

Supplementary Information for

Materials Discovery with Extreme Properties via

Reinforcement Learning-Guided Combinatorial Chemistry

Hyunseung Kim^{1†}, Haeyeon Choi^{2,3†}, Dongju Kang^{1†}, Won Bo Lee^{1*} and Jonggeol Na^{2,3*}

¹School of Chemical and Biological Engineering, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul, 08826, Republic of Korea.

²Department of Chemical Engineering and Materials Science, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Republic of Korea

³Graduate Program in System Health Science and Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Republic of Korea

† These authors contributed equally to this work

Correspondence should be addressed to W. Lee (email: wblee@snu.ac.kr) and J. Na (email: jgna@ehwa.ac.kr)

Supplementary Notes

1. Supplementary experiment on materials extrapolation to hit five extreme targets except for MW and DRD2. This note describes a supplementary experiment to the section in the manuscript: *Materials extrapolation to hit multiple extreme target properties*. Here, we conducted an experiment to generate molecules that hit five target properties: calculated octanol-water partition coefficient ($\log P$)¹, topological polar surface area (TPSA)², quantitative estimates of drug-likeness (QED)³, number of hydrogen bond acceptor (HBA), number of hydrogen bond donor (HBD). In this experiment, molecular weight (MW) and drug activity or dopamine receptor D2 (DRD2)⁴ which caused problems in the performance evaluation were excluded. Here, MOSES⁵ data sets were used for the training and test of GCT⁶. The training data contains 1,584,664 molecules and the test data contains 176,074 molecules. PubChem⁷ anticancer set, which covers a wider range than the MOSES data sets, was adopted to select extrapolated targets M1 to M10 (*Supplementary Figure 1a*). Targets M1 to M5 were selected from anticancer samples deviating from the TPSA- $\log P$ distribution of the training data and targets M6 to M10 were selected from samples deviating from the TPSA-QED distribution.

Before conducting the materials extrapolation, we evaluated the original target-hitting performance for GCT on the MOSES test data set whose distribution is similar to the trained data. To this end, the root mean squared error of each target property i ($RMSE_i$) for GCT was analysed to set the target bound of the target property i as $\pm RMSE_i$ (*Supplementary Table 5*).

The experimental results on materials extrapolation and interpolation are depicted in *Supplementary Figure 1*. Since the number of attempted molecular generation for materials interpolation and extrapolation is not the same, we rescaled the results based on 10,000 trials for easy comparison. The results are summarized in the left-hand side of *Supplementary Figure*

1b. In the experiment of materials interpolation, GCT discovered 5,365 five target-hitting molecules out of 10,000 trials for materials interpolation (target M0 in *Supplementary Figure 1b*). However, GCT failed to discover the five target-hitting molecules for all extrapolated targets, except for only six trials for the target M9—which is the closest target from the trained region. In contrast, AI-driven combinatorial chemistry worked well on materials extrapolation. It generated target-hitting molecules for all the extrapolated targets except the target M6. We believe that the cause lies in the way the fragments set was constructed. The x-makers shown in *Supplementary Figure 1c* denote the MWs of the reference molecules for the target M1 to M10. The other o-markers denote the MWs of molecular fragments constituting the fragments set. The problem is that the MW of the target M6 is 167 *Da*. It is significantly smaller than the other MWs of extrapolated targets. Note that the logP, TPSA, HBA, and HBD are scores that are calculated by summing each score of molecular fragments constituting the molecule. Hence, the descriptors have very large correlations with molecular size. It means that it is highly correlated with MW. Therefore, the combinations of properties are likely to be found only in materials with a certain level of MW. Since the MW of 167 *Da* is a level that can be composed of one or two molecular fragments in the fragments set, it might be difficult to complete target-hitting materials with a randomly selected initial fragment. In addition, the reference molecule of the target M6 was the only molecule that are not completely fragmented into molecular fragments with the used fragments set. Therefore, the fragment set should be designed with smaller fragments if the targets are smaller.

2. Size of the fragment set. We selected the size of the fragment set according to the result of the performance benchmark. To this end, we trained our model with various sizes of the fragment sets on a randomly sampled target (MW: 321.35, logP: 0.653, QED: 0.830, TPSA:

73.74, HBA: 4, HBD: 1). By subsampling BRICS⁸ 40k fragments in the order of their appearance in the MOSES database⁵, eight fragment sets with various sizes were constructed: ~40k fragments (41,153), ~20k fragments (21,477), ~12k fragments (12,768), ~8k fragments (8,835), ~6k fragments (6,434), ~2k fragments (2,547), ~0.5k fragments (565), ~0.2k fragments (166).

To select the efficient size of the fragment set, we compared the ratio of unique molecules after generating 3,000 molecules with the various fragment sets. As the results, we confirmed that the uniqueness ratio according to the size of the fragment sets as follows: ~40k fragments (98.87%), ~20k fragments (96.1%), ~12k fragments (93.73%), ~8k fragments (90.87%), ~6k fragments (88.47%), ~2k fragments (74.7%), ~0.5k fragments (38.83%), and ~0.2k fragments (16%). In addition, we compared the rewards achieved for each size of the fragment set (*Supplementary Figure 2*). Considering both results of the uniqueness ratio and the reward, we decided to use a fragment set with a size of around 2k.

More precisely, we used a fragment set of 2,207 BRICS fragments which appear more than 100 times in the curated ChEMBL training dataset⁴ for the experiments in the manuscript. In the case of the supplementary experiment (*Supplementary Note 1*), we used a fragment set of 2,547 BRICS fragments which appear more than 150 times in the MOSES training dataset⁵.

3. Benchmark on reinforcement learning algorithms and action masking. There are many kinds of algorithms in reinforcement learning. We benchmarked the performance of state-of-the-art RL algorithms on our materials extrapolation problem. In addition, we benchmarked performance according to whether or not action masking was applied. To this end, we introduced a scoring metric as follows:

$$Score = \sum_{y \in prop.} \frac{100}{\frac{|y_{gen.} - y_{trg.}|}{RMSE_y} + 1}$$

where $y_{gen.}$ and $y_{trg.}$ denote the value of generated molecule's property \mathcal{Y} and the value of the target property \mathcal{Y} , respectively. Here, *prop.* is a set of logP, TPSA, QED, HBA, HBD, and MW. We calculated the average score for 10,000 molecules generated by each algorithm. The results are shown in *Supplementary Figure 3*. As shown in the results, the proximal policy optimization (PPO)⁹ algorithm with action masking showed the best performance for our problem. Hence, we used the PPO algorithm with action masking to train our models.

4. Parameter setting for QuickVina2. Configuration for molecular docking simulation with QuickVina2¹⁰ is summarized in *Supplementary Table 6*. Here, the concept of exhaustiveness factor is similar to the concept of the population of the genetic algorithm. As the value of the exhaustiveness factor increases, the possibility of deriving an accurately optimized conformation increases. However, as the value of the exhaustiveness factor increases, the calculation time increases linearly. Hence, a strategy of the step increase for the exhaustiveness factor was adopted. The exhaustiveness factor is initially set to 1, and if two or more molecules with a docking score less than -12.5 are generated, the value is set to 2. Thereafter, if two or more of the generated molecules' docking scores are less than -14, the exhaustiveness factor is set to 4. After that, if two or more molecules with a docking score less than -15.5 are generated, the value is set to 8.

5. Active molecules of protein docking problem. As a result of generating 10,000 molecules with a low docking score for the 5-HT_{1B} receptor, 23 molecules matched the molecular

structure present in the ChEMBL¹¹ database. When investigating their pharmaceutical activity in the PubChem Bioassay database⁷, five out of the 23 molecules (21.7%) were the molecules that have been reported as pharmaceutically active molecules for specific targets. Except for the CHEMBL1726441 which is stated in the manuscript, the detailed information for the active molecules is described in *Supplementary Table 7*.

When analyzing the ratio of pharmaceutical activity for a set of 10,000 randomly sampled molecules from the ChEMBL¹¹ database, it was 9.6% (961 out of 10,000 molecules); we achieved 21.7% (five out of 23 molecules). Considering the above results, we believe that some of the generated molecules whose drug activity is identified may also have potential medicinal effects.

6. Determination of the maximum number of fragments to use in HIV-related targets.

Of the all cases, we limited the maximum number of fragments to less than six. This is because the total molecular weight of the generated molecule becomes too large to be used as a drug if a greater number of fragments is used. In general, a molecule with a molecular weight between 200 and 600 *Da* is used for drugs¹². Note that the average molecular weight of the fragments constituting the fragment set is about 110 *Da*. Therefore, if more than six fragments are used, it may be too large for a drug. The predicted pIC₅₀ values for three HIV-related targets are summarized in *Supplementary Table 8*.

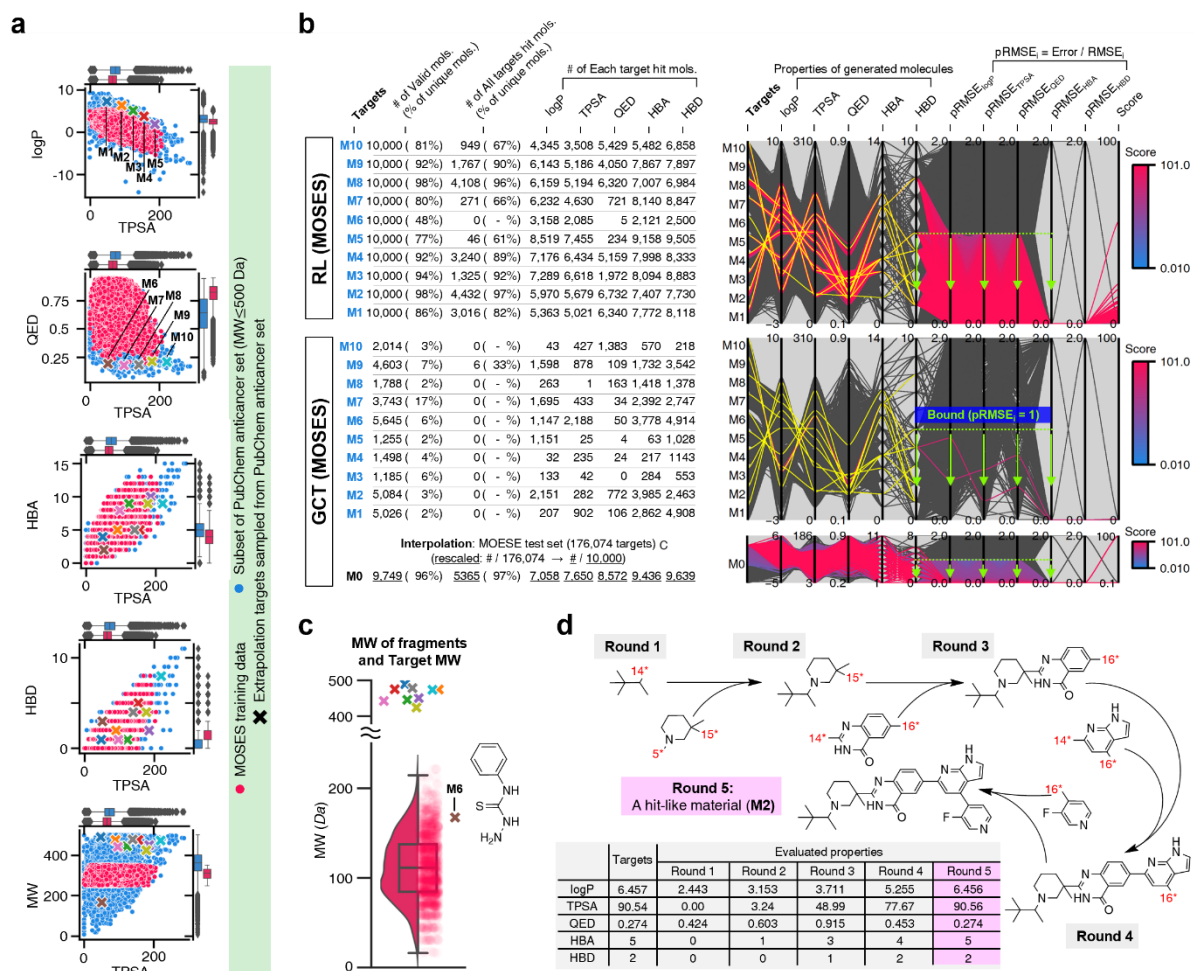
7. Model accuracy of prediction model. For the experiment on *Materials extrapolation to hit multiple extreme target properties*, 7 target properties were considered: logP, TPSA, QED, HBA, HBD, MW, and DRD2 activity. All properties except for DRD2 activity were calculated

using RDKit. Using RDKit, HBA, HBD, and MW are values that are directly calculated once the molecule is defined, while logP is computed based on the Wildman-Crippen method¹, and TPSA is determined using the summation of fragment contributions². RDKit calculates the Quantitative Estimation of Drug-likeness (QED) using a method that considers logP, TPSA, HBA, HBD, and MW. Lastly, DRD2 activity is predicted using the QSAR model introduced in the reference paper⁴ which is made using the support vector machine classification model. According to the reference paper, the QSAR model performance on active molecules of the test set shows that it misclassifies 10% of all cases as inactive.

For the *Application to the discovery of protein docking molecules*, QVina2 was used to predict the docking score of a molecule. In terms of model accuracy, it can depend on the specific target and parameter used in the program. Compared to Vina, it has Pearson's correlation coefficient(r) of 0.967, and the comparison of predicted binding energy can be seen in the reference paper¹⁰.

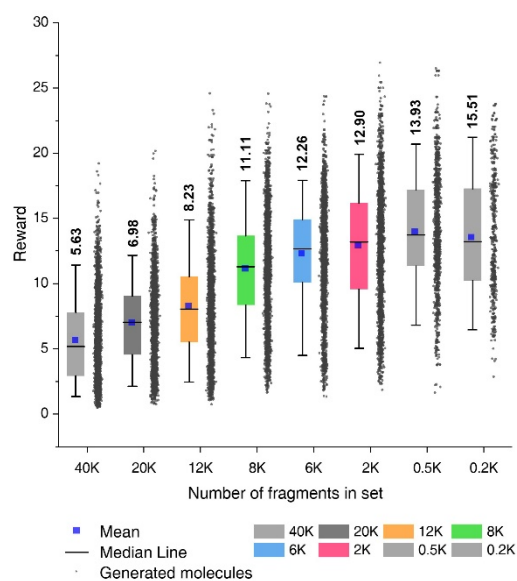
For the *Application to discovery of HIV inhibitors*, the QSAR model from the reference paper was used, which aims to predict the pIC50 value of each HIV-related target. The performance of the model can be found on the ESI† Table S10, which was analyzed from the reference paper¹³.

Based on the accuracy of the evaluation model, it is possible to narrow down the candidates more precisely within the chemical space that one aims to explore. Hence, the accuracy of evaluators affects how precisely narrow down the search space.

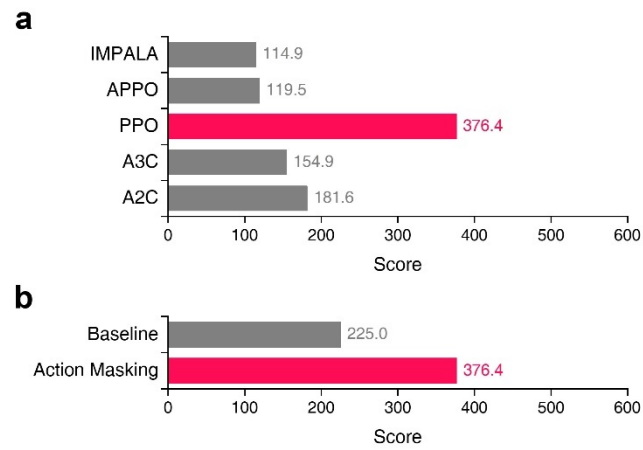


Supplementary Figures

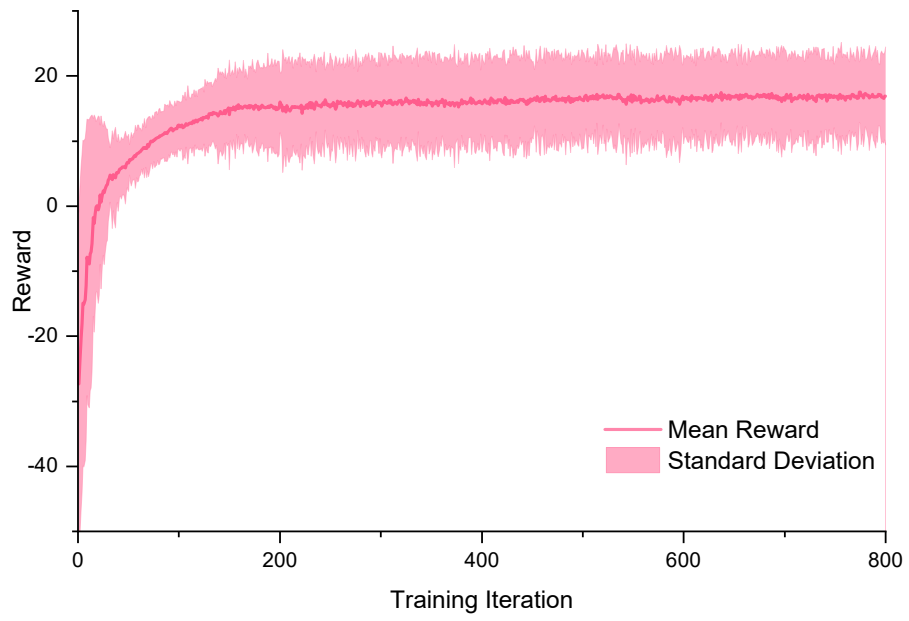
Supplementary Figure 1. Results for a supplementary experiment to generate the five target-hitting molecules. **a**, Distribution of MOSES⁵ training set and the extrapolation targets M1 to M10. **b**, Results comparison of AI-driven combinatorial chemistry and GCT. The left table shows the number of molecules that hit the target bounds. Each blue-red line in the right parallel coordinates plot indicates a molecule within the target bounds (green) of logP, TPSA, QED, HBA, and HBD at the same time. Blue-red lines were colored according to the log-scale score. The yellow line indicates each target. **c**, MW distribution of fragments in the fragments set and extrapolated targets. **d**, A molecular generation path example for target M2.



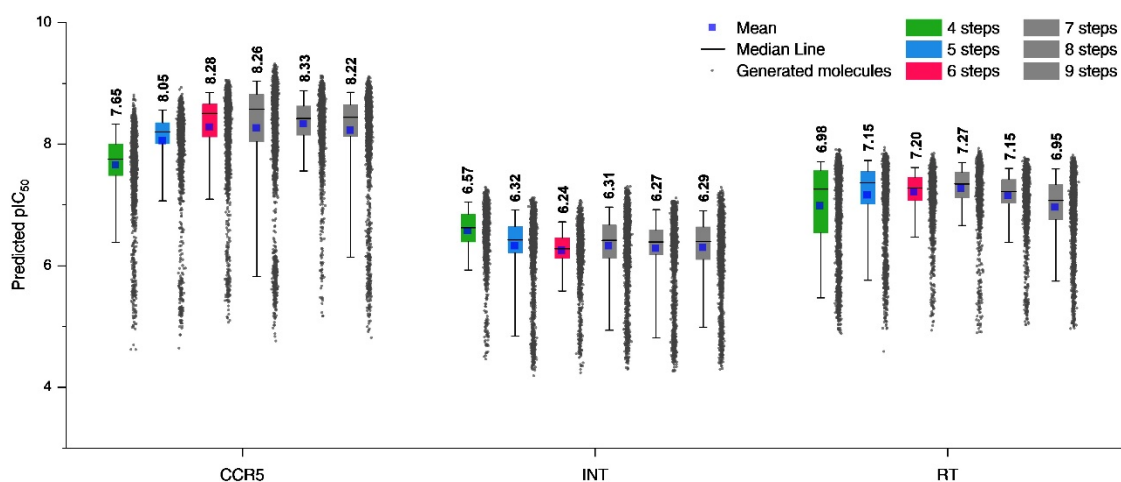
Supplementary Figure 2. The performance by the number of fragments constituting a fragment set. The fragment set consists of the fragments derived from the MOSES dataset. 3,000 molecules were generated with various sizes of fragment sets, and a box plot was drawn for unique molecules. The mean reward for each outcome is represented numerically, while the box is set with percentiles 25 and 75, and the whiskers extend to the 5th and 95th percentiles.



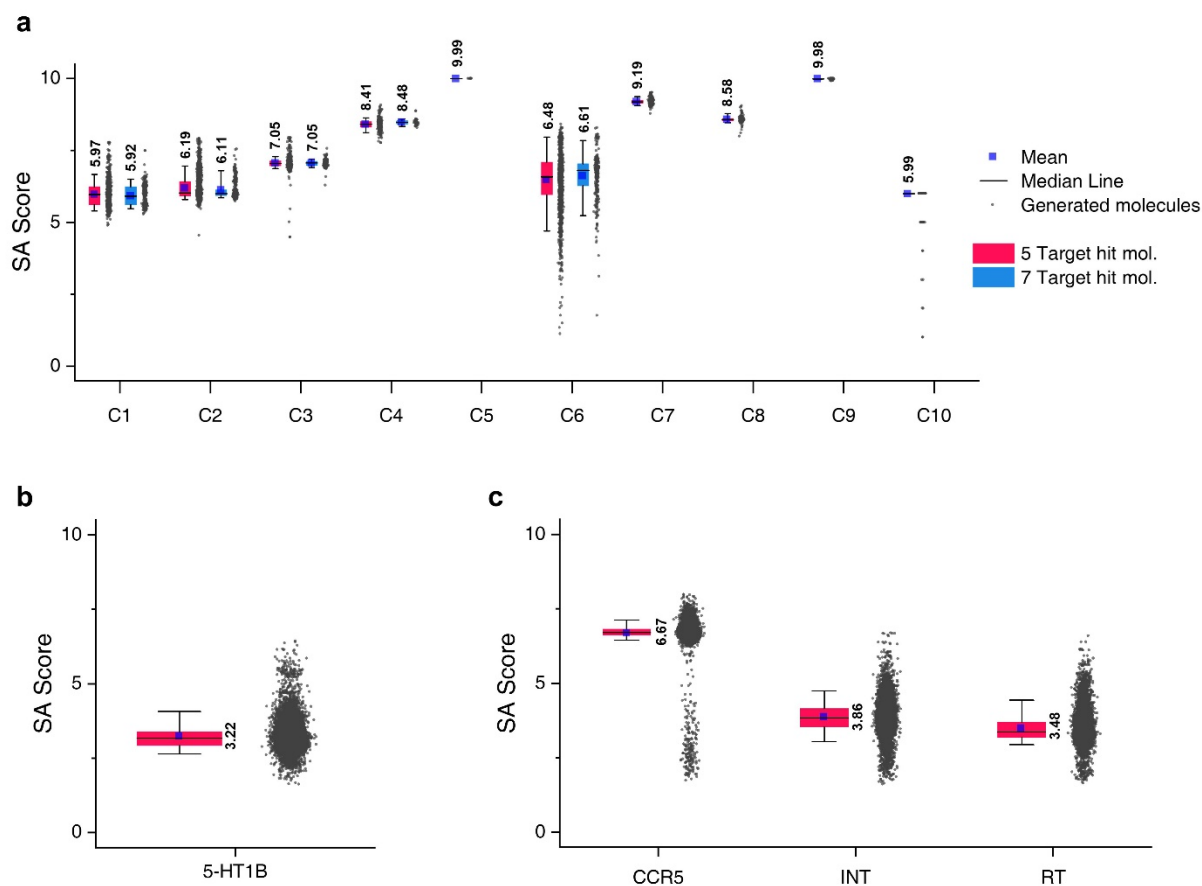
Supplementary Figure 3. Benchmark results for model performance. a, Results according to the type of reinforcement learning algorithms. **b,** Results according to the presence of action masking.



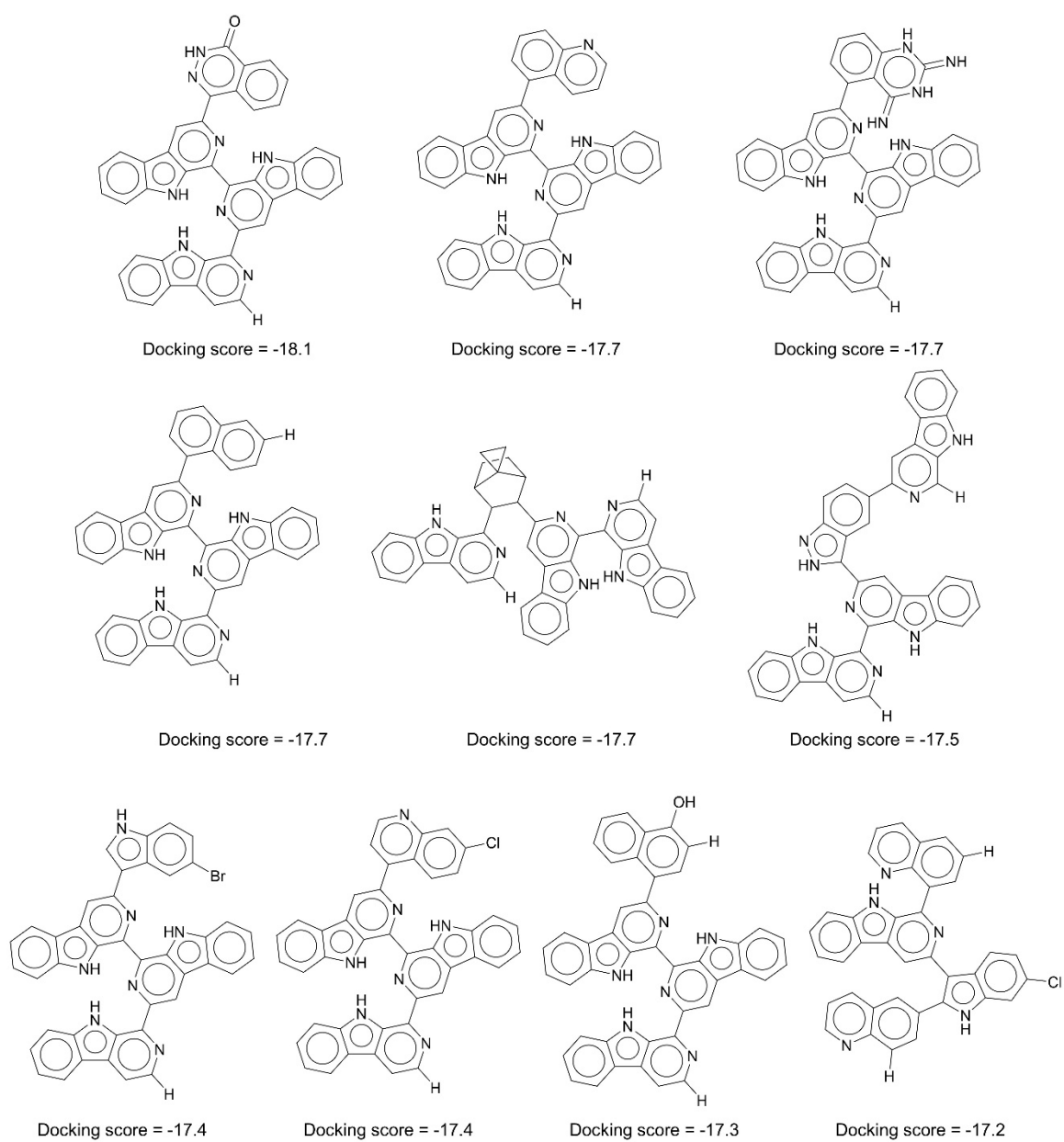
Supplementary Figure 4. The learning curve of the model. The graph illustrates changes in rewards according to training iterations.



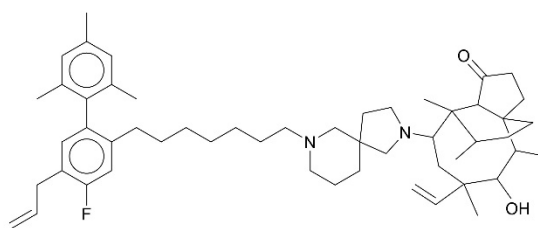
Supplementary Figure 5. The performance by the number of fragments used for HIV inhibitor discovery. For 10,000 generated molecules for each case, pIC₅₀ for each case was compared by setting the maximum fragments from 4 to 9. The mean of predicted pIC₅₀ value for each outcome is represented numerically, while the box is set with percentiles 25 and 75, and the whiskers extend to the 5th and 95th percentiles.



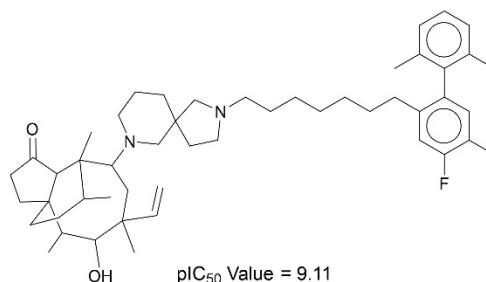
Supplementary Figure 6. Synthetic Accessibility (SA) score of each experiment. a, Materials extrapolation to hit multiple extreme target properties. **b,** Application to the discovery of protein docking molecules. **c,** Application to discovery of HIV inhibitors. The mean SA score for each outcome is represented numerically, while the box is set with percentiles 25 and 75, and the whiskers extend to the 5th and 95th percentiles.



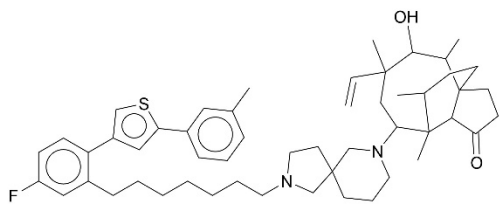
Supplementary Figure 7. Top 10 molecular structures with low docking score.



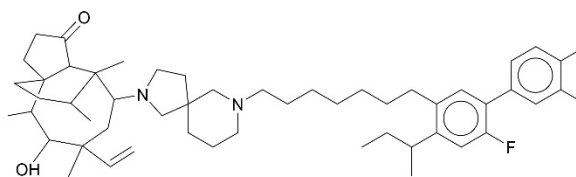
pIC₅₀ Value = 9.12



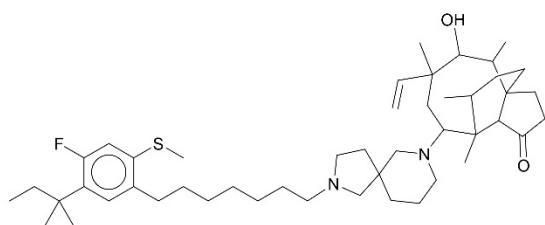
pIC₅₀ Value = 9.11



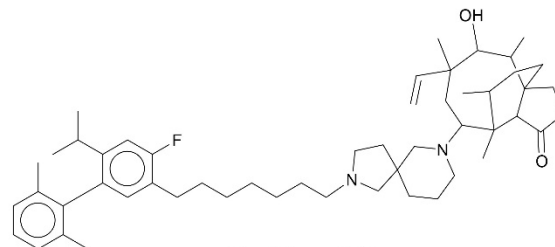
pIC₅₀ Value = 9.11



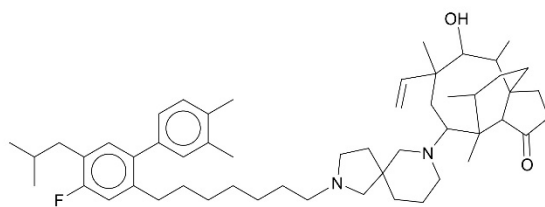
pIC₅₀ Value = 9.10



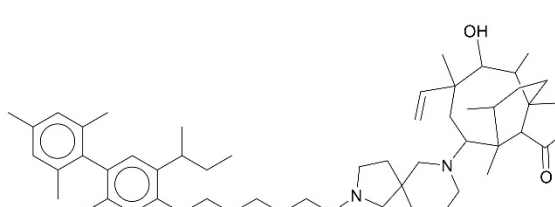
pIC₅₀ Value = 9.10



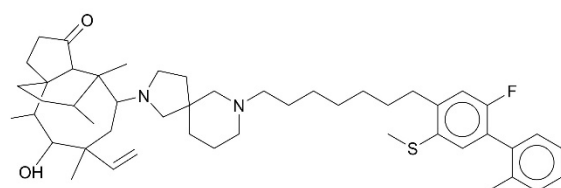
pIC₅₀ Value = 9.10



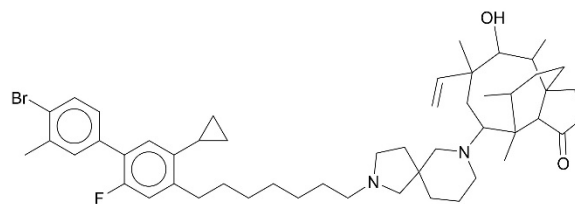
pIC₅₀ Value = 9.10



pIC₅₀ Value = 9.09

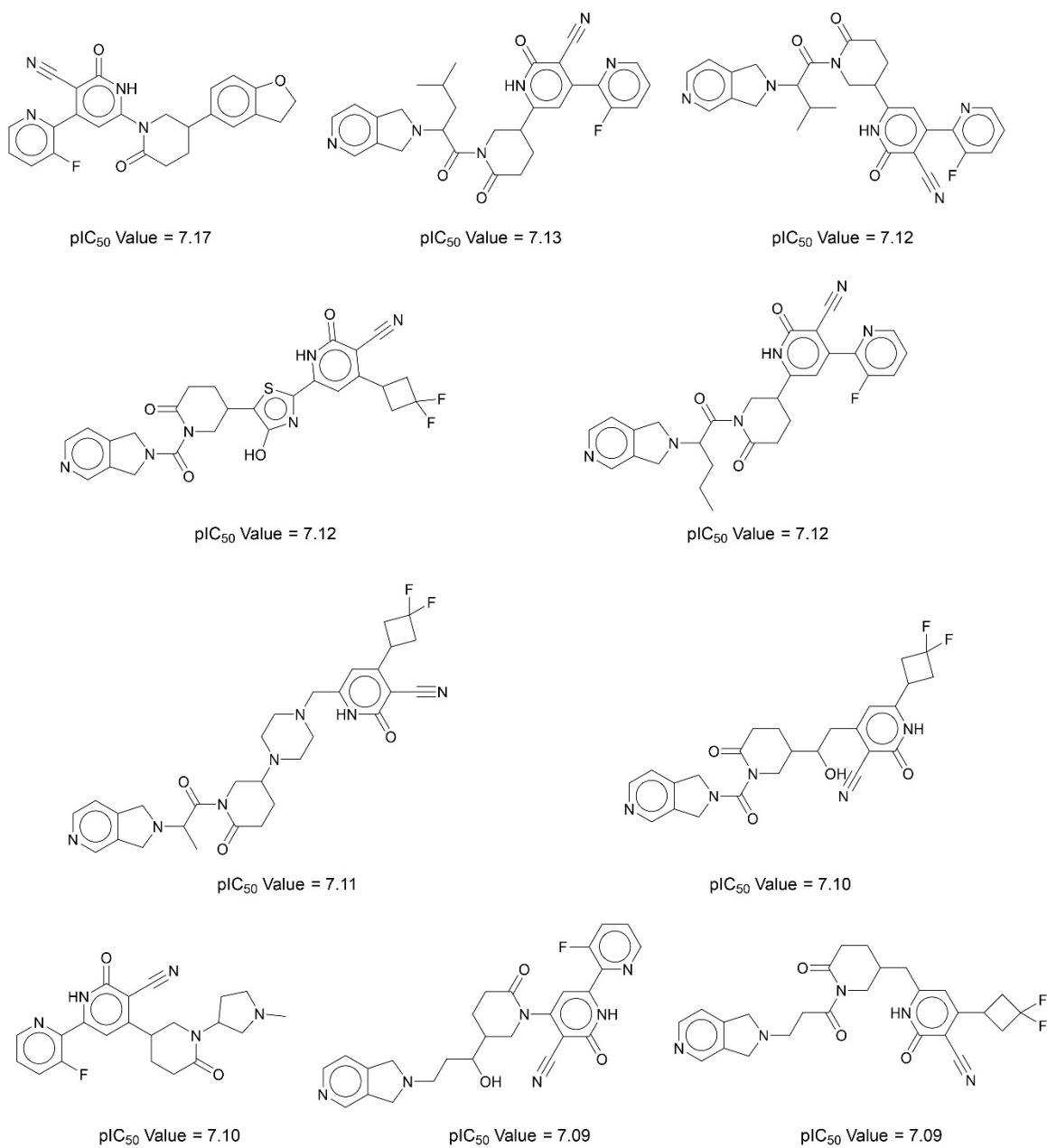


pIC₅₀ Value = 9.09

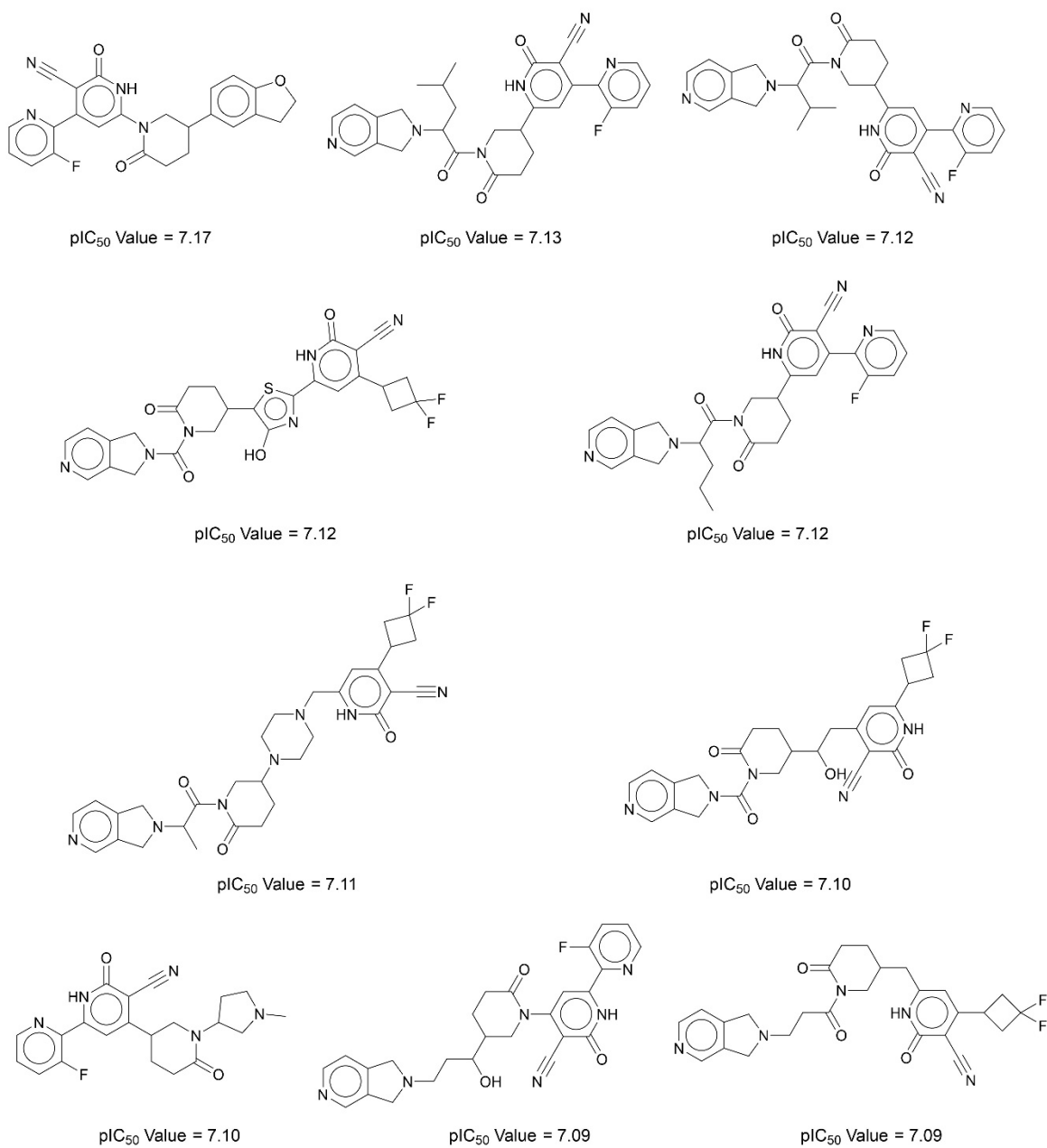


pIC₅₀ Value = 9.08

Supplementary Figure 8. Top 10 molecular structures with high pIC₅₀ value for CCR5.



Supplementary Figure 9. Top 10 molecular structures with high pIC₅₀ value for INT.



Supplementary Figure 10. Top 10 molecular structures with high pIC_{50} value for RT.

@@H](CCC(=O)O)C(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CCC(=O)O)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](CCC(=O)O)C(=O)N[C@@H](CCC(=O)O)C(=O)N[C@@H](C)C(=O)N[C@@H](CCC(=O)O)C(=O)N[C@@H](CC(=O)N)C(=O)NC(=O)[C@H](CCC(=O)O)NC(=O)[C@H](CO)NC(=O)[C@H](CO)NC(=O)[C@H]([C@@H](C)O)NC(=O)[C@H](CC(=O)O)NC(=O)[C@H](C(C)C)NC(=O)[C@H](C)NC(=O)[C@H](C)NC(=O)[C@H](CC(=O)O)NC(=O)[C@H](CO)NC(=O)C

Supplementary Table 2. Materials extrapolation for ChEMBL database with cRNN.

Simplified molecular-input line-entry system (SMILES)¹⁴⁻¹⁶ is a line-notated molecular representation. The SMILES^{a-n} are summarized in *Supplementary Table 4*.

Target	Generated SMILES	# of mols. generated	logP	TPSA	QED	HBA	HBD	MW	DRD2	
cRNN[4]										
C10			Target t	-18.09	1456.35	0.0251	49	49	3106.50	0.0100
	S	421		0.1128	0.00	0.3564	0	0	34.083	0.0000
	SO	1		0.3892	20.23	0.3005	1	2	50.082	0.0000
C9			Target t	11.22	1336.9	0.0073	40	51	3123.68	0.1638
	None									
C8			Target t	3.4797	926.85	0.0192	31	40	2086.96	0.0315
	None									
C7			Target t	6.4769	775.42	0.0290	27	27	2467.83	0.0005
	None									
C6			Target t	-1.77	526.91	0.0391	20	16	1421.75	0.2229
	SMILES ^a	1		2.2083	460.53	0.0122	17	15	1339.07	0.2304
	SMILES ^b	1		0.4676	477.46	0.0147	22	15	1334.75	0.1872
	SMILES ^c	1		0.7446	474.37	0.0393	21	15	1371.59	0.1136
	SMILES ^d	1		2.3708	421.51	0.0254	16	13	1279.72	0.2207
	SMILES ^e	1		0.3763	460.29	0.0186	18	14	1355.71	0.2734
	SMILES ^f	1		0.9282	428.74	0.0188	23	13	1356.50	0.2532
C5			Target t	-14.62	1447.9	0.0111	48	51	3324.74	0.0679
	None									
C4			Target t	-7.83	810.5	0.0154	28	27	1921.81	0.2455
	S	9,836		0.1128	0.00	0.3564	0	0	34.08	0.0000
C3			Target t	-12.15	483.41	0.0682	29	18	1026.38	0.0007
	SMILES ^g	1		-10.29	476.94	0.0198	27	18	832.76	0.0129
	SMILES ^h	1		-13.12	565.83	0.0338	29	21	924.00	0.0520
	SMILES ⁱ	1		-14.13	568.17	0.0114	31	22	916.87	0.0276
	SMILES ^j	1		-10.05	533.43	0.0107	30	20	956.94	0.0390
	O	3,018		-0.8247	31.5	0.3277	0	0	18.02	0.0007
C2			Target t	3.3153	464.92	0.0610	19	9	1269.63	0.0422
	None									
C1			Target t	13.6112	293.63	0.0129	15	7	1312.84	0.0151
	SMILES ^k	1		12.8784	272.75	0.0101	14	8	1170.10	0.0246
	SMILES ^l	1		11.8235	255.42	0.0095	13	7	1179.52	0.0423
	SMILES ^m	1		11.6254	247.89	0.0044	13	9	1137.46	0.0930
	SMILES ⁿ	1		13.2309	247.5	0.0049	12	9	1145.41	0.0251
	C	1,340		0.6361	0	0.3598	0	0	16.04	0.0001
	CS	119		0.546	0	0.3795	0	1	48.11	0.0000
	S	77		0.1128	0.00	0.3564	0	0	34.08	0.0000
	S=O	1		-0.3363	17.07	0.3724	2	0	48.07	0.0000
	SO	1		0.3892	20.23	0.3005	1	2	50.08	0.0000
	SS	246		0.7610	0	0.3025	0	2	66.15	0.0001

Supplementary Table 3. Materials extrapolation for ChEMBL database with GCT.

SMILES^{o-u} are summarized in *Supplementary Table 4*.

Target	Generated SMILES	# of mols. generated	logP	TPSA	QED	HBA	HBD	MW	DRD2	
GCT[6]										
C10			Target t	-18.09	1456.35	0.0251	49	49	3106.50	0.0100
	None									
C9			Target t	11.22	1336.9	0.0073	40	51	3123.68	0.1638
	None									
C8			Target t	3.4797	926.85	0.0192	31	40	2086.96	0.0315
	None									
C7			Target t	6.4769	775.42	0.0290	27	27	2467.83	0.0005
	None									
C6			Target t	-1.77	526.91	0.0391	20	16	1421.75	0.2229
	None									
C5			Target t	-14.62	1447.9	0.0111	48	51	3324.74	0.0679
	None									
C4			Target t	-7.83	810.5	0.0154	28	27	1921.81	0.2455
	None									
C3			Target t	-12.15	483.41	0.0682	29	18	1026.38	0.0007
	SMILES ^o	1		-6.6307	360.53	0.0284	21	16	992.43	0.0346
	SMILES ^p	1		-7.3916	409.53	0.0418	27	17	1011.58	0.0889
	SMILES ^q	1		-7.4463	411.61	0.0233	29	17	997.59	0.0671
C2			Target t	3.3153	464.92	0.0610	19	9	1269.63	0.0422
	None									
C1			Target t	13.6112	293.63	0.0129	15	7	1312.84	0.0151
	SMILES ^r	1		17.1370	71.06	0.1302	8	1	1007.17	0.0155
	SMILES ^s	1		10.2772	114.65	0.2391	8	2	724.12	0.0001
	SMILES ^t	1		8.3365	128.45	0.1319	11	4	765.92	0.0164
	SMILES ^u	1		12.7473	144.57	0.0307	8	2	949.66	0.0015

Supplementary Table 4. Discovered materials from cRNN and GCT. The molecules summarized in this table are referred to in *Supplementary Table 2* and *Supplementary Table 3*.

Index	SMILES
a	<chem>SC1n(C2=NCNS2(=O)=O)NC(C#N)C1Cc1ccc(CN(NCCCNC(N2CCC3N(C(C=CC4(N)CC(N(C(CCCCC)=O)C)CC(C(=O)N)N4C(Nc4ccc(Cl)c(NC(=NCC)N)c4)=O)C3CCCCNC(NC)=O)C2=N)=O)(=O)C)cc1NC(N)=O</chem>
b	<chem>S(CSC2=CCC(=NCCNC3C(O)OC(CN)=C(N)C3N)C(=O)NC3N(CCCCNC(=O)C4CCC(C)CN4CCCCOC(=O)NC(C)C(OC)=CC45CSC(C)(C)N5C(=O)N4)C(=O)NC(C)CC2(C(=O)N)C3C#N)(=O)C(CCCCCC)(N)NC(=N)N</chem>
c	<chem>Se1c(N=C(C2C(=O)N3C(C(=O)NC4C(=O)NC(CCCN)N(C(=O)C(CC(C)C)C(=O)NC(C(C)C)C(=O)N5CCCC5)N4C(=O)C4N(O)CCCC4)CCCCC2)C3)ccc(NC(=O)N2C3C(O)C(C2=O)C(O)C2(C(C(=O)O)=C(OC(C#N)C)CC2)NC3=N)c1NC(=N)NC</chem>
d	<chem>Se1c(N=C(C(C(N=c2ccc(Br)c[nH]2)=NS(C2C(=O)C(N3C(=O)N(CC(C)C)CCCNCNC(=O)NCCNC(=O)N3)C=CC3CN(C)CCC23)[O-])=O)N)c(=O)c2e1n(CCCC(CC(CC(O)=O)C)NC(N)=N)c(Cl)c2[N+](=O)[O-]</chem>
e	<chem>Se1c(CN=C(N)N)cc2c(c1)OC(C)C=CC(OC)C(C)OC(=O)C(C)=NC(=O)C1CC(C)NC(=O)C3(Cc4cc(ccc4)C[N+])45CCCC=CC[N+](C)CC[N+](CCN=C(N)N)CCCCN4CC(C(CN=NNC(N)=N)C1O)C(O)C3O)NC5NC(=O)N=C2N=C</chem>
f	<chem>S(CSSCCN(c2cccce2)C2=CC(=O)C(Nc3ccc(S(=O)(N)=O)cc3)=C(C)C(=O)NC(CBr)C(=O)NCCCCNCCCN(C(=O)C(S(=O)(O)=O)=CCC1C=CC[N+])1(C)C(=NS(=O)(N)=O)C2(O)C(O)(C)C(O)=O)C(O)(C)NC</chem>
g	<chem>NOCC(O)COC(OC)(C=C(C)OC(C)O)OC(COC(O)C(O)O)C(O)C(O)C(OC(O)C(CO)O)(O)OC(C(O)C(O)C2O)C(N)(C)N=C2N)C(O)O</chem>
h	<chem>NOCCC(NC(O)C(N)C(NCC(O)C(N)CO)O)OC(C(ON=c1c(=C2N(C(N)C(N)=O)C(CO)(N)C(O)C(CO)N2)C(N)C(N2COCC2N2C(N)=NN=C2N)=N1)CNC)N</chem>
i	<chem>NOC1SC2N(NN2)C1C(O)OC(C(C(N)C(O)O)O)OC(O)C(C(O)C(C(O)C(CO)O)OC(C(O)O)C(C(O)C(C(O)CO)O)ON(C(=O)C(N)C(N)=O)N)CC#N</chem>
j	<chem>NCC1SSC(SC2NC(N)=NC(CO)C2NC(O)=O)C(O)C1(O)C(C(=O)C(=C(C(O)O)C(C(C(C(C(=O)O)O)OC)O)O)OC(O)C(O)O)(N)OCCNC3=NN3N=O</chem>
k	<chem>SNS(CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC(=O)(=O)N=C(NSNe1ccc(S(=O)(=O)NCCOC(=O)N)c(C(O)=O)c1)c1cc(Cl)c(O)cc1CCC(=O)=O</chem>
l	<chem>SNS2(SSSSCC(C(=O)NC#N)=NC(SSC(C(=O)N)Ce3ccc(Cl)cc3)C(Cl)C(C)(C)C(=O)Ne3cc(C(C)C)c(S(=O)(O)=O)cc3CC=CC=C(C)CCCC=CC2=NC(=N[N+](C)=N)C)C</chem>
m	<chem>SNS(CSCSSCC(C(N=Cc1ccc(Cl)cc1Cl)=NOC(CCCC)(CCCCCCCCCCCCC)N)c1ccc(C(N)=O)cc1NC(NC(=O)C(CC)CC(=O)O)C=CC(=O)O)(=O)C</chem>
n	<chem>SNNS(=O)(OCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC[N+](CCCOc1cc(Cl)c(Cl)c(C(CNC(=O)C(C)C[N+](O)[O-])c2ccc(N)cc2)c1)(c1ccc(O)cc1)C[N+](=O)[O-])O</chem>
o	<chem>COC1OC(C(C)C(O)C(O)CONC(=SBr)C(OC2)C(NC(=O)CBr)C(O)C(O)C2OC2OC(O)C(Br)C(O)C(O)C(O)C(O)C(NO)O)OC12O</chem>
p	<chem>COC1OC(C(C)O)C(c2)C(NNC(=O)NN2Br)C(OC2OC(CC(O)C(O)C(O)C(O)C(O)COc2(Br)C(O)C2OOO)O)c2N(CC(NO)C(O)CNO)C2O1)C2O</chem>
q	<chem>COC1OC(C(Br)C(Br)CO)OC(C2OC(OCC(O)C(O)C(O)C(O)C(O)C(O)CO)OC2O)OC(NC2SSCC(ONO)O2)C(O)C(NO)O)OC1O</chem>
r	<chem>CC1CCCCCCC(=O)OCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC(=O)C(OSCc1)COc1C(Br)ccc(Br)cc1OS</chem>
s	<chem>CC1CCCCCCC(=O)OCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC(=O)C(O)C(=NC)C(O)COCCOc1</chem>
t	<chem>CC1CCCCCCC(=O)CCCCCCCC(=O)CCCC(=NC(=S)SSc2ccc(Br)cc2O)C(O)CCNC(O)COs1</chem>
u	<chem>O=C(O)CCCCCCCCCCCCCCCCCCCC(=O)OCCCCCCCCCCCC(=N)c1ccc(OCCBr)cc1OC(=O)N=C(Br)c1OCc1Br</chem>

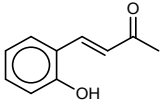
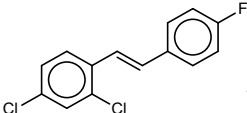
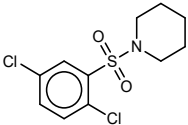
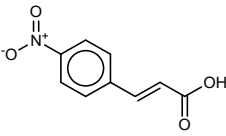
Supplementary Table 5. $RMSE_i$ for materials interpolation for the MOSES database with GCT.

Property i	$RMSE_i$ for GCT⁶
logP	0.214
TPSA	3.225
QED	0.037
HBA	0.180
HBD	0.106

Supplementary Table 6. Parameter settings for QuickVina2.

Parameter	Configuration
Exhaustiveness	Changes over time, from 1 to 8
Modes	10
CPU per subprocess	1
Box center	-26.602, 5.277, 17.898
Box size	22.5, 22.5, 22.5

Supplementary Table 7. The structure and role of molecules found to be active in the database among the molecules generated through our methodology.

Molecular Structure (ChEMBL ID)	Target Name Target Organism	Function
 <p>CHEMBL412355</p>	<p>Latent membrane protein 1</p> <hr/> <p>Human herpesvirus 4 (strain B95-8)</p>	<p>Regulates its expression and the expression of human genes¹⁷. Induces many changes associated with Epstein-Barr virus (EBV) infection and activation of primary B cells¹⁸. It is expressed in most EBV-related human cancers, such as various malignant EBV-related lymphoproliferative diseases¹⁹.</p>
 <p>CHEMBL2261013</p>	<p>Penicillium chrysogenum</p> <hr/> <p>Penicillium chrysogenum</p>	<p>Penicillium chrysogenum is a species of fungus in the genus penicillium and is the source of penicillin. Plus, the airborne asexual spores of penicillium chrysogenum are important human allergens. Vacuolar and alkaline serine proteases are associated with major allergenic proteins²⁰.</p>
 <p>CHEMBL1583499</p>	<p>TAR DNA-binding protein 43 (TDP-43)</p> <hr/> <p>Homo sapiens</p>	<p>TDP-43 is a transcription inhibitor that binds to TAR DNA integrated into chromosomes and inhibits HIV-1 transcription²¹ and thus may have efficacy in inhibiting HIV disease. Moreover, several neurodegenerative diseases, including amyotrophic lateral sclerosis, frontotemporal dementia, Alzheimer's disease, and limbic predominant age-related TDP-43 encephalopathy, share the common trait of the accumulation of TDP-43 aggregates in the central nervous system²².</p>
 <p>CHEMBL99068</p>	<p>Sporosarcina pasteurii</p> <hr/> <p>Unchecked</p>	<p>It has effects of inhibiting the urase of sporosarcina pasteurii CCM 2056²³. Urase is found in plant reservoirs and soils, which hydrolyze or decompose urea with water to form ammonium. The urase inhibitor can protect the process of ammonia volatilization²⁴.</p>

Supplementary Table 8. The pIC₅₀ comparison by the number of fragments of 3,000 generated molecules of each termination conditions. The case of showing the best performance for each condition is indicated in blue bold text.

Steps	CCR5		INT		RT	
	maximum	median	maximum	median	maximum	median
4 steps	8.81	7.75	7.29	6.62	7.91	7.26
5 steps	8.93	8.20	7.10	6.42	7.94	7.36
6 steps	9.05	8.51	7.06	6.28	7.84	7.28
7 steps	9.31	8.57	7.29	6.42	7.92	7.35
8 steps	9.12	8.42	7.10	6.39	7.77	7.22
9 steps	9.11	8.44	7.28	6.40	7.81	7.07

Supplementary Table 9. Comparison with other methods. We compared our model against a few other models. The best case in each section was marked in blue bold. Our proposed model has produced compounds with the highest maximum pIC₅₀ scores in two of the three domains, demonstrating that it outperforms others. The benchmark results of GCPN²⁴, JT-VAN²⁵, MSO²⁶, PGFS²⁷ were borrowed from ref. ²⁷.

Method	CCR5	INT	RT
GCPN ²⁵	8.20 (8.62)	6.45	7.42 (7.45)
JT-VAE ²⁶	8.15 (8.23)	7.25	7.58
MSO ²⁷	8.68 (8.77)	7.28	7.76
PGFS ¹³	9.05	7.5	7.89
RL (Ours)	9.11	7.17	8.01

Supplementary Table 10. Performance evaluation of trained QSAR models for predicting HIV-related targets using cross-validation¹³. R^2 refers to coefficient of determination, MAE refers to mean absolute error, and $Range$ refers to range of the values in the dataset.

Dataset	R^2		MAE		Range
	aggregated	average	aggregated	average	
CCR5	0.72	0.69 ± 0.03	0.51	0.54 ± 0.02	4.04-10.30
HIV-INT	0.69	0.65 ± 0.04	0.45	0.48 ± 0.03	4.00-8.15
HIV-RT	0.55	0.52 ± 0.05	0.51	0.53 ± 0.03	4.00-8.66

Supplementary References

- 1 Wildman, S. A. & Crippen, G. M. Prediction of physicochemical parameters by atomic contributions. *Journal of chemical information and computer sciences* **39**, 868-873 (1999).
- 2 Ertl, P., Rohde, B. & Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of medicinal chemistry* **43**, 3714-3717 (2000).
- 3 Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S. & Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nature chemistry* **4**, 90-98 (2012).
- 4 Kotsias, P.-C. *et al.* Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. *Nature Machine Intelligence* **2**, 254-265 (2020).
- 5 Polykovskiy, D. *et al.* Molecular sets (MOSES): a benchmarking platform for molecular generation models. *Frontiers in pharmacology* **11**, 565644 (2020).
- 6 Kim, H., Na, J. & Lee, W. B. Generative chemical transformer: neural machine learning of molecular geometric structures from chemical language via attention. *Journal of chemical information and modeling* **61**, 5804-5814 (2021).
- 7 Kim, S. *et al.* PubChem in 2021: new data content and improved web interfaces. *Nucleic acids research* **49**, D1388-D1395 (2021).
- 8 Degen, J., Wegscheid-Gerlach, C., Zaliani, A. & Rarey, M. On the Art of Compiling and Using 'Drug-Like' Chemical Fragment Spaces. *ChemMedChem: Chemistry Enabling Drug Discovery* **3**, 1503-1507 (2008).
- 9 Schulman, J., Wolski, F., Dhariwal, P., Radford, A. & Klimov, O. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- 10 Alhossary, A., Handoko, S. D., Mu, Y. & Kwok, C.-K. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics* **31**, 2214-2216 (2015).
- 11 Mendez, D. *et al.* ChEMBL: towards direct deposition of bioassay data. *Nucleic acids research* **47**, D930-D940 (2019).
- 12 Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews* **23**, 3-25 (1997).
- 13 Gottipati, S. K. *et al.* in *International conference on machine learning*. 3668-3679 (PMLR).
- 14 Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of chemical information and computer sciences* **28**, 31-36 (1988).
- 15 Weininger, D. SMILES. 3. DEPICT. Graphical depiction of chemical structures. *Journal of chemical information and computer sciences* **30**, 237-243 (1990).
- 16 Weininger, D., Weininger, A. & Weininger, J. L. SMILES. 2. Algorithm for generation of unique SMILES notation. *Journal of chemical information and computer sciences* **29**, 97-101 (1989).
- 17 Pratt, Z. L., Zhang, J. & Sugden, B. The latent membrane protein 1 (LMP1) oncogene of Epstein-Barr virus can simultaneously induce and inhibit apoptosis in B cells. *Journal of virology* **86**, 4380-4393 (2012).
- 18 Gupta, S. *et al.* Latent Membrane Protein 1 as a molecular adjuvant for single-cycle lentiviral vaccines. *Retrovirology* **8**, 1-12 (2011).
- 19 Ersing, I., Bernhardt, K. & Gewurz, B. E. NF- κ B and IRF7 pathway activation by

- Epstein-Barr virus latent membrane protein 1. *Viruses* **5**, 1587-1606 (2013).
- 20 Shen, H. D. *et al.* Molecular and immunological characterization of Pen ch 18, the vacuolar serine protease major allergen of *Penicillium chrysogenum*. *Allergy* **58**, 993-1002 (2003).
- 21 Ou, S., Wu, F., Harrich, D., García-Martínez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *Journal of virology* **69**, 3584-3596 (1995).
- 22 Jo, M. *et al.* The role of TDP-43 propagation in neurodegenerative diseases: integrating insights from clinical and experimental studies. *Experimental & molecular medicine* **52**, 1652-1662 (2020).
- 23 Macegoniuk, K. *et al.* Potent covalent inhibitors of bacterial urease identified by activity-reactivity profiling. *Bioorganic & Medicinal Chemistry Letters* **27**, 1346-1350 (2017).
- 24 Upadhyay, L. S. B. Urease inhibitors: A review. (2012).
- 25 You, J., Liu, B., Ying, Z., Pande, V. & Leskovec, J. Graph convolutional policy network for goal-directed molecular graph generation. *Advances in neural information processing systems* **31** (2018).
- 26 Jin, W., Barzilay, R. & Jaakkola, T. in *International conference on machine learning*. 2323-2332 (PMLR).
- 27 Winter, R. *et al.* Efficient multi-objective molecular optimization in a continuous latent space. *Chemical science* **10**, 8016-8024 (2019).