Benchmarking Deep Graph Generative Models for Optimizing New Drug Molecules for COVID-19

Logan Ward
Argonne National Laboratory
lward@anl.gov

Jenna A. Billrey
Pacific Northwest National Laboratory
jenna.pope@pnnl.gov

Sutanay Choudhury
Pacific Northwest National Laboratory
sutanay.choudhury@pnnl.gov

Neeraj Kumar
Pacific Northwest National Laboratory
neeraj.kumar@pnnl.gov

Ganesh Sivaraman
Argonne National Laboratory
gsivaraman@anl.gov

ABSTRACT
Design of new drug compounds with target properties is a key area of research in generative modeling. We present a small drug molecule design pipeline based on graph-generative models and a comparison study of two state-of-the-art graph generative models for designing COVID-19 targeted drug candidates: 1) a variational autoencoder-based approach (VAE) that uses prior knowledge of molecules that have been shown to be effective for earlier coronavirus treatments and 2) a deep Q-learning method (DQN) that generates optimized molecules without any proximity constraints. We evaluate the novelty of the automated molecule generation approaches by validating the candidate molecules with drug-protein binding affinity models. The VAE method produced two novel molecules with similar structures to the antiretroviral protease inhibitor Indinavir that show potential binding affinity for the SARS-CoV-2 protein target 3-chymotrypsin-like protease (3CL-protease).

1 INTRODUCTION
Deep learning has demonstrated the potential to revolutionize drug design by reducing the initial search space in the early stages of discovery [25, 33, 49, 51]. By applying the appropriate algorithm trained on the appropriate data, novel molecules can be generated with target properties [13, 24, 29, 43]. Here, we evaluate methods for the high-performance intelligent design of small-molecule drug candidates with anti-SARS activity, with a specific focus on SARS-CoV-2, otherwise known as COVID-19. However, discovering potential lead candidates for COVID-19 presents a challenge to the scientific community due to the long timescale of the drug development process. There is a need to accelerate the design process through AI-driven workflows for effective drug compound development.

The potential drug space is composed of over $10^{20}$ compounds, and candidates with suitable activity against specific proteins only narrows the search space to $10^3 − 10^5$ structures. Machine learning (ML) methods are actively used in this search-and-screen process. Candidate molecules generated by ML methods are passed to downstream verification via virtual high-throughput drug-protein binding techniques, drug synthesis, biological assay, and finally clinical trials [5, 16, 44].

Heterogeneous graphs provide a natural representation for small-molecule organic compounds, with nodes representing atoms in the molecular structure and edges representing bonds between the atoms [48]. This approach motivated the exploration of graph-generative models such as graph convolutional policy networks [49], variational autoencoders [26, 27, 40], and variants of deep reinforcement learning [45, 52] for the target-driven optimization of drug molecules. The drug-molecule design task is defined as generating a set of graphs $G_{opt}$ such that for each graph $g \in G_{opt}$, $f_{opt}(g) \geq \delta$, where $f_{opt}$ is a property optimization function and $\delta$ is a user-specified threshold. Most methods optimize target molecules for properties that can be derived from the molecular structure, such as the octanol-water partition coefficient (logP) [2] and quantitative estimate of drug likeness (QED) [7].

The focus of our work is two-fold to generate compounds for drug discovery, specifically for SARS-CoV-2. In the first similarity approach, a junction-tree based variational autoencoder (JT-VAE) [27] is trained on a database of known drug molecules that has pIC$_{50}$ activity [15]. The trained VAE is used to generate novel molecules optimized towards specific properties using Bayesian optimization (BO). The second approach extends a graph-based deep reinforcement learning (DQN) method [52] to generate molecules that are not constrained by their proximity to past anti-SARS compounds. For comparison, we use the same set of property scoring functions ($f_{opt}$) as the respective optimization and reward functions for benchmarking the JT-VAE and DQN approaches.

Contributions. 1) The goal of our study is to perform multi-objective optimization for generating molecular structures by considering critical bioactivity properties that are typically not considered in structure-activity relationship approaches. Specifically, we focus on pIC$_{50}$ [42], which captures the potency of the drug towards a protease target and cannot be accurately estimated from the chemical structure [3]. 2) In this context, we assembled a new protease dataset with molecules that are active against various enzymatic assays. This is considered to be one of the key properties while generating new drug molecules. These molecules are filtered from experimental pharmacology data in ChEMBL, BindingDB, and ToxCat [6]. 3) Finally, we validate all top-ranking molecules against a Drug Target Binding Affinity (DBTA) classifier [25] to assess potential anti-SARS-CoV-2 activity. Full details of our implementation and source code are available at https://github.com/exalearn/covid-drug-design.
Figure 1: Depiction of workflow developed for this work. The anti-SARS database is used to train a MPNN to predict pIC$_{50}$ and the JT-VAE model. The trained MPNN is used as the scoring function in both JT-VAE (top row) and DQN-based molecular generation (bottom). Candidate molecules are screened by pIC$_{50}$ (>8) and validated by a Drug Target Binding Affinity classifier.

2 BACKGROUND AND RELATED WORK

In response to the COVID-19 global pandemic, researchers have pushed to identify marketed drugs that can be repurposed for SARS-CoV-2 treatment [18, 20, 35, 44, 51]. Born et al. amended their Pac-Mann RL approach, originally intended to generate anticancer drugs, to generate molecules with affinity to viral target proteins and controlled toxicity [10]. Batra et al. applied an ML-based screening approach to find known compounds with binding affinity to either the isolated SARS-CoV-2 S-protein at its host receptor region or to the S-protein-human ACE2 interface complex [5]. Huang et al. developed a deep learning toolkit for drug repurposing, called DeepPurpose, with the goal of recommending candidates with high binding affinity to target amino acid sequences from known drugs [25]. These approaches screen large databases of known compounds, which have the potential to miss novel molecules with anti-SARS activity. Chenthamarakshan et al. developed the generative modeling framework CogMol to design drug candidates specific to a given target protein [13].

Here we discuss which properties should be considered during lead optimization [14]. LogP is a measure of lipophilicity, which provides an understanding of the behavior of a drug in the body. Compounds intended for oral administration should have a logP no greater than 5, according to Lipinski’s Rule of 5 [34]. Further analysis has shown that logP values between 1 and 3 may be more appropriate considering the effect of logP on absorption, distribution, metabolism, elimination, and toxicology (ADMET) properties [47]. Though oral bioavailability is an important factor, a sole focus on logP has the potential to screen out otherwise useful compounds [50]. QED has been proposed as a more holistic druglikeness metric [8], from 0 (low) to 1 (high). Druglikeness provides a general metric for ADMET properties, but does not indicate the activity or effectiveness of a drug towards a specific target. The half maximal inhibitory concentration (IC$_{50}$), on the other hand, provides a quantitative measure of the potency of a compound to inhibit a specific biological process. IC$_{50}$ is obtained by measurement, and no universal ab initio method of computing its value exists. A number of methods have been developed to approximate IC$_{50}$, many based on QSPR and recently some based on machine learning [1,3,9,38]. Similarity to known drugs is also an important factor in drug discovery [30,36], as is the ability to synthesize the molecule, which can be estimated by the synthetic accessibility (SA) score [from 1 (easy) to 10 (difficult)] [17].

3 METHOD

3.1 Surrogate Model for pIC$_{50}$ Prediction

We trained a message-passing neural network (MPNN) [19, 41] to predict pIC$_{50}$ (the inverse log of IC$_{50}$) for a given molecular structure. Following the formalism of Gilmer et al. [19], our network is composed of message, update and readout operations (eqns. 1-3) and our choices for these functions are based on networks developed by St. John et al. for polymer property prediction. [28]

The original state of each atom ($h_v$) and bond ($\alpha_{vw}$) in our molecule (G) is a 256-length vector with values defined by an embedding table based on the atomic number and bond type (e.g., single, double, aromatic). The states of these atoms are modified by eight successive “message” layers. Each message layer uses a two-layer multi-layer perceptron (MLP) with sigmoid activations to compute a message that uses the state of an atom ($h_v$), the state of the neighboring atom ($h_w$) and the bond which joins them ($\alpha_{vw}$). The atom and bond states are updated according to the following equations:
We use a junction-tree (JT) based variational autoencoder (VAE). We tested both "molecular fingerprints," where the states of each node are combined before using a multi-layer perceptron (MLP) to reduce to compute pIC<sub>50</sub>, and an "atomic contribution," where we combine the nodes after MLP to compute a per-node contribution to pIC<sub>50</sub>. We experimented with the use of five different functions to reduce the atomic state/contributions to a single value for each graph: summation, mean, maximum, softmax, and attention. The attention functions are created by learning an attention map by passing the node states through a MPNN. We tested all combinations of "molecular fingerprint" vs. "atomic contribution" and the five readout functions, for a total of 10 networks, training each on network the same 90% of our pIC<sub>50</sub> dataset and comparing its performance on the withheld 10% of the data. We used an MPNN that uses attention maps to reduce contributions from each atom to a single pIC<sub>50</sub> of a molecule in all subsequent experiments.

### 3.2 Junction Tree Variational Autoencoder

We use a junction-tree (JT) based variational autoencoder (VAE) method [27] for generating molecules with high proximity to anti-SARS drug molecules. This model generates novel molecular graphs by laying a tree-structured scaffold over substructures in the molecule, which are then combined into a valid molecule using a MPNN. The JT-VAE model allows for the incremental expansion of a molecular graph through the addition of subgraphs, or "vocabulary" of valid components (denoted <i>V</i><sub>D</sub>) derived from the training set (Figure 1). The subgraphs are used to encode a molecule into its vector representation and decode latent vectors into valid molecular graphs. The use of subgraphs, rather than building a molecule atom-by-atom, maintains chemical validity at each step, while also incorporating functional groups common to the training set. Chemical graphs generated from the vocabulary will be structurally similar to those in the training set, which is a benefit when attempting to design drugs with similar properties to those in the training set.

Given a molecular graph <i>G</i> = (<i>V</i>, <i>E</i>), JT-VAE coarsens <i>G</i> into a junction tree data structure <i>T</i> = (<i>V</i><sub>T</sub>, <i>E</i><sub>T</sub>) such that each node <i>v</i> ∈ <i>V</i><sub>T</sub> in <i>T</i> represents a subgraph of <i>G</i>. The subgraphs can be coarsened by using either automated JT construction algorithms from the graphical model literature [32] or a vocabulary based mapping approach that reduces to important building blocks of chemical structures.

Next, the JT-VAE method uses a message-passing network (as described in section 3.1) to encode <i>T</i> (Figure 1) into a vector representation <i>z</i><sub>T</sub>. We refer the reader to [27] for specific implementations of the MPNN-based encoders for the input graph and the junction tree. The decoder component of the VAE learns to generate the same junction tree structure from <i>Z</i><sub>T</sub> by maximizing the likelihood function <i>p(T|Z</i><sub>T</sub>).</n>

Once the VAE model is trained, the next phase involves a two-step process for generating a molecular graph structure that is optimized for a target property. The first step involves drawing a sample from the latent space learnt by the VAE and transforming the sampled vector representation into a corresponding junction tree structure. The second step pursues a Bayesian optimization (BO) strategy to map the junction tree to a molecular graph that maximizes the target properties. Using the JT-VAE trained on our SARS-related database, we perform Bayesian optimization (BO), using the method of Kusner et al. [31] to produce novel molecules with target properties described in section 3.4.

### 3.3 Deep Reinforcement Learning

We follow the Q-learning approach of Zhou et al. for our deep reinforcement learning approach [52]. We consider tasks in which an agent interacts with an environment <i>𝔼</i> represented as a molecular graph. The agent starts with a null graph. At each time-step the agent selects an action <i>a</i> from an action space <i>A</i> that includes addition of single atoms, changing the type of bonds or removing bonds from the graph. The agent also receives a reward <i>R</i> at each time step depending on a property scoring function. All the property scoring functions are described in section 3.4.

In this setting, we cast the molecule generation problem as a Markov Decision Process (MDP) [37] to learn a policy network <i>𝜋</i> that determines the best sequence of actions that start with an initial molecular graph and transform it to a larger graph with desirable properties in a step-by-step fashion (Figure 1). At each step, we enumerate all possible actions and then select those which produce valid molecules (e.g., respect valency rules). Next, we train a multi-layer perceptron (MLP) to predict the value of a certain action by passing the Morgan fingerprints [39] as input. The MLP approximates the value of an action computed using the Bellman equation, where the score of a state and the maximum score of the subsequent state is multiplied by a decay factor. As established with other Deep Q-Learning approaches, the addition of the value of the next state increases the value of moves which will lead to higher future rewards.

### 3.4 Scoring Functions

The scoring functions described in this section are used for both Bayesian optimization in the VAE based approach and reward computation for the deep reinforcement-learning based approach. Following Jin et al. [27], we first compute a score that penalizes logP by the SA score (recall from section 2 that higher SA values are discouraged) and number of cycles with more than 6 atoms (eqn. 5). Considering that QED is a more comprehensive heuristic than logP, we also use a similar scoring function composed of QED penalized by the SA score and number of long cycles (eqn. 6). We then examine the utility of a SARS-specific scoring function based on the pIC<sub>50</sub> predicted by our MPNN penalized by the SA score and number of long cycles. Finally, we examine a multi-objective scoring function that accommodates both pIC<sub>50</sub> and penalized QED.

\[
\text{logP}^P(m) = \text{logP}(m) - SA(m) - \text{cycle}(m)
\]
We tested multiple variants of the pIC\textsubscript{50}\textsuperscript{p} which only uses the contribution from a single atom, performs worse than the whole-molecule fingerprint. These trends give us confidence that the MPNNs are operating based on well-known physics, but we would require comparison of which atoms are contributing to the predicted pIC\textsubscript{50} to results from molecular binding simulations to determine if the networks are indeed interpretable.

4.3 Effect of the scoring function

Molecule generation setup using JT-VAE. We trained JT-VAE for 8,300 iterations on the full database, with the following hyperparameters: hidden state dimension of 450, latent code dimension of 56, and graph message passing depth of 3. Analysis of the trained JT-VAE is given in Figure S1. To optimize towards the specified scoring functions, we trained a sparse Gaussian process (SGP) to predict a score given the latent representation learned by JT-VAE and then perform 10 iterations of batched BO (sampling = 50) based on the expected improvement.

4 EXPERIMENTAL ANALYSIS

We conduct experiments to answer two questions: Q1: What are the best possible ways to generate molecules with targeted activity towards SARS-CoV-2? Q2: How do we evaluate the novelty of our generated molecules?

4.1 Dataset and Analysis

We assembled a dataset with molecules active against various enzymatic assays filtered from experimental pharmacology data logged in CheMBL, BindingDB, and ToxCat [6]. Details of this dataset are provided in section 5.1. Following the workflow in Figure 1, we first train a pIC\textsubscript{50} model, and subsequently use it to train the JT-VAE and DQN models. We begin with a discussion of the pIC\textsubscript{50} model given its key role in training the JT-VAE and DQN models.

4.2 Training pIC\textsubscript{50} surrogate model

We tested multiple variants of the pIC\textsubscript{50} prediction model. We found that MPNNs that limited the computation of the pIC\textsubscript{50} to contributions from only a few specific atoms in the molecule performed the best. As shown in Table S1, the models which use max and softmax functions for aggregating the atom-level representations to a molecule-level representation (equation 4) have higher $R^2$ scores than those which use summation or mean outputs.

The relative performance of different networks can be explained by the physical mechanism behind the performance of anti-viral drugs. The presence or absence of a specific pattern in the molecular structure (e.g., functional groups, substituents) controls whether the molecule will bind with a certain portion of a target virus protein. We hypothesize that the “maximum” function, in particular, models this “all-or-nothing” physics well. The other atoms in the molecule play a role in determining whether the molecule will stay affixed at the target site. The reduced but non-negligible effect of the side groups could explain why the “atomic”-contribution model, which only uses the contribution from a single atom, performs less well than the whole-molecule fingerprint. These trends give us confidence that the MPNNs are operating based on well-known physics, but we would require comparison of which atoms are contributing to the predicted pIC\textsubscript{50} to results from molecular binding simulations to determine if the networks are indeed interpretable.

4.3 Effect of the scoring function

Molecule generation setup using JT-VAE. Our DQN approach builds up molecules from a single atom to large molecules atom-by-atom and bond-by-bond. Each “episode” starts with a blank slate and the RL agent is allowed up to 40 steps. We update a model of the Q-function to predict the value of each move after each step in each episode and, as this model improves during training, we smoothly turn down the probability that we will choose a random move over the prediction of this model. The DQN finds tens of thousands of candidate molecules with high pIC\textsubscript{50}, as shown in Figure 2 and Table 1. Using the multi-objective reward function led to fewer molecules with pIC\textsubscript{50} > 8, but with increased druglikeness. Approximately two-thirds of the molecules with pIC\textsubscript{50} > 8 found with the multi-objective reward also show QED > 0.5. In contrast, only 5% of the molecules in the pIC\textsubscript{50} based search show QED > 0.5. The addition of QED reduces the total number of high-pIC\textsubscript{50} molecules found by 25%, but increases the number of high-pIC\textsubscript{50} molecules found by over 10 times. Therefore, we recommend incorporating synthesizability and/or druglikeness into RL-driven searches for drug-like molecules.

Comparing JT-VAE and DQN. Table 1 shows the three highest-scoring molecules generated by the two generative models with each scoring function. The DQN models always outperform JT-VAE in finding a molecule with a superior value of the scoring function being optimized. The performance disparity is particularly apparent when optimizing for logP: the maximum logP from DQN is 12.6 compared to only 4.1 for JT-VAE. We attribute the difference in optimization performance to JT-VAE implicitly sampling from a distribution of drug-like molecules and DQN having no such constraints.

The candidate molecules generated by JT-VAE have consistently better druglikeness and SA scores even when those values are not explicitly optimized for. When optimizing towards logP, the top-3 molecules generated by JT-VAE have moderate-to-high QEDs, while the top-3 from DQN are below 0.11. We attribute the exceptionally large disparity in optimal logP and associated QED values between the two methods to the fact that drugs typically have logP values between -0.4 and 5.6. The molecules from which JT-VAE was trained were all drug-like molecules, which makes it improbable to sample molecules with logP values far outside that range. Similarly, the
Table 1: Properties of the three highest-scoring molecules generated by each model using the specified scoring function.

| Scoring function | pIC<sub>50</sub> 1st | pIC<sub>50</sub> 2nd | pIC<sub>50</sub> 3rd | QED 1st | QED 2nd | QED 3rd | logP 1st | logP 2nd | logP 3rd | SA Score 1st | SA Score 2nd | SA Score 3rd |
|------------------|---------------------|---------------------|---------------------|---------|---------|---------|---------|---------|---------|-------------|-------------|-------------|
| logP<sup>P</sup> (JT-VAE) | 4.93     | 4.57     | 4.60     | 0.45    | 0.78    | 0.71    | 4.05    | 4.27    | 4.20    | 1.70        | 2.08        | 2.08        |
| logP<sup>P</sup> (DQN)   | 6.10     | 8.17     | 4.98     | 0.04    | 0.07    | 0.11    | 13.86   | 12.64   | 12.52   | 3.59        | 2.97        | 2.93        |
| QED<sup>P</sup> (JT-VAE) | 4.15     | 4.23     | 4.71     | 0.91    | 0.84    | 0.91    | 3.72    | 2.19    | 2.20    | 1.80        | 1.71        | 2.25        |
| QED<sup>P</sup> (DQN)   | 6.80     | 6.80     | 6.80     | 0.77    | 0.77    | 0.77    | 3.46    | 3.46    | 3.46    | 2.05        | 2.05        | 2.05        |
| pIC<sub>50</sub> (JT-VAE) | 10.22    | 10.07    | 10.07    | 0.15    | 0.12    | 0.12    | 4.86    | 3.80    | 3.06    | 4.52        | 4.94        | 4.98        |
| pIC<sub>50</sub> (DQN)   | 10.57    | 10.56    | 10.39    | 0.09    | 0.11    | 0.41    | 1.51    | 2.03    | −0.44   | 6.90        | 6.71        | 5.57        |
| pIC<sub>50</sub>+QED<sup>P</sup> (JT-VAE) | 8.58     | 5.98     | 8.18     | 0.83    | 0.87    | 0.70    | 4.02    | 3.37    | 3.82    | 1.93        | 1.74        | 1.50        |
| pIC<sub>50</sub>+QED<sup>P</sup> (DQN) | 10.27    | 10.27    | 10.27    | 0.80    | 0.80    | 0.80    | 3.02    | 3.02    | 3.02    | 2.90        | 2.90        | 2.90        |

4.4 Qualitative analysis from a drug discovery perspective

We observed an interesting structural trend in the molecules generated by JT-VAE when using pIC<sub>50</sub> as the scoring function. Figure 3 shows the structures of both the molecules with pIC<sub>50</sub> > 8 and the anti-HIV drug Indinavir, an antiretroviral protease inhibitor. A common backbone is shared between Indinavir and the top-6 predictions. The Tanimoto similarity scores [4] of these six generated molecules against Indinavir range from 0.65 to 0.91. Indinavir has been proposed as a drug to treat SARS-CoV-2 due to favorable docking to the coronavirus 3-chymotrypsin-like protease (3CL-protease), a promising drug target for combating coronavirus infections [11, 21, 22]. Notably, three of the generated molecules have a higher predicted pIC<sub>50</sub> than does Indinavir.

Experimental confirmation of Drug-Target-Interaction (DTI) is challenging and time-consuming [46]. In silico Drug-Target Binding Affinity (DTBA) methods offer an alternative to evaluate DTI [23]. We employ a ML-based DTBA model to validate the interaction of molecules generated by JT-VAE against a target SARS-CoV2-L protease [12]. We trained a DTBA binary classification model using extended connectivity fingerprint [39] encoding for the drug molecule and the target protease sequence encoding using a Convolutional Neural Network (CNN) as implemented in the DeepPurpose toolkit [25]. The default hyperparameters provided in the DeepPurpose toolkit were found to be sufficient. The DBTA model classified four of the top 11 molecules (including the top two in Figure 3) with probability > 0.5 to have interaction with SARS-CoV2-3CL protease.

5 CONCLUSIONS

We compared two graph generative models, JT-VAE and DQN, for the task of discovering potential small-molecule candidates with activity against SARS-CoV-2. DQN always outperformed JT-VAE in finding a molecule with a superior value of the scoring function being optimized. However, JT-VAE generated molecules that were more structurally similar to those in the database due to substructure representation, which produced a lower SA score and logP < 5. JT-VAE tended to produce what looked to be drug molecules, while DQN produced precursor-like candidates with optimized properties, which could be used as starting structures to add additional substituents aimed at the specific target.
REFERENCES

[1] P. Armutlu, M. E. Ozdemir, F. Uney-Yuksetekte, I. H. Kavakli, and M. Turkyay. 2008. Classification of drug molecules considering their IC50 values using mixed-integer linear programming-based hyper-boxes method. BMC Bioinformatics 9 (2008), 411. https://doi.org/10.1186/1471-2105-9-411

[2] A Avdeen, KJ Box, JEA Comer, C Hibbert, and KY Tam. 1998. pH-Metric logP. Determination of liposomal membrane-water partition coefficients of ionizable compounds. Journal of Pharmaceutical Research 15, 2 (1998), 209–215.

[3] A. Bag and P. K. Ghori. 2016. Development of Quantum Chemical Method to Calculate Half Maximal Inhibitory Concentration (IC50). Mol Inform 35, 5 (2016), 199–206. https://doi.org/10.1002/minf.201500004

[4] David Bajusz, Ana Rácz, and Károly Héberger. 2015. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? Journal of cheminformatics 7, 1 (2015), 20.

[5] Rohit Batra, Henry Chan, Ganesh Katham, Rampi Ramprasad, Mathew J. Cherukara, and Subramanian Sankaranarayanan. 2020. Screening of Therapeutic Agents for COVID-19 using Machine Learning and Ensemble Docking Simulations. arXiv preprints, Article arXiv:2004.03766 (April 2020), arXiv:2004.03766 pages. arXiv:q-bio.BM/2004.03766

[6] A Patricia Bento, Anna Gavotin, Anne Hersey, Louisa J Bellis, Jon Chambers, Mark Davies, Felix A Kruger, Yvonne Light, Lora Mak, Shaun McGlinchey, et al. 2014. The CHEMBL bioactivity database: an update. Nucleic Acids Research 42, D1 (2014), D1083–D1090.

[7] G Richard Bickerton, Gaia V Paolini, Jérémie Benard, Sorel Muresan, and Andrew L Hopkins. 2012. Quantifying the chemical beauty of drugs. Nature Chemistry 4, 2 (2012), 90.

[8] G. Richard Bickerton, Gaia V. Paolini, Jérémie Benard, Sorel Muresan, and Andrew L. Hopkins. 2012. Quantifying the chemical beauty of drugs. Nature chemistry 4, 2 (2012), 90–98. https://doi.org/10.1038/nchem.1243

[9] Esben Jannik Bjerrum. 2017. SMILES Enumeration as Data Augmentation Agents for COVID-19 using Machine Learning and Ensemble Docking Simulations. arXiv preprints, Article arXiv:2004.01215 (2020), arXiv:2004.01215 pages. arXiv:q-bio.BM/2004.01215

[10] Yu Wai Chen, Chin-Fang Bengu Yu, and Kowk-Yin Wong. 2020. Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CLpro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. F1000Research 9 (2020).

[11] Viji Chemthamarakshen, Fayel Das, Inkot Padhi, Hendrik Strobelt, Kar Wai Lim, Thomas Vallet, Alice Mac Kain, Lisa Miorin, Elena Moreno, Zun Zar Chi Naing, Jacqueline M. Fabius, Manon Eckhardt, Margaret Soucheray, Melanie J. Bennett, Merve Cakir, Michael Emerman, Verena V. Rezelj, Jeffrey Z. Guo, Theodore L. Roth, Ujjwal Rathore, Advait Subramanian, Julia Noack, Mark Davies, Felix A Krüger, Yvonne Light, Lora Mak, Shaun McGlinchey, et al. 2020. Coronavirus puts drug repurposing on the fast track. Nature Biotechnology 38, 4 (2020), 379–384.

[12] Tong He, Marten Heidegger, Fuqiang Ban, Artem Cherkesov, and Martin Estner. 2017. SimBoost: a read-across approach for predicting drug-target binding affinities using gradient boosting machines. Journal of Cheminformatics 9, 1 (2017), 1–14.

[13] Julian Horwood and Emmanuel Noutschi. 2020. Molecular Design in Synthetically Accessible Chemical Space via Deep Reinforcement Learning. arXiv preprint arXiv:2004.14308 (2020).

[14] Keixin Huang, TianFan Fu, Cao Xiao, Lucas Glass, and Jieming Sun. 2020. DeepPurpose: A Deep Learning Based Drug Repurposing Toolkit. arXiv preprint arXiv:2004.08919 (2020).

[15] Wengong Jin, Regina Barzilai, and Tommi Jaakkola. 2020. Multi-Objective Molecular Generation using Interpretable Substructures. arXiv preprints, Article arXiv:2002.03244 (Feb. 2020), arXiv:2002.03244 pages. arXiv:q-bio.BM/2002.03244

[16] Wengong Jin, Regina Barzilai, and Tommi S. Jaakkola. 2018. Junction Tree Variational Autoencoder for Molecular Graph Generation. CoRR abs/1802.03436 (2018). arXiv:1802.03436 http://arxiv.org/abs/1802.03436

[17] Peter-F. St. John, Caleb Phillips, Travis W. Kemper, A Nolan Wilson, Yanfei Guan, Michael F. Crowley, Mark R. Nimlos, and Ross E. Larsen. 2019. Message-passing neural networks for high-throughput polymer screening. Journal of Chemical Physics 150, 23 (June 2019), 234111. https://doi.org/10.1063/1.5099132

[18] Yash Khemchandani, Steve O’Hagan, Soumitra Samanta, Neil Swainston, Timothey J Roberts, Danielle Balagka, and Douglas B Kell. 2020. DeepGraphMol, a multi-objective, computational strategy for generating molecules with desirable properties: a graph convolution and reinforcement learning approach. bioRxiv (2020).

[19] Ashutosh Kumar and Kam Y. J. Zhang. 2018. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. Frontiers in chemistry 6 (2018), 315–315. https://doi.org/10.3389/fchem.2018.00315

[20] Mark Davies, Felix A Krüger, Yvonne Light, Lora Mak, Shaun McGlinchey, et al. 2020. Coronavirus puts drug repurposing on the fast track. Nature Biotechnology 38, 4 (2020), 379–384.

[21] Peter-C. St John, Caleb Phillips, Travis W. Kemper, A Nolan Wilson, Yanfei Guan, Michael F. Crowley, Mark R. Nimlos, and Ross E. Larsen. 2019. Message-passing neural networks for high-throughput polymer screening. Journal of Chemical Physics 150, 23 (June 2019), 234111. https://doi.org/10.1063/1.5099132

[22] Yash Khemchandani, Steve O’Hagan, Soumitra Samanta, Neil Swainston, Timothey J Roberts, Danielle Balagka, and Douglas B Kell. 2020. DeepGraphMol, a multi-objective, computational strategy for generating molecules with desirable properties: a graph convolution and reinforcement learning approach. bioRxiv (2020).

[23] Ashutosh Kumar and Kam Y. J. Zhang. 2018. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. Frontiers in chemistry 6 (2018), 315–315. https://doi.org/10.3389/fchem.2018.00315

[24] Mark Davies, Felix A Krüger, Yvonne Light, Lora Mak, Shaun McGlinchey, et al. 2020. Coronavirus puts drug repurposing on the fast track. Nature Biotechnology 38, 4 (2020), 379–384.

[25] Peter-F. St. John, Caleb Phillips, Travis W. Kemper, A Nolan Wilson, Yanfei Guan, Michael F. Crowley, Mark R. Nimlos, and Ross E. Larsen. 2019. Message-passing neural networks for high-throughput polymer screening. Journal of Chemical Physics 150, 23 (June 2019), 234111. https://doi.org/10.1063/1.5099132

[26] Yash Khemchandani, Steve O’Hagan, Soumitra Samanta, Neil Swainston, Timothey J Roberts, Danielle Balagka, and Douglas B Kell. 2020. DeepGraphMol, a multi-objective, computational strategy for generating molecules with desirable properties: a graph convolution and reinforcement learning approach. bioRxiv (2020).

[27] Ashutosh Kumar and Kam Y. J. Zhang. 2018. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. Frontiers in chemistry 6 (2018), 315–315. https://doi.org/10.3389/fchem.2018.00315

[28] Mark Davies, Felix A Krüger, Yvonne Light, Lora Mak, Shaun McGlinchey, et al. 2020. Coronavirus puts drug repurposing on the fast track. Nature Biotechnology 38, 4 (2020), 379–384.
molecular graphs. In AAAI Conference on Artificial Intelligence, Vol. 33. 1110–1117.

[41] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. IEEE Transactions on Neural Networks 20, 1 (2008), 61–80.

[42] JL Sebaugh. 2011. Guidelines for accurate EC50/IC50 estimation. Pharmaceutical statistics 10, 2 (2011), 128–134.

[43] Ganesh Sivaraman, Nicholas Jackson, Benjamin Sanchez-Lengeling, Alvaro Vasquez-Mayagoitia, Alan Aspuru-Guzik, Venkatram Vishwanath, and Juan de Pablo. 2020. A machine learning workflow for molecular analysis: application to melting points. Machine Learning: Science and Technology (2020).

[44] Nicholas Stahl, Göran Falkman, Alexander Karlsson, Gunnar Mathiason, and Jonas Bostrom. 2019. Deep reinforcement learning for multiparameter optimization in de novo drug design. Journal of Chemical Information and Modeling 59, 7 (2019), 3166–3176.

[45] Maha Thafar, Arwa Bin Raies, Somayah Albaradei, Magbubah Essack, and Vladimir R Bajic. 2019. Comparison Study of Computational Prediction Tools for Drug-Target Binding Affinities. Frontiers in Chemistry 7 (2019).

[46] Michael J Waring. 2010. Lipophilicity in drug discovery. Expert Opinion on Drug Discovery 5, 3 (2010), 235–248. https://doi.org/10.1517/17460441003605098 arXiv:https://doi.org/10.1517/17460441003605098 PMID: 22823020.

[47] David Weiming. 1990. SMILES. 3. DEPICT. Graphical depiction of chemical structures. Journal of Chemical Information and Computer Sciences 30, 3 (1990), 237–243.

[48] Jiaxuan You, Bowen Liu, Zhitao Ying, Vijay Pande, and Jure Leskovec. 2018. Graph convolutional policy network for goal-directed molecular graph generation. In Advances in Neural Information Processing Systems. 6410–6421.

[49] Ming-Qiang Zhang and Barrie Wilkinson. 2007. Drug discovery beyond the ‘rule-of-five’. Current Opinion in Biotechnology 18, 6 (2007), 478 – 488. https://doi.org/10.1016/j.copbio.2007.10.005 Chemical biotechnology / Pharmaceutical biotechnology.

[50] Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N. Zare, and Patrick Riley. 2019. Optimization of Molecules via Deep Reinforcement Learning. Scientific Reports 9, 1 (2019), 10752. https://doi.org/10.1038/s41598-019-47148-x
SUPPLEMENTARY MATERIAL

5.1 Dataset Description

Preparation. We prepared and assembled the protease datasets with molecules active against various protease in enzymatic assays filtered from experimentally pharmacology data such as CheMBL, BindingDB, and ToxCat [6]. The database was filtered out with the IC\textsubscript{50} activity standard types and their potency. Molecules with size larger than 1000 Dalton were removed due to the limitation of the representation of large molecules in cheminformatics. We also filtered out non-drug like molecules containing metals and polypeptides. The curated data was standardized using the logarithmic scale –log10 of a numeric value in nM for all compounds. We used the mean average for a molecule with more than one IC\textsubscript{50} value. The resulting dataset contains 6545 unique molecules accompanied by their SMILES strings and experimental IC\textsubscript{50} values.

Quantitative Characterization. We computed the metrics included in our scoring functions for each molecule in the database. The range of values can be seen in Figure S1. The highest pIC\textsubscript{50} is 10.89, while the lowest pIC\textsubscript{50} is 1.22. The most common pIC\textsubscript{50} is 4.0, which is shared by 320 structures, and the vast majority of structures (91.5%) have pIC\textsubscript{50} values greater than 4.0. LogP values range from -10.36 to 16.65 in a near Gaussian distribution with a mean of 3.70; 77.0% of all structures meet the requirement of Lipinski’s Rule of 5 that logP be no greater than 5. QED values range from 0.01 to 0.94, while SA values range from 1.35 to 8.24, with 96.3% being below 5.

We computed the Tanimoto similarity for all pairs of compounds to gain insight into the structural diversity of molecules in our database (Figure S2). The entries in the matrix were ordered in increasing pIC\textsubscript{50} values. The similarity is represented by the color bar, with yellow representing low similarity (0) and red high similarity (1). We observe that structures tend to become more similar to their neighbors as pIC\textsubscript{50} increases, indicating that compounds with high pIC\textsubscript{50} values tend to be structurally similar, supporting the consideration of molecular similarity during drug discovery.

Figure S1: We generated 1,000 molecules using the trained JT-VAE, of which 560 were unique. The figure shows the comparison of the QED, logP, SA score, and pIC\textsubscript{50} of compounds in the database and those generated by the JT-VAE. The JT-VAE reproduced the range of values present in the database, minus outliers. The similar values indicate that the JT-VAE is able to reproduce the wide range of structures present in the database. The pIC\textsubscript{50} values for generated molecules were estimated by our MPNN.
Figure S2: Heat map of the Tanimoto similarity between all compounds in the database. Entries are ordered in increasing pIC$_{50}$. The Tanimoto similarity is generally higher among structures with high pIC$_{50}$ (located towards bottom right of the matrix), indicating the importance of considering structural similarity in drug discovery.
Table S1: $R^2$ scores for MPNN models trained to predict the pIC$_{50}$ of drugs from their molecular structure. Each model was trained using a different readout function (columns) to combine atomic contributions to pIC$_{50}$ or to create a single molecular fingerprint. Bold indicates the model used in our experiments and underscore indicates the best-performing model.

| readout   | atomic | molecular |
|-----------|--------|-----------|
| attention | 0.57   | 0.56      |
| max       | 0.61   | 0.71      |
| mean      | 0.57   | 0.61      |
| softmax   | 0.61   | 0.70      |
| sum       | 0.51   | 0.54      |

Figure S3: Performance of DQN agent using the multi-objective (MO) and pIC$_{50}$ reward functions. The solid lines indicate the number of unique molecules with pIC$_{50}$ > 8 found after a certain number of steps. The dashed lines indicate unique molecules with pIC$_{50}$ > 8 and QED > 0.5.
Figure S4: Comparison of the properties of molecules optimized with different scoring functions by JT-VAE method. The pIC$_{50}$ value was estimated using our MPNN.