DeepPurpose: a Deep Learning Based Drug Repurposing Toolkit

Kexin Huang [1] **Tianfan Fu** [2] **Cao Xiao** [3] **Lucas M. Glass** [3] **Jimeng Sun** [4]

Abstract

We present DeepPurpose, a deep learning toolkit for simple and efficient drug repurposing. With a few lines of code, DeepPurpose generates drug candidates based on aggregating five pretrained state-of-the-art models while offering flexibility for users to train their own models with 15 drug/target encodings and 50+ novel architectures. We demonstrated DeepPurpose using case studies, including repurposing for COVID-19 where promising candidates under trials are ranked high in our results.

1. Introduction

Drug repurposing is about investigating existing drugs for new therapeutic purposes which can potentially speed up drug development (Pushpakom et al., 2019). With a large number of existing drugs, it is important to quickly and accurately identify promising candidates for new indications. Especially in facing COVID-19 pandemic today, drug repurposing become particularly relevant as a potentially much faster way to discover effective and safe drugs for treating COVID-19.

Deep learning has recently demonstrated its superior performance than classic methods to assist computational drug discovery (Öztürk et al., 2018; Lee et al., 2019), thanks to its expressive power in extracting, processing and extrapolating patterns in molecular data. For example, an existing study shows with deep learning we can quickly generate a drug candidate to act against untreatable strain of bacteria (Stokes et al., 2020). There are many deep learning models designed for drug target interaction prediction (Öztürk et al., 2018; Lee et al., 2019; Nguyen et al., 2019), which can be extended to drug repurposing. However, to acquire the results for drug repurposing and test their validity require substantial code development skills and biochemical domain knowledge, which are hard to acquire both. For example, the existing tools (Öztürk et al., 2018; Lee et al., 2019; Nguyen et al., 2019) are designed for computer scientists and hard to use by domain researchers with limited machine learning and coding experience. Furthermore, each individual open-sourced tool was designed and coded differently, which prevents easy integration of the diverse methods or model ensembles (Yang et al., 2019).

In this work, we aim to break the technical barrier by introducing DeepPurpose, a powerful and easy-to-use Python toolkit that can recommend the top binding drug candidates. With only one line of code that specifies the target amino acid sequences and the drug candidates, DeepPurpose loads and processes input molecular data, feeds them into multiple deep learning models pretrained on the large BindingDB datasets with different encoding schemes, aggregates the prediction results, and generates a descriptive ranked list of drug candidates with top binding scores. Then biomedical researchers can inspect this short list for further wet-lab validation. The ensemble of many models allow DeepPurpose to use different encoding algorithms (i.e., input representations), which broaden the search horizons and catches drugs that were missed by existing works due to the bias from only using one particular encoding model.

DeepPurpose is a python package with a Jupyter Notebook interface. It can run locally to ease the concern of processing proprietary drug data. It can also run on the cloud which alleviate the computational resource burdens faced by some users. DeepPurpose uses the most accessible input format: SMILES string for drugs and amino acid sequence for target. The output of DeepPurpose is a score that measures the binding activity of the input drug target pair.

In addition to the aforementioned design for helping domain scientists, DeepPurpose also offers a flexible framework for computer scientists to innovate new models for repurposing. Particularly, DeepPurpose includes multiple molecular encodings, varying from deep neural networks on classic computational biology and cheminformatics descriptors, to convolutional neural network (Krizhevsky et al., 2012), convolutional recurrent neural network (CNN-RNN) (Hochreiter & Schmidhuber, 1997), Transformer encoders (Shin et al., 2019) and Message-Passing Neural Network (Gilmer

[1]Health Data Science, Harvard T.H. Chan School of Public Health, Harvard University, Boston, USA [2]College of Computing, Georgia Institute of Technology, Atlanta, USA [3]Analytic Center of Excellence, IQVIA, Cambridge, USA [4]Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, USA.. Correspondence to: Jimeng Sun <jimeng@illinois.edu>.

Submitted to the ICML 2020 Workshop on Computational Biology (WCB).

et al., 2017; Yang et al., 2019). In total, by combining 7 encodings for proteins and 8 encodings for drugs, DeepPurpose offer 50+ models. To the best of our knowledge, majority of the encodings and models are novel for drug repurposing. Besides the functionality, DeepPurpose is simple to use. With less than 10 lines of codes, DeepPurpose downloads, prepossesses the training data, trains the model with multiple encodings, and evaluates various performance metrics on the test set. We also provide 10+ pretrained models trained on large benchmark datasets that are ready to use.

2. Result

Next, we showcase DeepPurpose's functionalities.

Drug Repurposing for SARS-CoV2 3CL Protease. The first case considers drug repurposing for COVID-19. Suppose a biomedical student want to identify which existing antiviral drugs can be repurposed to target SARS-CoV2 3CL Protease (3CLPro), which is related to virus replication. With only one line of code (A, Figure. 2), DeepPurpose aggregate multiple pretrained models and outputs top-ranked antiviral drug candidates. We present the results in Figure 2. Out of the 81 antiviral drugs in the library, DeepPurpose recommend 13 potentially active drugs that have Kd values within 500 (Beck et al., 2020) units. We conduct literature search for the 13 drugs and find that Ritonavir, Darunavir, Lopinavir are three of few drug candidates that are currently undergo clinical trials for SARS-CoV2-3CLPro (Harrison, 2020; Liping, 2020; Lu, 2020). Overall, 6 recommendations out of 13 have promising evidence based on literature or clinical trials, which confirms the potential of DeepPurpose for suggesting high-quality repurposing candidates.

Virtual Screening. Now, we consider virtual screening for a list of drug-target pairs. After inputting the drug target sequence information, DeepPurpose generates a ranked list of drug-target pairs based on the predicted binding affinity. To evaluate the predictive efficacy, we report the test set performance for the binding affinity prediction for each pretrained model, which is trained on BindingDB dataset with Kd values (Table 1. Figure. 2) using 7:1:2 train:validation:test split. We see all five models achieve high performance on all the metrics in the test set. Since in many cases the users may use drug-target pairs that are different from the training dataset, to test model's generalizability, we test on the DAVIS dataset (Davis et al., 2011), which has zero overlap with the pretraining BindingDB dataset. We then sample 1,000 unseen drug-target pairs from DAVIS and feed them into the pretrained models. We find the predicted Kd values of DeepPurpose is highly correlated with the true values with pearson correlation 0.7789, indicating the reliable predictions from DeepPurpose, even on new data (B, Figure. 2).

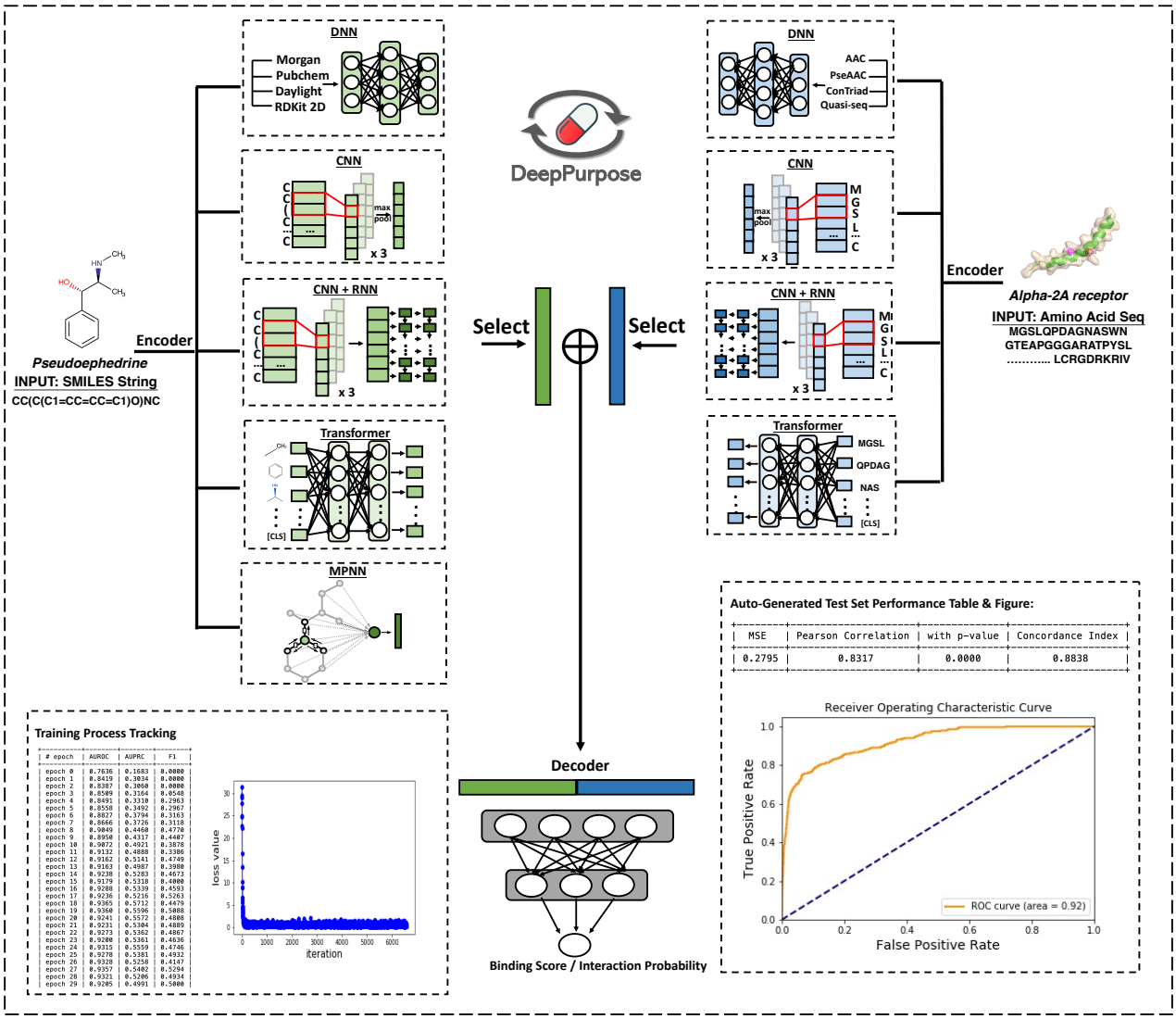
Repurposing using Customized Training Data. The third case illustrates the flexibility of DeepPurpose for more customized biomedical use case. We still consider the 3CL Protease for repurposing for COVID-19, but in this case, a biomedical scientist wants to train a deep learning model from past bioassay data such as high throughput screening (HTS) assay on SARS-CoV 3CL Protease (The Scripps Research Institute Molecular Screening Center, 2009), which conserves 96% of gene with SARS-CoV-2. This is potentially a better training dataset than the general BindingDB dataset for COVID-19. Specifically, DeepPurpose takes the customized training dataset as input. DeepPurpose then trains multiple deep learning models using this assay data to score drug candidates from the antiviral library or any proprietary data (C, Figure. 2). This resulting candidate list has support from literature. We find most of them are protein synthesis inhibitors, which is consistent with the general 3CLPro inhibition mechanism. Specifically, in addition to the Ritonavir, it surprisingly outputs Remdesivir with high confidence, which is a star candidate for COVID19 by blocking the RNA polymerase and has shown initial clinical effects (Grein et al., 2020). Since this bioassay has a binary label, DeepPurpose automatically switches from regression to binary classification loss.

A machine learning framework for DTI prediction. The fourth case is about developing new methods for drug-target interaction prediction. The researcher can benefit from DeepPurpose in the following way. First, DeepPurpose provides benchmark dataset loaders that obviates extensive searching and preprocessing. It supports various input data including drug-target pairs, one target, drug property prediction, and accommodates both continuous and binary label. Second, DeepPurpose supports various settings such as cold target, cold drug setup to test the robustness of the model. Third, DeepPurpose provides 50+ novel model architectures, which can be trained, and evaluated on user data using less than 10 lines of codes. We reproduce the state of the art methods DeepDTA (Öztürk et al., 2018) using our CNN+CNN encodings, and report the results on two benchmark dataset DAVIS (Davis et al., 2011) and KIBA (Tang et al., 2014) in (D, Figure. 2), along with other encodings. This confirms DeepPurpose's usability and generalizability.

3. Conclusion

Through these case studies, we demonstrate DeepPurpose 's functionality and easy usage for drug repurposing and screening for both biomedical scientists and machine learning researchers. We hope DeepPurpose can increase the accessibility of deep learning tools for drug repurposing and create valuable insights that can benefit the patients. We also call for domain and method researchers to contribute to this open-sourced project [1].

[1] https://github.com/kexinhuang12345/DeepPurpose



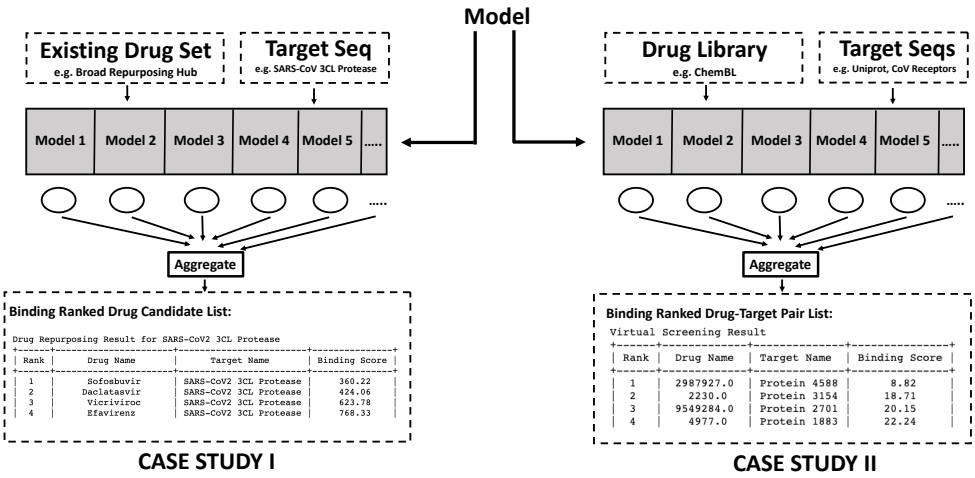


Figure 1. DeepPurpose method illustration. DeepPurpose takes the accessible drug's SMILES string and target's amino acid sequence and encode them through one of the selected 15+ encoder model. Then, the latent representations are concatenated and fed into a decoder to classify. The model is trained end-to-end and the training process, along with the test set performance is automatically-generated and reported. For drug repurposing & virtual screening, given a drug library and a new target of interest or new drug-target pairs of interest, DeepPurpose feeds them into five pretrained models and the predictions are aggregated and ranked to generate a drug-target list. This list can then be used for wet-lab validation. This entire process can done using one line of code in DeepPurpose.

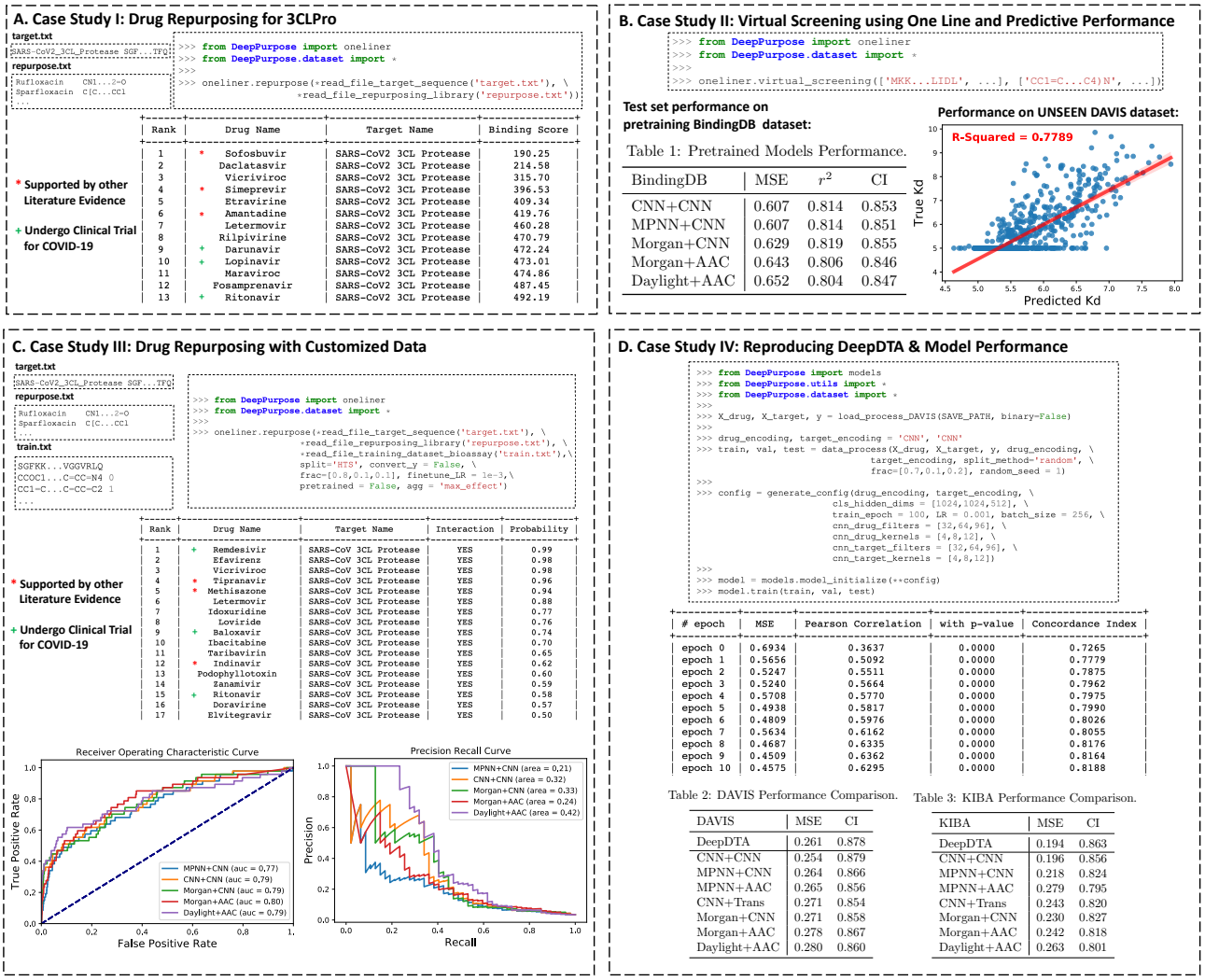


Table 1: Pretrained Models Performance.

BindingDB	MSE	r^2	CI
CNN+CNN	0.607	0.814	0.853
MPNN+CNN	0.607	0.814	0.851
Morgan+CNN	0.629	0.819	0.855
Morgan+AAC	0.643	0.806	0.846
Daylight+AAC	0.652	0.804	0.847

Table 2: DAVIS Performance Comparison.

DAVIS	MSE	CI
DeepDTA	0.261	0.878
CNN+CNN	0.254	0.879
MPNN+CNN	0.264	0.866
MPNN+AAC	0.265	0.856
CNN+Trans	0.271	0.854
Morgan+CNN	0.271	0.858
Morgan+AAC	0.278	0.867
Daylight+AAC	0.280	0.860

Table 3: KIBA Performance Comparison.

KIBA	MSE	CI
DeepDTA	0.194	0.863
CNN+CNN	0.196	0.856
MPNN+CNN	0.218	0.824
MPNN+AAC	0.279	0.795
CNN+Trans	0.243	0.820
Morgan+CNN	0.230	0.827
Morgan+AAC	0.242	0.818
Daylight+AAC	0.263	0.801

Figure 2. Case study illustration. **A.** Using one line of code, DeepPurpose generates a drug candidate list for SARS-CoV2-3CLPro target, aggregated from five pretrained model. **B.** With one line of code, DeepPurpose can obtain a list of drug-target pairs' predicted binding score. The figure showcases that DeepPurpose pretrained model is able to generalize to unseen drug-target pairs and the table shows that DeepPurpose has strong predictive performance on test set. **C.** Still using one line of code, DeepPurpose is able to take in customized training data, such as AID1706 bioassay data in this case and train five new models to generate a drug candidate list. The ROC-AUC and PR-AUC curves are automatically generated and they compare different training model's predictive performance. **D.** For method researchers, DeepPurpose provides a flexible framework to try on different encodings for drug and proteins, using less than 10 lines of code. It is able to reproduce the state-of-the-art DeepDTA's performance on two benchmark datasets. Table 2 & 3 report the predictive performance using various encodings and we see DeepPurpose has competitive predictive performance.

References

Beck, B. R., Shin, B., Choi, Y., Park, S., and Kang, K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (sars-cov-2) through a drug-target interaction deep learning model. *Computational and Structural Biotechnology Journal*, 2020.

Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. Comprehensive analysis of kinase inhibitor selectivity. *Nature biotechnology*, 29(11):1046, 2011.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 1263–1272. JMLR. org, 2017.

Grein, J., Ohmagari, N., Shin, D., Diaz, G., Asperges, E., Castagna, A., Feldt, T., Green, G., Green, M. L., Lescure, F.-X., Nicastri, E., Oda, R., Yo, K., Quiros-Roldan, E., Studemeister, A., Redinski, J., Ahmed, S., Bernett, J., Chelliah, D., Chen, D., Chihara, S., Cohen, S. H., Cunningham, J., D'Arminio Monforte, A., Ismail, S., Kato, H., Lapadula, G., L'Her, E., Maeno, T., Majumder, S., Massari, M., Mora-Rillo, M., Mutoh, Y., Nguyen, D., Verweij, E., Zoufaly, A., Osinusi, A. O., DeZure, A., Zhao, Y., Zhong, L., Chokkalingam, A., Elboudwarej, E., Telep, L., Timbs, L., Henne, I., Sellers, S., Cao, H., Tan, S. K., Winterbourne, L., Desai, P., Mera, R., Gaggar, A., Myers, R. P., Brainard, D. M., Childs, R., and Flanigan, T. Compassionate use of remdesivir for patients with severe covid-19. *New England Journal of Medicine*, 2020.

Harrison, C. Coronavirus puts drug repurposing on the fast track. *Nature biotechnology*, 2020.

Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.

Lee, I., Keum, J., and Nam, H. Deepconv-dti: Prediction of drug-target interactions via deep learning with convolution on protein sequences. *PLoS computational biology*, 15(6):e1007129, 2019.

Liping, D. A randomised, open, controlled trial for darunavir/cobicistat or lopinavir/ritonavir combined with thymosin a1 in the treatment of novel coronavirus pneumonia (covid-19). *Chinese Clinical Trial Registry*, 2020.

Lu, H. Efficacy and safety of darunavir and cobicistat for treatment of pneumonia caused by 2019-ncov (daconcov). *U.S. National Library of Medicine, ClinicalTrials.gov*, 2020.

Nguyen, T., Le, H., and Venkatesh, S. Graphdta: prediction of drug–target binding affinity using graph convolutional networks. *BioRxiv*, pp. 684662, 2019.

Öztürk, H., Özgür, A., and Ozkirimli, E. Deepdta: deep drug–target binding affinity prediction. *Bioinformatics*, 34(17):i821–i829, 2018.

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. Drug repurposing: progress, challenges and recommendations. *Nature reviews Drug discovery*, 18(1):41–58, 2019.

Shin, B., Park, S., Kang, K., and Ho, J. C. Self-attention based molecule representation for predicting drug-target interaction. In *Proceedings of the 4th Machine Learning for Healthcare Conference*, volume 106 of *Proceedings of Machine Learning Research*, pp. 230–248, Ann Arbor, Michigan, 09–10 Aug 2019. PMLR.

Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackerman, Z., et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702, 2020.

Tang, J., Szwajda, A., Shakyawar, S., Xu, T., Hintsanen, P., Wennerberg, K., and Aittokallio, T. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *Journal of Chemical Information and Modeling*, 54(3):735–743, 2014.

The Scripps Research Institute Molecular Screening Center. National center for biotechnology information. pubchem database. *AID 1706*, 2009.

Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., et al. Analyzing learned molecular representations for property prediction. *Journal of chemical information and modeling*, 59(8):3370–3388, 2019.