Supplementary Information

Universal approach to de novo drug design for target proteins using deep reinforcement learning

Yunjiang Zhang, Shuyuan Li, Miaojuan Xing, Qing Yuan, Hong He and Shaorui Sun**

Beijing Key Laboratory for Green Catalysis and Separation, The Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, PR China

∗. Corresponding author.Email: sunsr@bjut.edu.cn; hehong@bjut.edu.cn

Evaluation metrics of DTA models

This DTA models are trained by minimizing the loss function defined by the mean square error (MSE) between the outputs P of this network and depth values Y included:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - Y_i)^2$$

concordance index (CI) evaluates the ranking performance of the models that output continuous values.

$$CI = \frac{1}{Z}\sum_{\delta_x > \delta_y} h(b_x - b_y)$$

where b_x is the prediction value for the larger affinity δ_x, b_y is the prediction value for the smaller affinity δ_y, Z is a normalization constant, and h(m) is the step function.

$$h(m) = \begin{cases} 1; if\ m > 0 \\ 0.5; if\ m = 0 \\ 0; if\ m < 0 \end{cases}$$

DTA model for proteins and related inhibitors

Before the target protein was added to our model, proteins and inhibitors related to the target protein were deleted, and the DTA neural network was retrained. We used the optimal DTA model to predict the four proteins with known inhibitors, and the results are shown in Table 1. The highest known affinity of the inhibitor to the protein was (example Afatinib) 8.286. The molecules produced an affinity of up to 12. Next, molecules with high affinity can be further screened through molecular docking and virtual screening.

Table S1. DTA model affinity data for proteins and related inhibitors

Proteins	Inhibitors	affinity
BTK	Ibrutinib	6.180
	Calquence	5.156
	Zanubrutinib	7.201
PARP	Olaparib	7.004
	Rubraca	6.450
	Zejula	5.285
BRAF	Sorafenib	6.603
	Pazopanib	6.429
	Regorafenib	6.595
	Vemurafenib	6.232
	Dabrafenib	6.913
EGFR	Gefitinib	7.285
	Erlotinib	6.524
	Icotinib	6.959
	Afatinib	8.286
	Dacomitinib	7.922
	Osimertini	6.520

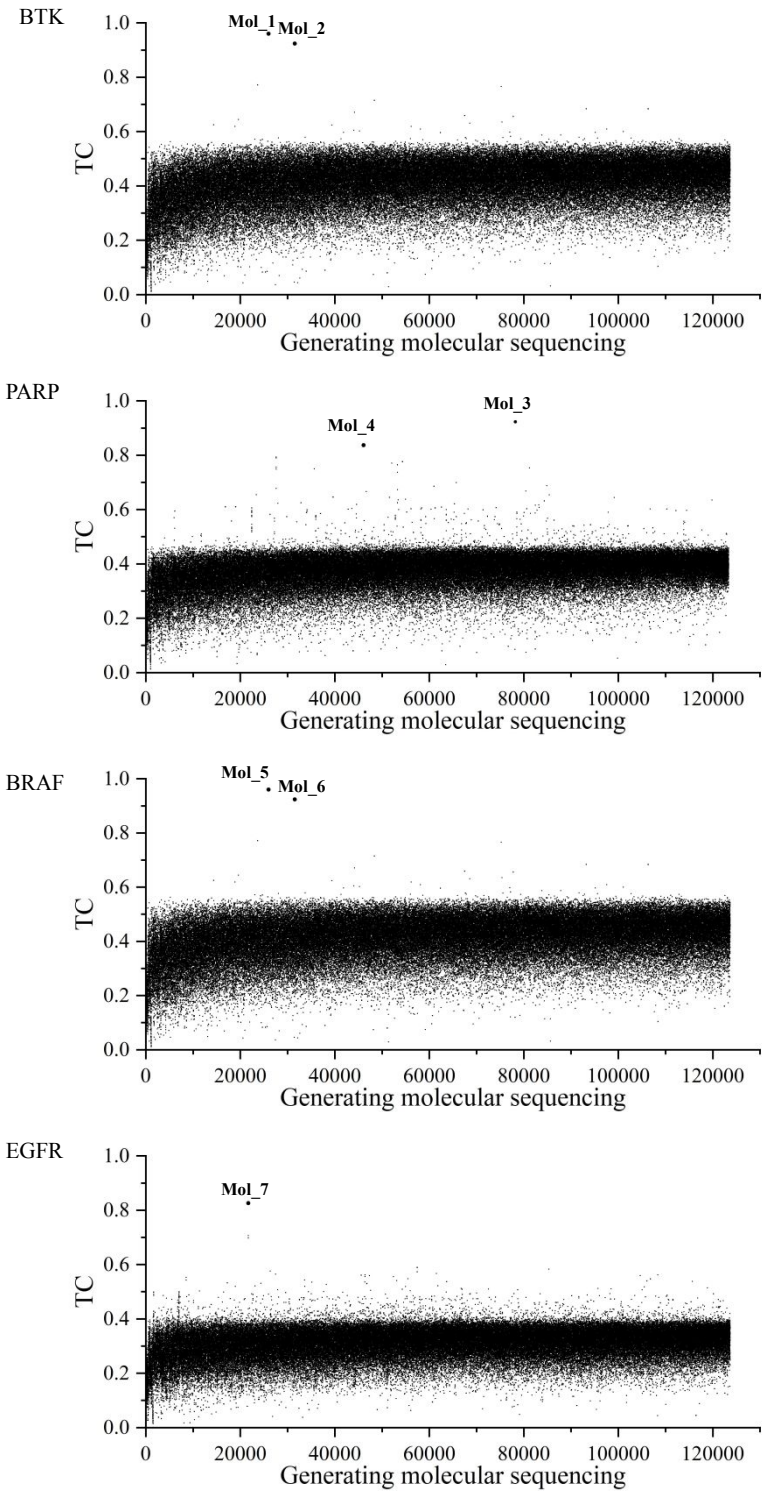

Figure S1. TC of the molecules generated in order of generation and known inhibitors among the four proteins.

Candidate inhibitors of BTK

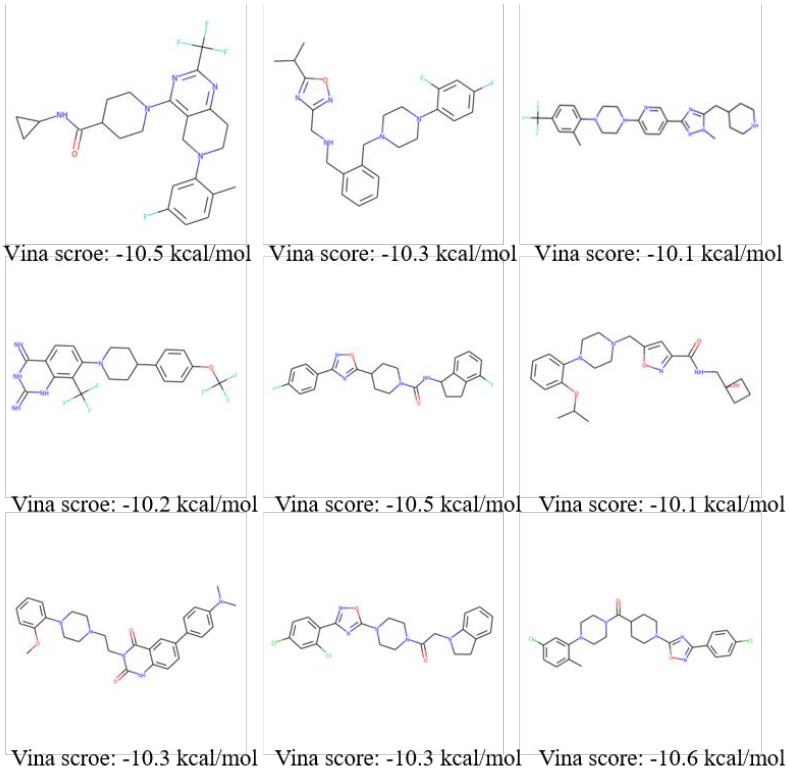

Figure S2. The 9 hit compounds obtained using Autodock Vina virtual screenings with ≤ -10 kcal·mol^{-1} of binding free energy for BTK.

Detailed computational methods for molecular assessment

Validity mean that the generated SMILES strings conform to the chemical structure.

Novelty is the part of the generated molecules that is not present in the trainset.

Filters is the proportion of valid molecules that can pass the custom medicinal chemistry filters and PAINS[1] filters.

Internal diversity is calculated by the following formula, $T(m_1, m_2)$ is the tanimoto similarity of the two molecules. For a set of molecules G, the interval of

internal diversity is [0,1], and the higher its value, the higher the diversity of molecules.

$$Internal\ deiversity(G) = 1 - \sqrt[p]{\frac{1}{|G|^2}\sum_{m_1,m_2 \in G} T(m_1,m_2)^p}$$

(1) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem*. **2010**, 53, 2719−2740.