoglu, M.; Brown, S.; Babb, D. A.; Pentel, P. R.; Hatsukami, D. K. Carvedilol affects the physiological and behavioral response to smoked cocaine in humans. *Drug Alcohol Dependence* **2000**, *60*, 69−76.
(34) Economidou, D.; Dalley, J. W.; Everitt, B. J. Selective norepinephrine reuptake inhibition by atomoxetine prevents cue-induced heroin and cocaine seeking. *Biol. Psychiatry* **2011**, *69*, 266−274.
(35) Szerman, N.; Peris, L.; Mesias, B.; Colis, P.; Rosa, J.; Prieto, A. Reboxetine for the treatment of patients with Cocaine Dependence Disorder. *Human Psychopharmacol.: Clin. Exp.* **2005**, *20*, 189−192.
(36) Kampman, K. M. New medications for the treatment of cocaine dependence. *Psychiatry (Edgmont)* **2005**, *2*, 44.
(37) Food and Drug Administration; HHS. International Conference on Harmonisation; guidance on S7B nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals; availability. Notice. *Fed. Reg.* **2005**, *70*, 61133−61134.
(38) Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.;

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719

Bork, P.; Jensen, L. J.; Mering, C. v. STRING v11: protein−protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.h* **2019**, *47*, D607−D613.
(39) Gao, K.; Chen, D.; Robison, A. J.; Wei, G.-W. Proteome-informed machine learning studies of cocaine addiction. *J. Phys. Chem. Lett.* **2021**, *12*, 11122−11134.
(40) Wu, K.; Wei, G.-W. Quantitative toxicity prediction using topology based multitask deep neural networks. *J. Chem. Inf. Model.* **2018**, *58*, 520−531.
(41) Weinshenker, D.; Schroeder, J. P. There and back again: a tale of norepinephrine and drug addiction. *Neuropsychopharmacology* **2007**, *32*, 1433−1451.
(42) Burmeister, J. J.; Lungren, E. M.; Kirschner, K. F.; Neisewander, J. L. Differential roles of 5-HT receptor subtypes in cue and cocaine reinstatement of cocaine-seeking behavior in rats. *Neuropsychopharmacology* **2004**, *29*, 660−668.
(43) Filip, M. Role of serotonin (5-HT) 2 receptors in cocaine self-administration and seeking behavior in rats. *Pharmacol. Rep.* **2005**, *57*, 35−46.
(44) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic acids research* **2012**, *40*, D1100−D1107.
(45) Gao, K.; Nguyen, D. D.; Sresht, V.; Mathiowetz, A. M.; Tu, M.; Wei, G.-W. Are 2D fingerprints still valuable for drug discovery? *Phys. Chem. Chem. Phys.* **2020**, *22*, 8373−8390.
(46) Nguyen, D. D.; Cang, Z.; Wei, G.-W. A review of mathematical representations of biomolecular data. *Phys. Chem. Chem. Phys.* **2020**, *22*, 4343−4367.
(47) Winter, R.; Montanari, F.; Steffen, A.; Briem, H.; Noé, F.; Clevert, D.-A. Efficient multi-objective molecular optimization in a continuous latent space. *Chem. Sci.* **2019**, *10*, 8016−8024.
(48) Gao, K.; Nguyen, D. D.; Tu, M.; Wei, G.-W. Generative network complex for the automated generation of drug-like molecules. *J. Chem. Inf. Model.* **2020**, *60*, 5682−5698.
(49) 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*, 513−530.
(50) Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078,* 2014.
(51) Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735−1780.
(52) Cang, Z.; Wei, G.-W. Integration of element specific persistent homology and machine learning for protein-ligand binding affinity prediction. *Int. J. Numer. Methods Biomed. Eng.* **2018**, *34*, No. e2914.
(53) Lee, K. H.; Fant, A. D.; Guo, J.; Guan, A.; Jung, J.; Kudaibergenova, M.; Miranda, W. E.; Ku, T.; Cao, J.; Wacker, S.; Duff, H. J.; Newman, A. H.; Noskov, S. Y.; Shi, L. Toward reducing hERG affinities for DAT inhibitors with a combined machine learning and molecular modeling approach. *J. Chem. Inf. Model.* **2021**, *61*, 4266−4279.
(54) Wager, T. T.; Hou, X.; Verhoest, P. R.; Villalobos, A. Central nervous system multiparameter optimization desirability: application in drug discovery. *ACS Chem. Neurosci.* **2016**, *7*, 767−775.
(55) Xiong, G.; Wu, Z.; Yi, J.; Fu, L.; Yang, Z.; Hsieh, C.; Yin, M.; Zeng, X.; Wu, C.; Lu, A.; Chen, X.; Hou, T.; Cao, D. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res.* **2021**, *49*, W5.
(56) Glick, S. D.; Maisonneuve, I. M. Mechanisms of antiaddictive actions of ibogaine a. *Ann. N.Y. Acad. Sci.* **1998**, *844*, 214−226.
(57) Coleman, J. A.; Yang, D.; Zhao, Z.; Wen, P.-C.; Yoshioka, C.; Tajkhorshid, E.; Gouaux, E. Serotonin transporter−ibogaine complexes illuminate mechanisms of inhibition and transport. *Nature* **2019**, *569*, 141−145.
(58) Zubaran, C.; Shoaib, M.; Stolerman, I.; Pablo, J.; Mash, D. Noribogaine generalization to the ibogaine stimulus: correlation with noribogaine concentration in rat brain. *Neuropsychopharmacology* **1999**, *21*, 119−126.
(59) Wei, D.; Maisonneuve, I.; Kuehne, M.; Glick, S. Acute iboga alkaloid effects on extracellular serotonin (5-HT) levels in nucleus accumbens and striatum in rats. *Brain Res.* **1998**, *800*, 260−268.
(60) Efange, S. M.; Mash, D. C.; Khare, A. B.; Ouyang, Q. Modified ibogaine fragments: Synthesis and preliminary pharmacological characterization of 3-ethyl-5-phenyl-1,2,3,4,5,6-hexahydroazepino [4, 5-b] benzothiophenes. *J. Med. Chem.* **1998**, *41*, 4486−4491.
(61) Bulling, S.; Schicker, K.; Zhang, Y.-W.; Steinkellner, T.; Stockner, T.; Gruber, C. W.; Boehm, S.; Freissmuth, M.; Rudnick, G.; Sitte, H. H.; Sandtner, W. The mechanistic basis for noncompetitive ibogaine inhibition of serotonin and dopamine transporters. *J. Biol. Chem.* **2012**, *287*, 18524−18534.
(62) Koenig, X.; Hilber, K. The anti-addiction drug ibogaine and the heart: a delicate relation. *Molecules* **2015**, *20*, 2208−2228.
(63) Glick, S. D.; Maisonneuve, I. M. Development of novel medications for drug addiction: The legacy of an African shrub. *Ann. N.Y. Acad. Sci.* **2000**, *909*, 88−103.
(64) Glick, S.; Kuehne, M.; Maisonneuve, I.; Bandarage, U.; Molinari, H. 18-Methoxycoronaridine, a non-toxic iboga alkaloid congener: effects on morphine and cocaine self-administration and on mesolimbic dopamine release in rats. *Brain Res.* **1996**, *719*, 29−35.
(65) Walsh, S. L.; Preston, K. L.; Sullivan, J. T.; Fromme, R.; Bigelow, G. E. Fluoxetine alters the effects of intravenous cocaine in humans. *J. Clin. Psychopharmacol.* **1994**, *14*, 396−407.
(66) Schmitz, J. M.; Rhoades, H. M.; Elk, R.; Creson, D.; Hussein, I.; Grabowski, J. Medication take-home doses and contingency management. *Exp. Clin. Psychopharmacol.* **1998**, *6*, 162.
(67) Moeller, F. G.; Schmitz, J. M.; Steinberg, J. L.; Green, C. M.; Reist, C.; Lai, L. Y.; Swann, A. C.; Grabowski, J. Citalopram combined with behavioral therapy reduces cocaine use: a double-blind, placebo-controlled trial. *Am. J. Drug Alcohol Abuse* **2007**, *33*, 367−378.
(68) Davidson, C.; Lazarus, C.; Lee, T. H.; Ellinwood, E. H. Ondansetron, given during the acute cocaine withdrawal, attenuates oral cocaine self-administration. *Eur. J. Pharmacol.* **2004**, *503*, 99−102.
(69) Sofuoglu, M.; Sewell, R. A. Norepinephrine and stimulant addiction. *Addiction Biol.* **2009**, *14*, 119−129.
(70) Camarasa, X.; Lopez-Martinez, E.; Duboc, A.; Khazaal, Y.; Zullino, D. F. *Progress Neuro-Psychopharmacol. Biol. Psychiat.* **2005**, *29*, 165−168.
(71) Cantilena, L.; Kahn, R.; Duncan, C. C.; Li, S.-H.; Anderson, A.; Elkashef, A. Safety of atomoxetine in combination with intravenous cocaine in cocaine-experienced participants. *J.Addiction Med.* **2012**, *6*, 265.
(72) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2009**, *31*, 455−461.
(73) Gajewski, P. A.; Eagle, A. L.; Williams, E. S.; Manning, C. E.; Lynch, H.; McCornack, C.; Maze, I.; Heller, E. A.; Robison, A. Epigenetic regulation of hippocampal fosb expression controls behavioral responses to cocaine. *J. Neurosci.* **2019**, *39*, 8305−8314.
(74) Doyle, M. A.; Bali, V.; Eagle, A. L.; Stark, A. R.; Fallon, B.; Neve, R. L.; Robison, A.; Mazei-Robison, M. S. Serum-and glucocorticoid-inducible kinase 1 activity in ventral tegmental area dopamine neurons regulates cocaine conditioned place preference but not cocaine self-administration. *Neuropsychopharmacology* **2021**, *46*, 1−10.
(75) Prus, A. J.; James, J. R.; Rosecrans, J. A. Conditioned place preference. In *Methods of Behavioral Analysis in Neuroscience*; Buccafusco, J. J., Ed.; CRC Press/Routledge/Taylor & Francis Group, 2009; pp 59−76.
(76) Thomsen, M.; Caine, S. B. Intravenous drug self-administration in mice: practical considerations. *Behavior Genetics* **2007**, *37*, 101−118.

https://doi.org/10.1021/acs.jctc.2c00002
J. Chem. Theory Comput. 2022, 18, 2703−2719