Supporting Information

RediscMol: Benchmarking molecular generation models in biological properties

Gaoqi Weng[1,†], Huifeng Zhao[1,†], Dou Nie[1], Haotian Zhang[1], Liwei Liu[2,*], Tingjun Hou[1,*], Yu Kang[1,*]

[1]Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China
[2]Advanced Computing and Storage Laboratory, Central Research Institute, 2012 Laboratories, Huawei Technologies Co., Ltd., Shenzhen 518129, Guangdong, China

[†]Equivalent authors

Corresponding authors:

Yu Kang	**E-mail**: yukang@zju.edu.cn
Tingjun Hou	**E-mail**: tingjunhou@zju.edu.cn
Liwei Liu	**E-mail:** liuliwei5@huawei.com

Contents of SI

Table S1. Data statistics of the datasets.

Name	Number	Generic murcko scaffolds	A	B	C	D	E	ratio A	ratio B	ratio C	ratio D	ratio E
EGFR	8747	1749	19	1794	2536	1795	2603	0.00217	0.20510	0.28993	0.20521	0.29759
CDK	12238	2305	8	1415	3320	3775	3720	0.00065	0.11562	0.27129	0.30847	0.30397
JakA	19731	3337	17	10638	4224	3079	1773	0.00086	0.53915	0.21408	0.15605	0.08986
VEGFR	11861	2561	31	2180	3637	3147	2866	0.00261	0.18380	0.30664	0.26532	0.24163
PDGFR	11006	2594	53	3078	2531	2624	2720	0.00482	0.27967	0.22997	0.23842	0.24714
AR	11931	2266	159	4581	3063	2467	1661	0.01333	0.38396	0.25673	0.20677	0.13922
5-HTR	27427	4351	119	10349	8721	5669	2569	0.00434	0.37733	0.31797	0.20669	0.09367
DR	19141	3335	100	6529	5606	4577	2329	0.00522	0.34110	0.29288	0.23912	0.12168
ChEMBL	1873888	215512	/	/	/	/	/	/	/	/	/	/
ChEMBL_no_Kinase	1000478	150881	/	/	/	/	/	/	/	/	/	/
ChEMBL_no_GPCR	925805	133623	/	/	/	/	/	/	/	/	/	/

Table S2. Pretraining results on the dataset without similar compounds of the GPCR datasets.

Pretrain	VAE	AAE	CharRNN	Reinvent	ORGAN
Validity	0.855	0.883	**0.959**	0.936	0.888
Uniqueness	0.997	0.997	0.996	**0.998**	0.991
IntDiv	0.885	0.887	0.885	0.886	0.887
logP	3.26 ± 1.87	3.30 ± 1.95	3.44 ± 1.90	3.38 ± 1.95	3.27 ± 1.86
SA	3.13 ± 0.79	3.22 ± 0.84	3.16 ± 0.79	3.22 ± 0.89	3.01 ± 0.76
QED	0.56 ± 0.20	0.54 ± 0.21	0.53 ± 0.21	0.54 ± 0.21	0.57 ± 0.20
MW	385.19 ± 101.69	387.12 ± 105.33	400.48 ± 104.66	393.95 ± 109.07	367.13 ± 99.05
Novelty	0.935	0.921	0.868	**0.936**	0.924
SNN/Gen_train	0.540	0.564	0.634	0.549	0.559

Table S3. The fine-tuning results on the CDK 10%-fine-tuning datasets with RDKit filtering.

CDK	CharRNN	AAE	VAE	Reinvent	ORGAN	GraphAF	RNNAttn			TransVAE		
							rand	high entropy	k high entropy	rand	high entropy	k high entropy
IntDiv	0.864 ± 0.006	0.868 ± 0.003	0.861 ± 0.005	0.862 ± 0.005	0.866 ± 0.003	0.901 ± 0.001	1	1	1	1	1	1
SNN/Gen_train	0.476 ± 0.034	0.493 ± 0.022	0.531 ± 0.025	0.503 ± 0.026	0.455 ± 0.018	0.242 ± 0.009	0.244 ± 0.008	0.290 ± 0.011	0.244 ± 0.008	0.233 ± 0.009	0.265 ± 0.011	0.230 ± 0.009
SNN/Gen_goal	0.452 ± 0.019	0.451 ± 0.014	0.477 ± 0.015	0.464 ± 0.015	0.422 ± 0.012	0.276 ± 0.005	0.277 ± 0.006	0.340 ± 0.007	0.278 ± 0.006	0.266 ± 0.006	0.303 ± 0.007	0.263 ± 0.006
IntDiv_ Rediscovery	0.803 ± 0.015	0.814 ± 0.005	0.817 ± 0.008	0.810 ± 0.011	0.794 ± 0.021	0.480 ± 0.054	/	/	/	/	/	/
SNN/Rediscovery_ train	0.779 ± 0.016	0.775 ± 0.015	0.779 ± 0.014	0.774 ± 0.009	0.773 ± 0.021	0.524 ± 0.078	0.472 ± 0.000	/	0.340 ± 0.000	/	/	/
Rediscovery	0.005 ± 0.001	0.004 ± 0.001	0.005 ± 0.001	0.005 ± 0.001	0.002 ± 0.001	0	0	0	0	0	0	0
Rediscovery_ number	138.800 ± 21.940	111.400 ± 25.562	140.000 ± 26.226	140.300 ± 29.564	69.300 ± 19.920	1.500 ± 1.025	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Rediscovery_A_ number	0.500 ± 0.671	0.700 ± 0.458	0.900 ± 0.700	1.000 ± 0.447	0.300 ± 0.458	0	0	0	0	0	0	0
Rediscovery_B_ number	18.500 ± 5.239	15.000 ± 4.940	16.000 ± 4.243	20.600 ± 6.468	7.200 ± 3.219	0.200 ± 0.400	0	0	0	0	0	0
Rediscovery_C_ number	57.500 ± 13.366	43.400 ± 12.035	56.100 ± 15.261	55.100 ± 11.379	27.400 ± 10.575	0.300 ± 0.900	0	0	0	0	0	0
Rediscovery_D_ number	42.000 ± 12.337	31.300 ± 10.845	42.600 ± 11.395	40.300 ± 11.055	22.900 ± 5.787	0.100 ± 0.300	0	0	0	0	0	0

Rediscovery_E_number	20.300 ± 2.900	21.000 ± 4.000	24.400 ± 1.800	23.300 ± 5.478	11.500 ± 4.387	0.900 ± 0.700	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Rediscovery_0.7	0.239 ± 0.062	0.252 ± 0.051	0.257 ± 0.034	0.252 ± 0.046	0.300 ± 0.051	0.700 ± 0.407	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Rediscovery_0.7_number	32.800 ± 7.909	27.700 ± 6.543	35.500 ± 5.334	35.400 ± 9.604	21.000 ± 7.197	1.200 ± 0.748	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Rediscovery_0.7_A_number	0.100 ± 0.300	0.200 ± 0.400	0.100 ± 0.300	0.200 ± 0.400	0.100 ± 0.300	0	0	0	0	0	0	0
Rediscovery_0.7_B_number	3.200 ± 1.661	3.100 ± 1.972	2.500 ± 1.628	4.600 ± 1.960	1.700 ± 1.616	0.200 ± 0.400	0	0	0	0	0	0
Rediscovery_0.7_C_number	13.900 ± 4.482	11.400 ± 4.152	15.100 ± 4.323	12.800 ± 3.763	7.600 ± 4.341	0.100 ± 0.300	0	0	0	0	0	0
Rediscovery_0.7_D_number	9.100 ± 2.914	7.100 ± 1.758	9.500 ± 2.540	9.800 ± 3.458	7.100 ± 2.508	0	0	0	0	0	0	0
Rediscovery_0.7_E_number	6.500 ± 1.746	5.900 ± 2.879	8.300 ± 2.052	8.000 ± 2.933	4.500 ± 2.156	0.900 ± 0.700	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Sim_0.7	0.088 ± 0.013	0.084 ± 0.010	0.104 ± 0.012	0.096 ± 0.014	0.051 ± 0.013	0.001 ± 0.000	0	0	0	0	0	0
Sim_0.7_number	2632.100 ± 383.147	2530.300 ± 303.789	3126.200 ± 363.647	2894.700 ± 432.065	1533.500 ± 375.330	18.900 ± 14.591	0.900 ± 0.943	1.100 ± 0.831	0.300 ± 0.458	0.700 ± 0.781	0.800 ± 0.600	1.500 ± 0.806
Sim_0.7_A_number	14.400 ± 13.147	13.300 ± 13.682	23.700 ± 19.282	22.400 ± 17.351	14.800 ± 17.325	0	0	0	0	0	0	0
Sim_0.7_B_number	278.300 ± 50.815	296.000 ± 76.463	353.700 ± 82.321	353.800 ± 92.454	157.200 ± 47.543	1.100 ± 1.300	0	0	0	0	0	0
Sim_0.7_C_number	1007.500 ± 221.375	924.200 ± 152.064	1135.400 ± 257.740	1073.600 ± 244.186	572.800 ± 211.134	4.800 ± 8.244	0.100 ± 0.300	0	0	0	0	0

Sim_0.7_D_number	902.000 ± 216.488	836.300 ± 166.707	1062.900 ± 203.210	936.300 ± 218.504	538.000 ± 137.844	5.000 ± 4.405	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0
Sim_0.7_E_number	429.900 ± 122.871	460.500 ± 110.186	550.500 ± 125.241	508.600 ± 106.708	250.700 ± 85.808	8.000 ± 4.171	0.700 ± 1.005	1.000 ± 0.775	0.300 ± 0.458	0.700 ± 0.781	0.800 ± 0.600	1.500 ± 0.806
Sim_0.7_train_0.7	0.204 ± 0.037	0.192 ± 0.034	0.187 ± 0.030	0.206 ± 0.044	0.262 ± 0.058	0.700 ± 0.263	0.500 ± 0.500	0.700 ± 0.458	0.300 ± 0.458	0.500 ± 0.500	0.700 ± 0.458	0.900 ± 0.300
Sim_0.7_train_0.7_number	527.200 ± 85.040	485.000 ± 103.276	578.700 ± 100.282	580.500 ± 101.575	402.800 ± 140.198	10.400 ± 6.037	0.800 ± 0.980	1.100 ± 0.831	0.300 ± 0.458	0.700 ± 0.781	0.800 ± 0.600	1.500 ± 0.806
Sim_0.7_train_0.7_A_number	3.100 ± 4.679	2.500 ± 2.837	5.100 ± 4.437	6.500 ± 6.830	3.300 ± 2.492	0	0	0	0	0	0	0
Sim_0.7_train_0.7_B_number	62.000 ± 18.515	58.800 ± 23.421	60.600 ± 18.602	74.800 ± 24.770	47.000 ± 26.050	0.800 ± 1.400	0	0	0	0	0	0
Sim_0.7_train_0.7_C_number	203.700 ± 43.343	187.600 ± 52.343	226.200 ± 60.526	220.800 ± 43.584	155.500 ± 74.075	1.600 ± 1.800	0	0	0	0	0	0
Sim_0.7_train_0.7_D_number	165.500 ± 25.935	143.300 ± 36.562	175.200 ± 27.455	168.000 ± 32.790	133.200 ± 46.925	2.300 ± 2.193	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0
Sim_0.7_train_0.7_E_number	92.900 ± 15.636	92.800 ± 12.319	111.600 ± 13.908	110.400 ± 23.234	63.800 ± 15.164	5.700 ± 2.283	0.700 ± 1.005	1.000 ± 0.775	0.300 ± 0.458	0.700 ± 0.781	0.800 ± 0.600	1.500 ± 0.806
Sim_0.8	0.025 ± 0.003	0.022 ± 0.003	0.028 ± 0.005	0.026 ± 0.004	0.013 ± 0.003	0	0	0	0	0	0	0
Sim_0.8_number	751.500 ± 90.944	670.300 ± 93.121	853.600 ± 139.405	792.700 ± 105.514	404.400 ± 88.550	3.700 ± 2.283	0.100 ± 0.300	0	0.100 ± 0.300	0.300 ± 0.458	0	0
Sim_0.8_A_number	4.200 ± 4.094	5.400 ± 4.363	7.700 ± 4.196	7.900 ± 5.735	4.400 ± 4.477	0	0	0	0	0	0	0
Sim_0.8_B_number	87.600 ± 24.063	88.600 ± 36.341	101.800 ± 32.227	111.200 ± 42.126	44.200 ± 17.820	0.300 ± 0.640	0	0	0	0	0	0

Sim_0.8_C_number	287.700 ± 62.944	240.700 ± 49.766	304.700 ± 87.759	295.400 ± 74.297	140.900 ± 53.784	0.800 ± 1.249	0	0	0	0	0	0
Sim_0.8_D_number	267.900 ± 61.989	227.800 ± 47.271	306.000 ± 67.281	254.300 ± 49.806	160.000 ± 29.987	1.100 ± 1.044	0	0	0	0	0	0
Sim_0.8_E_number	104.100 ± 27.413	107.800 ± 24.584	133.400 ± 36.634	123.900 ± 29.225	54.900 ± 20.945	1.500 ± 1.025	0.100 ± 0.300	0	0.100 ± 0.300	0.300 ± 0.458	0	0
Sim_0.8_train_0.7	0.166 ± 0.033	0.160 ± 0.033	0.151 ± 0.018	0.169 ± 0.033	0.204 ± 0.039	0.684 ± 0.362	0.100 ± 0.300	0	0.100 ± 0.300	0.300 ± 0.458	0	0
Sim_0.8_train_0.7_number	124.000 ± 24.831	107.300 ± 25.613	128.000 ± 21.194	133.700 ± 30.932	83.600 ± 25.660	2.400 ± 2.010	0.100 ± 0.300	0	0.100 ± 0.300	0.300 ± 0.458	0	0
Sim_0.8_train_0.7_A_number	0.700 ± 1.552	1.100 ± 1.136	1.500 ± 1.857	2.600 ± 3.800	0.700 ± 1.187	0	0	0	0	0	0	0
Sim_0.8_train_0.7_B_number	14.100 ± 5.735	14.100 ± 7.382	11.800 ± 4.468	19.300 ± 8.821	9.700 ± 5.021	0.300 ± 0.640	0	0	0	0	0	0
Sim_0.8_train_0.7_C_number	48.000 ± 14.401	42.900 ± 15.043	52.000 ± 12.900	49.000 ± 14.540	31.200 ± 14.979	0.300 ± 0.640	0	0	0	0	0	0
Sim_0.8_train_0.7_D_number	38.500 ± 8.547	28.800 ± 6.882	37.200 ± 5.582	35.400 ± 7.826	28.900 ± 11.458	0.600 ± 0.800	0	0	0	0	0	0
Sim_0.8_train_0.7_E_number	22.700 ± 5.041	20.400 ± 5.389	25.500 ± 7.032	27.400 ± 7.826	13.100 ± 4.657	1.200 ± 1.077	0.100 ± 0.300	0	0.100 ± 0.300	0.300 ± 0.458	0	0
Sim_0.9	0.008 ± 0.001	0.006 ± 0.001	0.008 ± 0.001	0.008 ± 0.001	0.004 ± 0.001	0	0	0	0	0	0	0
Sim_0.9_number	233.500 ± 32.601	185.900 ± 28.399	243.600 ± 40.981	236.200 ± 39.957	117.500 ± 34.106	1.800 ± 1.327	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Sim_0.9_A_number	0.700 ± 1.005	1.100 ± 0.831	1.900 ± 1.044	1.700 ± 1.005	1.300 ± 1.487	0	0	0	0	0	0	0

Sim_0.9_B_number	35.000 ± 10.789	30.600 ± 9.604	35.300 ± 8.798	40.600 ± 13.048	15.400 ± 6.711	0.200 ± 0.400	0	0	0	0	0	0
Sim_0.9_C_number	90.400 ± 19.043	67.600 ± 16.070	91.500 ± 24.748	87.200 ± 21.311	42.100 ± 17.734	0.300 ± 0.900	0	0	0	0	0	0
Sim_0.9_D_number	72.900 ± 17.824	52.400 ± 12.043	72.600 ± 13.507	66.900 ± 13.538	40.400 ± 10.641	0.200 ± 0.600	0	0	0	0	0	0
Sim_0.9_E_number	34.500 ± 6.515	34.200 ± 5.828	42.300 ± 9.023	39.800 ± 7.054	18.300 ± 6.664	1.100 ± 0.700	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7	0.188 ± 0.046	0.199 ± 0.043	0.196 ± 0.027	0.195 ± 0.043	0.236 ± 0.036	0.693 ± 0.408	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7_number	43.200 ± 9.379	36.500 ± 7.606	47.300 ± 7.281	46.000 ± 11.472	28.100 ± 9.944	1.300 ± 0.900	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7_A_number	0.100 ± 0.300	0.200 ± 0.400	0.300 ± 0.900	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0	0	0	0
Sim_0.9_train_0.7_B_number	4.600 ± 2.154	4.500 ± 2.335	3.900 ± 1.300	6.000 ± 2.000	2.800 ± 1.166	0.200 ± 0.400	0	0	0	0	0	0
Sim_0.9_train_0.7_C_number	17.200 ± 5.212	13.600 ± 4.341	18.900 ± 4.784	16.200 ± 5.828	9.800 ± 5.913	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.9_train_0.7_D_number	12.300 ± 3.662	9.800 ± 2.561	12.600 ± 2.615	12.400 ± 4.055	9.900 ± 3.780	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.9_train_0.7_E_number	9.000 ± 2.000	8.400 ± 2.764	11.600 ± 2.538	11.200 ± 3.487	5.400 ± 2.498	0.900 ± 0.700	0.100 ± 0.300	0	0.100 ± 0.300	0	0	0

Table S4. The fine-tuning results on the EGFR 10%-fine-tuning datasets with RDKit filtering.

EGFR	CharRNN	AAE	VAE	Reinvent	ORGAN	GraphAF	RNNAttn			TransVAE		
							rand	high entropy	k high entropy	rand	high entropy	k high entropy
IntDiv	0.856 ± 0.006	0.863 ± 0.006	0.848 ± 0.007	0.855 ± 0.006	0.855 ± 0.004	0.900 ± 0.002	1	1	1	1	1	1
SNN/Gen_train	0.463 ± 0.021	0.477 ± 0.017	0.529 ± 0.017	0.467 ± 0.037	0.450 ± 0.010	0.237 ± 0.007	0.234 ± 0.003	0.283 ± 0.006	0.235 ± 0.003	0.220 ± 0.003	0.254 ± 0.005	0.217 ± 0.003
SNN/Gen_goal	0.450 ± 0.016	0.448 ± 0.015	0.487 ± 0.015	0.451 ± 0.021	0.429 ± 0.010	0.276 ± 0.006	0.270 ± 0.003	0.336 ± 0.003	0.271 ± 0.003	0.254 ± 0.003	0.298 ± 0.003	0.252 ± 0.003
IntDiv_ Rediscovery	0.764 ± 0.019	0.774 ± 0.022	0.768 ± 0.018	0.764 ± 0.021	0.750 ± 0.034	0.459 ± 0.026	/	/	/	/	/	/
SNN/Rediscovery_ train	0.761 ± 0.024	0.767 ± 0.019	0.761 ± 0.016	0.757 ± 0.014	0.748 ± 0.030	0.545 ± 0.046	/	0.312 ± 0.004	0.652 ± 0.000	/	/	/
Rediscovery	0.003 ± 0.000	0.003 ± 0.001	0.004 ± 0.001	0.003 ± 0.000	0.001 ± 0.000	0	0	0	0	0	0	0
Rediscovery_ number	94.700 ± 11.393	86.700 ± 15.120	121.800 ± 16.981	84.300 ± 13.624	44.200 ± 12.319	9.600 ± 6.989	0	0.200 ± 0.400	0.100 ± 0.300	0	0	0
Rediscovery_A_ number	0.600 ± 0.663	0.700 ± 0.640	0.800 ± 0.600	0.800 ± 0.748	0.200 ± 0.600	0	0	0	0	0	0	0
Rediscovery_B_ number	29.100 ± 7.449	27.400 ± 8.924	40.500 ± 5.371	25.300 ± 4.713	13.700 ± 5.515	3.100 ± 2.809	0	0	0	0	0	0
Rediscovery_C_ number	36.000 ± 6.856	29.900 ± 8.191	38.700 ± 7.537	28.900 ± 7.341	15.400 ± 5.571	4.300 ± 2.571	0	0	0.100 ± 0.300	0	0	0
Rediscovery_D_ number	14.800 ± 4.094	13.500 ± 4.295	20.600 ± 4.477	14.000 ± 4.626	6.700 ± 3.035	1.200 ± 1.536	0	0	0	0	0	0

Rediscovery_E_number	14.200 ± 4.045	15.200 ± 4.308	21.200 ± 7.626	15.300 ± 4.001	8.200 ± 2.227	1.000 ± 1.095	0	0.200 ± 0.400	0	0	0	0
Rediscovery_0.7	0.349 ± 0.051	0.325 ± 0.064	0.351 ± 0.043	0.342 ± 0.037	0.395 ± 0.110	0.924 ± 0.071	0	0.200 ± 0.400	0.100 ± 0.300	0	0	0
Rediscovery_0.7_number	32.900 ± 5.375	28.000 ± 6.738	42.800 ± 7.808	28.700 ± 4.562	17.500 ± 7.852	8.500 ± 5.852	0	0.200 ± 0.400	0.100 ± 0.300	0	0	0
Rediscovery_0.7_A_number	0	0.100 ± 0.300	0.200 ± 0.400	0	0.100 ± 0.300	0	0	0	0	0	0	0
Rediscovery_0.7_B_number	10.300 ± 3.579	8.800 ± 4.094	14.600 ± 4.271	8.500 ± 3.324	5.000 ± 3.435	2.400 ± 2.289	0	0	0	0	0	0
Rediscovery_0.7_C_number	10.100 ± 4.277	8.600 ± 3.470	11.800 ± 3.458	10.100 ± 3.506	6.900 ± 3.700	4.000 ± 2.324	0	0	0.100 ± 0.300	0	0	0
Rediscovery_0.7_D_number	6.000 ± 2.793	4.700 ± 2.685	7.900 ± 1.921	4.200 ± 1.077	2.300 ± 1.952	1.100 ± 1.375	0	0	0	0	0	0
Rediscovery_0.7_E_number	6.500 ± 2.872	5.800 ± 2.676	8.300 ± 3.900	5.900 ± 2.508	3.200 ± 1.939	1.000 ± 1.095	0	0.200 ± 0.400	0	0	0	0
Sim_0.7	0.081 ± 0.015	0.085 ± 0.015	0.109 ± 0.016	0.083 ± 0.015	0.045 ± 0.011	0.002 ± 0.001	0	0	0	0	0	0
Sim_0.7_number	2423.600 ± 443.933	2541.500 ± 436.924	3277.900 ± 490.851	2490.600 ± 454.132	1356.500 ± 343.104	64.500 ± 37.742	1.200 ± 1.166	5.700 ± 2.369	0.700 ± 0.640	0.100 ± 0.300	0.300 ± 0.458	0.100 ± 0.300
Sim_0.7_A_number	17.400 ± 14.305	20.200 ± 16.394	35.600 ± 25.784	20.800 ± 14.112	8.000 ± 6.971	0	0	0	0	0	0	0
Sim_0.7_B_number	679.300 ± 162.013	710.200 ± 140.609	954.800 ± 171.650	715.500 ± 146.181	387.400 ± 97.712	15.000 ± 10.900	0.200 ± 0.400	0.100 ± 0.300	0.100 ± 0.300	0	0	0
Sim_0.7_C_number	1013.000 ± 284.269	908.900 ± 320.473	1252.100 ± 304.898	956.900 ± 365.619	523.700 ± 260.115	22.700 ± 14.512	0.500 ± 0.671	0.100 ± 0.300	0.100 ± 0.300	0	0	0

Sim_0.7_D_number	387.500 ± 139.887	480.600 ± 130.722	552.500 ± 137.853	400.200 ± 99.236	216.000 ± 61.792	12.300 ± 7.100	0.100 ± 0.300	0.500 ± 0.671	0.300 ± 0.458	0	0	0.100 ± 0.300
Sim_0.7_E_number	326.400 ± 86.355	421.600 ± 117.002	482.900 ± 116.230	397.200 ± 105.446	221.400 ± 67.435	14.500 ± 7.890	0.400 ± 0.490	5.000 ± 1.949	0.200 ± 0.400	0.100 ± 0.300	0.300 ± 0.458	0
Sim_0.7_train_0.7	0.266 ± 0.034	0.232 ± 0.034	0.262 ± 0.031	0.251 ± 0.028	0.383 ± 0.060	0.887 ± 0.108	0.700 ± 0.458	0.887 ± 0.298	0.600 ± 0.490	0.100 ± 0.300	0.300 ± 0.458	0.100 ± 0.300
Sim_0.7_train_0.7_number	631.000 ± 58.657	583.000 ± 95.734	845.100 ± 75.252	620.300 ± 105.183	514.500 ± 136.903	57.100 ± 35.342	1.200 ± 1.166	5.500 ± 2.500	0.700 ± 0.640	0.100 ± 0.300	0.300 ± 0.458	0.100 ± 0.300
Sim_0.7_train_0.7_A_number	3.900 ± 2.508	6.000 ± 7.497	10.700 ± 8.486	5.900 ± 4.415	4.400 ± 4.128	0	0	0	0	0	0	0
Sim_0.7_train_0.7_B_number	213.000 ± 40.748	206.800 ± 53.224	293.300 ± 49.510	214.000 ± 43.914	169.500 ± 44.733	13.900 ± 10.084	0.200 ± 0.400	0.100 ± 0.300	0.100 ± 0.300	0	0	0
Sim_0.7_train_0.7_C_number	228.200 ± 40.948	180.300 ± 34.276	289.300 ± 49.592	214.500 ± 45.724	191.900 ± 79.241	19.700 ± 11.550	0.500 ± 0.671	0.100 ± 0.300	0.100 ± 0.300	0	0	0
Sim_0.7_train_0.7_D_number	85.200 ± 14.204	89.000 ± 17.567	117.700 ± 14.670	79.400 ± 19.319	70.100 ± 18.923	11.200 ± 7.414	0.100 ± 0.300	0.500 ± 0.671	0.300 ± 0.458	0	0	0.100 ± 0.300
Sim_0.7_train_0.7_E_number	100.700 ± 16.947	100.900 ± 23.985	134.100 ± 35.317	106.500 ± 28.643	78.600 ± 28.977	12.300 ± 8.403	0.400 ± 0.490	4.800 ± 2.088	0.200 ± 0.400	0.100 ± 0.300	0.300 ± 0.458	0
Sim_0.8	0.019 ± 0.002	0.019 ± 0.003	0.026 ± 0.003	0.019 ± 0.003	0.010 ± 0.002	0.001 ± 0.000	0	0	0	0	0	0
Sim_0.8_number	579.900 ± 70.026	562.400 ± 87.438	794.700 ± 101.538	575.300 ± 85.725	290.700 ± 65.224	18.200 ± 13.681	0	2.100 ± 1.044	0.100 ± 0.300	0	0	0
Sim_0.8_A_number	5.700 ± 5.934	6.300 ± 6.558	11.100 ± 7.842	6.700 ± 4.291	2.200 ± 2.272	0	0	0	0	0	0	0
Sim_0.8_B_number	161.600 ± 38.502	163.700 ± 40.259	247.400 ± 46.571	170.700 ± 40.147	80.000 ± 26.027	6.000 ± 5.550	0	0	0	0	0	0

Sim_0.8_C_number	250.000 ± 44.197	201.300 ± 55.249	291.400 ± 42.912	216.400 ± 56.763	110.200 ± 48.346	7.300 ± 5.515	0	0	0.100 ± 0.300	0	0	0
Sim_0.8_D_number	88.500 ± 28.454	103.000 ± 30.552	133.600 ± 30.180	95.200 ± 22.031	47.800 ± 15.924	2.000 ± 2.049	0	0	0	0	0	0
Sim_0.8_E_number	74.100 ± 14.970	88.100 ± 19.071	111.200 ± 29.305	86.300 ± 19.698	50.500 ± 11.902	2.900 ± 1.921	0	2.100 ± 1.044	0	0	0	0
Sim_0.8_train_0.7	0.247 ± 0.036	0.204 ± 0.034	0.229 ± 0.039	0.237 ± 0.032	0.333 ± 0.094	0.846 ± 0.186	0	0.800 ± 0.400	0.100 ± 0.300	0	0	0
Sim_0.8_train_0.7_number	141.900 ± 15.833	114.400 ± 22.699	179.200 ± 20.923	135.000 ± 20.881	96.400 ± 38.645	15.400 ± 11.271	0	2.000 ± 1.183	0.100 ± 0.300	0	0	0
Sim_0.8_train_0.7_A_number	1.100 ± 1.446	1.500 ± 2.335	1.900 ± 2.022	1.700 ± 1.676	1.200 ± 1.536	0	0	0	0	0	0	0
Sim_0.8_train_0.7_B_number	45.300 ± 9.675	41.200 ± 11.034	64.600 ± 15.743	45.100 ± 12.029	27.900 ± 10.251	5.300 ± 4.900	0	0	0	0	0	0
Sim_0.8_train_0.7_C_number	53.000 ± 12.474	33.900 ± 10.492	58.300 ± 12.822	45.800 ± 10.666	37.300 ± 23.529	5.900 ± 3.885	0	0	0.100 ± 0.300	0	0	0
Sim_0.8_train_0.7_D_number	17.800 ± 5.706	16.400 ± 6.453	24.400 ± 3.292	16.800 ± 3.544	13.100 ± 5.281	1.900 ± 1.921	0	0	0	0	0	0
Sim_0.8_train_0.7_E_number	24.700 ± 6.482	21.400 ± 4.673	30.000 ± 10.440	25.600 ± 6.873	16.900 ± 8.227	2.300 ± 1.900	0	2.000 ± 1.183	0	0	0	0
Sim_0.9	0.005 ± 0.000	0.005 ± 0.001	0.007 ± 0.001	0.005 ± 0.001	0.002 ± 0.001	0	0	0	0	0	0	0
Sim_0.9_number	147.300 ± 14.731	137.800 ± 18.766	202.200 ± 25.880	137.200 ± 20.702	69.400 ± 18.134	10.800 ± 7.481	0	0.600 ± 0.663	0.100 ± 0.300	0	0	0
Sim_0.9_A_number	0.900 ± 0.831	1.700 ± 1.345	3.000 ± 2.324	2.200 ± 1.470	0.500 ± 0.671	0	0	0	0	0	0	0

Sim_0.9_B_number	43.900 ± 14.377	42.000 ± 11.278	65.200 ± 13.121	38.300 ± 7.925	21.200 ± 7.871	3.300 ± 3.100	0	0	0	0	0	0
Sim_0.9_C_number	61.400 ± 10.547	49.000 ± 10.498	70.800 ± 9.908	50.700 ± 14.142	25.000 ± 12.474	4.900 ± 3.113	0	0	0.100 ± 0.300	0	0	0
Sim_0.9_D_number	22.100 ± 6.789	22.700 ± 6.310	34.500 ± 5.590	25.100 ± 7.543	11.500 ± 6.021	1.300 ± 1.735	0	0	0	0	0	0
Sim_0.9_E_number	19.000 ± 3.975	22.400 ± 4.695	28.700 ± 8.533	20.900 ± 4.527	11.200 ± 2.676	1.300 ± 1.005	0	0.600 ± 0.663	0	0	0	0
Sim_0.9_train_0.7	0.286 ± 0.044	0.250 ± 0.064	0.274 ± 0.038	0.282 ± 0.047	0.347 ± 0.106	0.856 ± 0.142	0	0.500 ± 0.500	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7_number	41.700 ± 5.442	34.000 ± 8.112	54.900 ± 8.396	38.000 ± 4.796	23.500 ± 9.615	9.200 ± 6.194	0	0.600 ± 0.663	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7_A_number	0.200 ± 0.400	0.300 ± 0.900	0.700 ± 0.900	0.600 ± 0.917	0.200 ± 0.400	0	0	0	0	0	0	0
Sim_0.9_train_0.7_B_number	13.600 ± 4.409	11.400 ± 3.527	20.500 ± 5.239	11.700 ± 4.196	7.000 ± 3.975	2.600 ± 2.577	0	0	0	0	0	0
Sim_0.9_train_0.7_C_number	13.100 ± 4.742	9.800 ± 4.261	15.200 ± 4.622	13.000 ± 3.950	9.500 ± 6.037	4.200 ± 2.358	0	0	0.100 ± 0.300	0	0	0
Sim_0.9_train_0.7_D_number	6.800 ± 2.960	5.200 ± 2.482	9.200 ± 2.400	5.600 ± 1.908	2.800 ± 2.135	1.200 ± 1.600	0	0	0	0	0	0
Sim_0.9_train_0.7_E_number	8.000 ± 2.098	7.300 ± 2.685	9.300 ± 3.743	7.100 ± 1.921	4.000 ± 1.844	1.200 ± 1.077	0	0.600 ± 0.663	0	0	0	0

Table S5. The fine-tuning results on the JakA 10%-fine-tuning datasets with RDKit filtering.

JakA	CharRNN	AAE	VAE	Reinvent	ORGAN	GraphAF	RNNAttn			TransVAE		
							rand	high entropy	k high entropy	rand	high entropy	k high entropy
IntDiv	0.849 ± 0.003	0.855 ± 0.003	0.845 ± 0.004	0.848 ± 0.004	0.853 ± 0.003	0.905 ± 0.001	1	1	1	1	1	1
SNN/Gen_train	0.547 ± 0.018	0.547 ± 0.015	0.582 ± 0.015	0.563 ± 0.018	0.494 ± 0.015	0.233 ± 0.004	0.245 ± 0.004	0.285 ± 0.007	0.246 ± 0.004	0.228 ± 0.003	0.256 ± 0.003	0.226 ± 0.003
SNN/Gen_goal	0.529 ± 0.014	0.517 ± 0.015	0.545 ± 0.013	0.538 ± 0.013	0.473 ± 0.012	0.264 ± 0.003	0.277 ± 0.003	0.332 ± 0.005	0.277 ± 0.003	0.257 ± 0.002	0.292 ± 0.003	0.256 ± 0.002
IntDiv_ Rediscovery	0.803 ± 0.005	0.802 ± 0.008	0.804 ± 0.007	0.797 ± 0.010	0.790 ± 0.011	0.478 ± 0.079	/	/	/	/	/	/
SNN/Rediscovery_ train	0.811 ± 0.011	0.802 ± 0.013	0.809 ± 0.008	0.805 ± 0.011	0.808 ± 0.012	0.662 ± 0.147	/	/	/	/	/	/
Rediscovery	0.010 ± 0.001	0.006 ± 0.001	0.009 ± 0.001	0.008 ± 0.001	0.005 ± 0.001	0	0	0	0	0	0	0
Rediscovery_ number	291.400 ± 28.232	188.000 ± 37.124	256.000 ± 32.336	248.300 ± 28.054	142.800 ± 27.051	3.100 ± 2.071	0	0	0	0	0	0
Rediscovery_A_ number	0.600 ± 0.663	0.300 ± 0.640	0.400 ± 0.490	0.500 ± 0.671	0.300 ± 0.458	0	0	0	0	0	0	0
Rediscovery_B_ number	177.500 ± 17.385	109.100 ± 25.078	160.400 ± 26.796	157.900 ± 22.832	89.900 ± 21.787	0.500 ± 0.671	0	0	0	0	0	0
Rediscovery_C_ number	56.900 ± 10.163	43.900 ± 11.743	50.800 ± 9.495	50.400 ± 10.062	27.200 ± 7.730	0.900 ± 1.044	0	0	0	0	0	0
Rediscovery_D_ number	45.200 ± 9.704	26.500 ± 5.536	34.600 ± 8.628	29.900 ± 6.640	18.700 ± 5.120	1.300 ± 1.100	0	0	0	0	0	0

Rediscovery_E_ number	11.200 ± 4.600	8.200 ± 3.341	9.800 ± 3.655	9.600 ± 2.375	6.700 ± 4.173	0.400 ± 0.663	0	0	0	0	0	0
Rediscovery_0.7	0.143 ± 0.044	0.177 ± 0.049	0.153 ± 0.024	0.157 ± 0.034	0.163 ± 0.036	0.612 ± 0.387	0	0	0	0	0	0
Rediscovery_0.7_ number	41.100 ± 10.540	34.200 ± 13.235	38.800 ± 5.963	39.000 ± 9.445	22.800 ± 5.075	2.200 ± 1.833	0	0	0	0	0	0
Rediscovery_0.7_A_ number	0.300 ± 0.640	0.300 ± 0.640	0.200 ± 0.400	0.100 ± 0.300	0.200 ± 0.400	0	0	0	0	0	0	0
Rediscovery_0.7_B_ number	21.800 ± 5.879	16.300 ± 8.149	19.600 ± 5.276	21.100 ± 4.571	13.500 ± 3.202	0.200 ± 0.600	0	0	0	0	0	0
Rediscovery_0.7_C_ number	8.900 ± 3.300	9.300 ± 2.759	9.500 ± 3.138	10.300 ± 3.318	4.400 ± 2.107	0.600 ± 0.917	0	0	0	0	0	0
Rediscovery_0.7_D_ number	7.600 ± 2.615	6.100 ± 3.360	7.000 ± 3.225	4.700 ± 2.795	3.900 ± 1.446	1.000 ± 1.000	0	0	0	0	0	0
Rediscovery_0.7_E_ number	2.500 ± 1.962	2.200 ± 1.600	2.500 ± 0.922	2.800 ± 1.720	0.800 ± 0.600	0.400 ± 0.663	0	0	0	0	0	0
Sim_0.7	0.216 ± 0.021	0.192 ± 0.024	0.224 ± 0.024	0.232 ± 0.021	0.116 ± 0.011	0.001 ± 0.001	0	0	0	0	0	0
Sim_0.7_number	6489.400 ± 643.140	5757.500 ± 733.929	6725.100 ± 724.132	6950.900 ± 630.187	3470.200 ± 336.633	38.200 ± 21.320	0.200 ± 0.600	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0.100 ± 0.300
Sim_0.7_A_ number	9.700 ± 6.116	5.000 ± 5.020	10.700 ± 8.672	6.800 ± 3.487	5.100 ± 4.110	0.600 ± 1.497	0	0	0	0	0	0
Sim_0.7_B_ number	4097.200 ± 600.339	3531.600 ± 570.642	4228.300 ± 662.964	4479.100 ± 642.559	2160.100 ± 324.395	11.700 ± 5.061	0.100 ± 0.300	0	0	0	0	0
Sim_0.7_C_ number	1257.200 ± 248.873	1269.100 ± 304.890	1348.400 ± 286.819	1337.500 ± 232.572	725.800 ± 198.237	12.100 ± 11.300	0	0	0	0	0	0

Sim_0.7_D_number	877.900 ± 180.952	734.600 ± 120.301	876.100 ± 141.933	880.000 ± 203.173	434.700 ± 63.619	10.800 ± 7.222	0.100 ± 0.300	0	0	0	0	0.100 ± 0.300
Sim_0.7_E_number	247.400 ± 78.332	217.200 ± 55.429	261.600 ± 78.550	247.500 ± 70.592	144.500 ± 60.232	3.000 ± 2.098	0	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0
Sim_0.7_train_0.7	0.167 ± 0.022	0.170 ± 0.020	0.157 ± 0.025	0.166 ± 0.026	0.207 ± 0.024	0.531 ± 0.250	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0.100 ± 0.300
Sim_0.7_train_0.7_number	1076.300 ± 109.879	977.000 ± 152.479	1041.900 ± 139.537	1147.500 ± 159.560	714.200 ± 88.429	16.900 ± 7.622	0.200 ± 0.600	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0.100 ± 0.300
Sim_0.7_train_0.7_A_number	5.800 ± 3.655	2.300 ± 2.759	6.000 ± 5.692	3.000 ± 1.549	2.600 ± 2.538	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.7_train_0.7_B_number	586.400 ± 66.443	536.300 ± 115.594	561.100 ± 85.988	690.400 ± 128.031	400.800 ± 62.705	5.100 ± 2.773	0.100 ± 0.300	0	0	0	0	0
Sim_0.7_train_0.7_C_number	237.600 ± 51.821	251.800 ± 48.773	249.100 ± 46.956	236.800 ± 45.851	167.500 ± 32.116	3.300 ± 2.100	0	0	0	0	0	0
Sim_0.7_train_0.7_D_number	188.000 ± 36.238	138.200 ± 26.396	165.300 ± 39.388	164.600 ± 30.693	108.600 ± 15.538	6.400 ± 4.609	0.100 ± 0.300	0	0	0	0	0.100 ± 0.300
Sim_0.7_train_0.7_E_number	58.500 ± 18.495	48.400 ± 10.346	60.400 ± 18.731	52.700 ± 17.298	34.700 ± 9.487	2.000 ± 1.844	0	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0
Sim_0.8	0.073 ± 0.008	0.057 ± 0.009	0.072 ± 0.009	0.074 ± 0.009	0.036 ± 0.004	0	0	0	0	0	0	0
Sim_0.8_number	2197.200 ± 247.632	1708.000 ± 274.462	2153.800 ± 268.121	2217.000 ± 278.861	1091.900 ± 121.114	10.400 ± 4.779	0	0	0	0	0	0
Sim_0.8_A_number	2.200 ± 1.990	1.300 ± 2.002	3.300 ± 3.407	1.600 ± 1.356	1.000 ± 1.183	0.200 ± 0.600	0	0	0	0	0	0
Sim_0.8_B_number	1475.700 ± 202.942	1128.100 ± 209.504	1447.300 ± 216.880	1551.400 ± 253.827	745.900 ± 129.523	2.600 ± 1.800	0	0	0	0	0	0

Sim_0.8_C_number	352.000 ± 78.616	324.000 ± 90.705	362.500 ± 91.610	350.000 ± 74.847	177.600 ± 51.411	3.500 ± 2.540	0	0	0	0	0	0
Sim_0.8_D_number	292.100 ± 68.645	198.100 ± 49.101	262.300 ± 59.621	245.000 ± 68.462	121.300 ± 39.875	3.300 ± 2.326	0	0	0	0	0	0
Sim_0.8_E_number	75.200 ± 26.555	56.500 ± 22.756	78.400 ± 25.629	69.000 ± 23.912	46.100 ± 24.489	0.800 ± 0.748	0	0	0	0	0	0
Sim_0.8_train_0.7	0.104 ± 0.018	0.103 ± 0.021	0.096 ± 0.015	0.099 ± 0.015	0.123 ± 0.014	0.495 ± 0.266	0	0	0	0	0	0
Sim_0.8_train_0.7_number	225.100 ± 30.969	174.000 ± 37.032	204.500 ± 26.440	216.100 ± 30.726	133.700 ± 14.423	5.200 ± 2.960	0	0	0	0	0	0
Sim_0.8_train_0.7_A_number	1.300 ± 1.487	0.700 ± 1.187	2.200 ± 2.482	0.500 ± 0.500	0.600 ± 1.020	0	0	0	0	0	0	0
Sim_0.8_train_0.7_B_number	120.400 ± 19.022	90.500 ± 21.722	105.000 ± 16.474	124.900 ± 17.997	77.600 ± 11.173	1.000 ± 0.894	0	0	0	0	0	0
Sim_0.8_train_0.7_C_number	50.400 ± 16.427	48.700 ± 16.493	51.600 ± 11.851	50.300 ± 12.570	28.700 ± 7.656	1.100 ± 1.136	0	0	0	0	0	0
Sim_0.8_train_0.7_D_number	41.300 ± 8.741	25.800 ± 5.741	33.300 ± 9.950	30.300 ± 4.691	20.900 ± 7.300	2.600 ± 2.200	0	0	0	0	0	0
Sim_0.8_train_0.7_E_number	11.700 ± 4.838	8.300 ± 2.452	12.400 ± 5.370	10.100 ± 2.427	5.900 ± 2.071	0.500 ± 0.671	0	0	0	0	0	0
Sim_0.9	0.021 ± 0.002	0.014 ± 0.002	0.019 ± 0.003	0.020 ± 0.002	0.011 ± 0.002	0	0	0	0	0	0	0
Sim_0.9_number	644.500 ± 61.391	429.900 ± 71.884	584.500 ± 75.484	589.400 ± 74.864	327.000 ± 50.164	6.300 ± 3.002	0	0	0	0	0	0
Sim_0.9_A_number	1.400 ± 1.281	0.600 ± 0.917	1.700 ± 1.792	0.700 ± 0.640	0.600 ± 0.917	0.100 ± 0.300	0	0	0	0	0	0

Sim_0.9_B_number	403.100 ± 50.091	260.800 ± 53.241	376.800 ± 71.366	392.100 ± 85.803	219.700 ± 49.477	1.500 ± 1.285	0	0	0	0	0	0
Sim_0.9_C_number	119.700 ± 28.517	96.000 ± 30.100	110.600 ± 29.104	104.600 ± 23.376	56.000 ± 18.396	1.400 ± 0.917	0	0	0	0	0	0
Sim_0.9_D_number	94.100 ± 26.838	57.100 ± 15.611	73.600 ± 20.096	70.400 ± 25.566	36.000 ± 14.772	2.600 ± 1.855	0	0	0	0	0	0
Sim_0.9_E_number	26.200 ± 11.241	15.400 ± 6.422	21.800 ± 6.823	21.600 ± 7.592	14.700 ± 7.760	0.700 ± 0.781	0	0	0	0	0	0
Sim_0.9_train_0.7	0.099 ± 0.025	0.115 ± 0.032	0.102 ± 0.019	0.100 ± 0.025	0.114 ± 0.018	0.549 ± 0.310	0	0	0	0	0	0
Sim_0.9_train_0.7_number	62.900 ± 13.034	50.200 ± 17.122	58.700 ± 10.546	57.500 ± 12.200	37.000 ± 5.710	3.900 ± 2.427	0	0	0	0	0	0
Sim_0.9_train_0.7_A_number	1.000 ± 1.265	0.500 ± 0.922	1.200 ± 1.327	0.300 ± 0.458	0.400 ± 0.917	0	0	0	0	0	0	0
Sim_0.9_train_0.7_B_number	30.500 ± 6.903	24.200 ± 9.673	27.400 ± 6.264	31.600 ± 7.003	20.900 ± 4.679	0.500 ± 0.671	0	0	0	0	0	0
Sim_0.9_train_0.7_C_number	16.000 ± 4.733	15.500 ± 7.131	16.200 ± 4.308	14.700 ± 2.968	8.500 ± 4.801	0.700 ± 0.781	0	0	0	0	0	0
Sim_0.9_train_0.7_D_number	12.000 ± 3.715	7.400 ± 3.072	9.400 ± 4.673	7.300 ± 4.001	5.800 ± 3.156	2.200 ± 1.887	0	0	0	0	0	0
Sim_0.9_train_0.7_E_number	3.400 ± 2.577	2.600 ± 1.497	4.500 ± 2.062	3.600 ± 1.908	1.400 ± 0.800	0.500 ± 0.671	0	0	0	0	0	0

Table S6. The fine-tuning results on the PDGFR 10%-fine-tuning datasets with RDKit filtering.

PDGFR	CharRNN	AAE	VAE	Reinvent	ORGAN	GraphAF	RNNAttn			TransVAE		
							rand	high entropy	k high entropy	rand	high entropy	k high entropy
IntDiv	0.855 ± 0.003	0.861 ± 0.002	0.850 ± 0.004	0.853 ± 0.004	0.855 ± 0.004	0.898 ± 0.002	1	1	1	1	1	1
SNN/Gen_train	0.495 ± 0.026	0.512 ± 0.016	0.543 ± 0.016	0.519 ± 0.023	0.471 ± 0.017	0.247 ± 0.004	0.251 ± 0.005	0.306 ± 0.009	0.252 ± 0.005	0.238 ± 0.005	0.280 ± 0.007	0.236 ± 0.005
SNN/Gen_goal	0.467 ± 0.016	0.464 ± 0.012	0.489 ± 0.011	0.481 ± 0.015	0.437 ± 0.011	0.279 ± 0.003	0.285 ± 0.005	0.356 ± 0.006	0.285 ± 0.005	0.270 ± 0.005	0.321 ± 0.006	0.269 ± 0.005
IntDiv_ Rediscovery	0.806 ± 0.007	0.813 ± 0.006	0.811 ± 0.007	0.810 ± 0.013	0.798 ± 0.009	0.427 ± 0.016	/	/	/	/	/	/
SNN/Rediscovery_ train	0.782 ± 0.017	0.768 ± 0.020	0.777 ± 0.016	0.774 ± 0.018	0.755 ± 0.026	0.468 ± 0.147	0.286 ± 0.000	0.311 ± 0.000	/	/	/	0.439 ± 0.000
Rediscovery	0.003 ± 0.001	0.003 ± 0.000	0.004 ± 0.000	0.003 ± 0.001	0.002 ± 0.000	0	0	0	0	0	0	0
Rediscovery_ number	94.100 ± 15.056	88.400 ± 10.920	113.200 ± 10.898	101.100 ± 16.967	52.600 ± 13.410	1.100 ± 0.831	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.100 ± 0.300
Rediscovery_A_ number	0.700 ± 0.458	1.200 ± 0.748	1.400 ± 0.800	1.200 ± 0.872	1.000 ± 0.632	0	0	0	0	0	0	0
Rediscovery_B_ number	34.400 ± 5.869	28.400 ± 6.785	41.400 ± 5.903	37.800 ± 8.818	18.600 ± 7.242	0	0	0	0	0	0	0
Rediscovery_C_ number	24.200 ± 8.268	22.300 ± 3.900	27.700 ± 4.076	23.200 ± 7.026	11.500 ± 4.822	0.200 ± 0.400	0	0	0	0	0	0.100 ± 0.300
Rediscovery_D_ number	22.400 ± 6.453	24.000 ± 7.335	27.800 ± 5.288	23.500 ± 4.319	13.300 ± 3.579	0.400 ± 0.663	0	0	0	0	0	0

Rediscovery_E_ number	12.400 ± 3.878	12.500 ± 4.129	14.900 ± 3.885	15.400 ± 2.764	8.200 ± 4.142	0.500 ± 0.671	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0
Rediscovery_0.7	0.217 ± 0.043	0.258 ± 0.051	0.210 ± 0.036	0.228 ± 0.048	0.296 ± 0.082	0.600 ± 0.436	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.100 ± 0.300
Rediscovery_0.7_ number	20.200 ± 4.190	23.100 ± 6.332	23.800 ± 4.854	22.800 ± 5.173	15.200 ± 4.771	0.900 ± 0.700	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.100 ± 0.300
Rediscovery_0.7_A_ number	0.300 ± 0.458	0.400 ± 0.663	0.300 ± 0.640	0.300 ± 0.640	0.500 ± 0.671	0	0	0	0	0	0	0
Rediscovery_0.7_B_ number	5.400 ± 1.281	5.200 ± 2.135	7.000 ± 2.366	7.000 ± 2.191	3.200 ± 1.778	0	0	0	0	0	0	0
Rediscovery_0.7_C_ number	4.500 ± 1.910	5.900 ± 3.048	4.400 ± 2.332	5.300 ± 2.452	2.800 ± 1.661	0.200 ± 0.400	0	0	0	0	0	0.100 ± 0.300
Rediscovery_0.7_D_ number	5.100 ± 2.508	8.000 ± 3.661	6.400 ± 2.010	5.300 ± 2.532	4.100 ± 0.943	0.300 ± 0.458	0	0	0	0	0	0
Rediscovery_0.7_E_ number	4.900 ± 2.548	3.600 ± 2.245	5.700 ± 2.369	4.900 ± 1.044	4.600 ± 2.973	0.400 ± 0.663	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0
Sim_0.7	0.090 ± 0.012	0.089 ± 0.012	0.107 ± 0.011	0.105 ± 0.014	0.052 ± 0.009	0	0	0	0	0	0	0
Sim_0.7_number	2710.000 ± 371.068	2671.300 ± 352.329	3206.800 ± 330.075	3161.400 ± 430.798	1562.200 ± 278.280	14.100 ± 10.242	1.200 ± 0.980	4.000 ± 2.049	0.600 ± 0.917	0.500 ± 0.671	0.400 ± 0.490	0.400 ± 0.490
Sim_0.7_A_ number	26.300 ± 10.286	21.100 ± 9.700	30.000 ± 10.982	26.600 ± 12.508	22.200 ± 12.205	0	0	0	0	0	0	0
Sim_0.7_B_ number	987.400 ± 253.048	917.200 ± 271.442	1154.100 ± 253.892	1143.800 ± 337.098	566.700 ± 170.084	2.200 ± 2.315	0.200 ± 0.600	0.200 ± 0.400	0	0.100 ± 0.300	0	0
Sim_0.7_C_ number	654.100 ± 162.558	676.700 ± 169.790	806.600 ± 186.199	805.900 ± 219.043	361.800 ± 117.586	1.700 ± 1.418	0.100 ± 0.300	0.600 ± 0.663	0.100 ± 0.300	0	0	0.200 ± 0.400

Sim_0.7_D_number	620.000 ± 179.126	604.700 ± 93.319	719.000 ± 184.152	703.800 ± 174.610	344.700 ± 80.304	3.600 ± 6.264	0.400 ± 0.490	0.500 ± 0.671	0	0	0	0.100 ± 0.300
Sim_0.7_E_number	422.200 ± 71.292	451.600 ± 84.506	497.100 ± 68.940	481.300 ± 82.275	266.800 ± 75.930	6.600 ± 3.980	0.500 ± 0.671	2.700 ± 1.676	0.500 ± 0.671	0.400 ± 0.490	0.400 ± 0.490	0.100 ± 0.300
Sim_0.7_train_0.7	0.179 ± 0.038	0.162 ± 0.030	0.165 ± 0.025	0.167 ± 0.036	0.247 ± 0.061	0.732 ± 0.221	0.700 ± 0.458	1.000 ± 0.000	0.400 ± 0.490	0.400 ± 0.490	0.400 ± 0.490	0.400 ± 0.490
Sim_0.7_train_0.7_number	473.800 ± 46.117	424.700 ± 49.242	523.200 ± 38.204	515.600 ± 80.345	374.700 ± 61.675	8.600 ± 3.382	1.200 ± 0.980	4.000 ± 2.049	0.600 ± 0.917	0.500 ± 0.671	0.400 ± 0.490	0.400 ± 0.490
Sim_0.7_train_0.7_A_number	10.600 ± 6.151	6.700 ± 3.226	7.600 ± 4.341	8.000 ± 6.245	8.300 ± 6.435	0	0	0	0	0	0	0
Sim_0.7_train_0.7_B_number	139.700 ± 29.793	122.000 ± 19.157	160.800 ± 26.555	160.300 ± 27.144	116.200 ± 30.426	0.400 ± 0.490	0.200 ± 0.600	0.200 ± 0.400	0	0.100 ± 0.300	0	0
Sim_0.7_train_0.7_C_number	110.200 ± 27.842	94.000 ± 22.009	119.200 ± 17.775	125.900 ± 36.190	86.800 ± 25.007	1.300 ± 1.418	0.100 ± 0.300	0.600 ± 0.663	0.100 ± 0.300	0	0	0.200 ± 0.400
Sim_0.7_train_0.7_D_number	116.900 ± 22.775	105.200 ± 19.072	129.600 ± 21.463	120.900 ± 24.143	86.200 ± 21.558	1.200 ± 1.400	0.400 ± 0.490	0.500 ± 0.671	0	0	0	0.100 ± 0.300
Sim_0.7_train_0.7_E_number	96.400 ± 19.226	96.800 ± 22.825	106.000 ± 22.751	100.500 ± 20.304	77.200 ± 22.529	5.700 ± 2.492	0.500 ± 0.671	2.700 ± 1.676	0.500 ± 0.671	0.400 ± 0.490	0.400 ± 0.490	0.100 ± 0.300
Sim_0.8	0.020 ± 0.002	0.019 ± 0.003	0.025 ± 0.003	0.023 ± 0.004	0.012 ± 0.003	0	0	0	0	0	0	0
Sim_0.8_number	609.800 ± 72.623	570.400 ± 83.730	748.000 ± 87.067	691.200 ± 126.962	352.500 ± 91.484	2.600 ± 2.375	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0.200 ± 0.400
Sim_0.8_A_number	6.300 ± 2.934	5.400 ± 2.289	6.300 ± 2.900	5.200 ± 2.358	4.300 ± 2.492	0	0	0	0	0	0	0
Sim_0.8_B_number	225.600 ± 41.057	195.100 ± 37.727	266.900 ± 37.822	252.100 ± 59.327	122.300 ± 37.283	0.200 ± 0.600	0	0	0	0	0	0

Sim_0.8_C_number	165.000 ± 39.749	158.100 ± 44.241	210.100 ± 46.105	192.300 ± 60.486	91.200 ± 36.633	0.500 ± 0.671	0	0	0	0	0	0.200 ± 0.400
Sim_0.8_D_number	135.100 ± 36.363	127.500 ± 29.090	163.200 ± 41.578	145.800 ± 38.070	77.600 ± 27.442	1.000 ± 1.612	0	0	0	0	0	0
Sim_0.8_E_number	77.800 ± 18.209	84.300 ± 19.152	101.500 ± 14.928	95.800 ± 19.245	57.100 ± 18.124	0.900 ± 1.221	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0
Sim_0.8_train_0.7	0.159 ± 0.036	0.159 ± 0.029	0.152 ± 0.023	0.152 ± 0.038	0.235 ± 0.042	0.638 ± 0.404	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0.200 ± 0.400
Sim_0.8_train_0.7_number	95.100 ± 16.628	89.600 ± 14.179	112.600 ± 14.759	103.400 ± 26.242	80.000 ± 13.107	1.500 ± 0.922	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0.200 ± 0.400
Sim_0.8_train_0.7_A_number	2.600 ± 2.458	2.200 ± 1.536	1.800 ± 1.077	2.000 ± 2.280	1.800 ± 1.939	0	0	0	0	0	0	0
Sim_0.8_train_0.7_B_number	29.600 ± 8.126	26.000 ± 8.283	35.400 ± 11.092	33.300 ± 11.568	24.900 ± 6.610	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.8_train_0.7_C_number	21.000 ± 8.866	20.600 ± 6.453	22.800 ± 4.643	23.000 ± 7.389	17.100 ± 5.356	0.300 ± 0.640	0	0	0	0	0	0.200 ± 0.400
Sim_0.8_train_0.7_D_number	22.800 ± 6.896	22.200 ± 5.896	29.100 ± 5.647	22.800 ± 7.305	18.800 ± 5.946	0.400 ± 0.490	0	0	0	0	0	0
Sim_0.8_train_0.7_E_number	19.100 ± 5.338	18.600 ± 3.137	23.500 ± 5.626	22.300 ± 5.311	17.400 ± 5.083	0.700 ± 0.781	0.200 ± 0.400	0.200 ± 0.400	0	0	0	0
Sim_0.9	0.005 ± 0.001	0.005 ± 0.001	0.006 ± 0.001	0.006 ± 0.001	0.003 ± 0.001	0	0	0	0	0	0	0
Sim_0.9_number	156.600 ± 29.179	139.900 ± 22.479	184.900 ± 23.420	166.800 ± 33.964	85.600 ± 28.566	1.200 ± 0.980	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.200 ± 0.400
Sim_0.9_A_number	1.800 ± 1.536	1.700 ± 1.345	2.200 ± 1.166	1.900 ± 1.375	1.500 ± 1.432	0	0	0	0	0	0	0

Sim_0.9_B_number	58.400 ± 11.706	46.500 ± 10.132	69.300 ± 11.568	64.200 ± 17.509	30.000 ± 13.364	0	0	0	0	0	0	0
Sim_0.9_C_number	45.100 ± 14.195	38.700 ± 8.810	49.400 ± 6.696	43.900 ± 13.989	22.300 ± 10.383	0.200 ± 0.400	0	0	0	0	0	0.200 ± 0.400
Sim_0.9_D_number	33.600 ± 9.583	33.900 ± 9.544	41.700 ± 10.668	34.500 ± 7.658	19.400 ± 6.453	0.400 ± 0.663	0	0	0	0	0	0
Sim_0.9_E_number	17.700 ± 5.533	19.100 ± 5.504	22.300 ± 6.404	22.300 ± 4.713	12.400 ± 5.064	0.600 ± 0.800	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0
Sim_0.9_train_0.7	0.157 ± 0.035	0.205 ± 0.033	0.169 ± 0.034	0.186 ± 0.057	0.237 ± 0.072	0.617 ± 0.435	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.200 ± 0.400
Sim_0.9_train_0.7_number	24.100 ± 5.147	28.800 ± 6.940	31.000 ± 6.017	30.200 ± 8.964	19.000 ± 5.727	1.000 ± 0.775	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0.200 ± 0.400
Sim_0.9_train_0.7_A_number	0.600 ± 0.917	0.700 ± 1.187	0.500 ± 0.671	0.700 ± 1.187	0.800 ± 1.249	0	0	0	0	0	0	0
Sim_0.9_train_0.7_B_number	6.500 ± 2.062	6.300 ± 2.492	9.000 ± 3.194	9.100 ± 4.253	3.900 ± 1.814	0	0	0	0	0	0	0
Sim_0.9_train_0.7_C_number	5.400 ± 2.375	7.500 ± 3.471	6.100 ± 2.844	7.000 ± 2.720	3.900 ± 2.022	0.200 ± 0.400	0	0	0	0	0	0.200 ± 0.400
Sim_0.9_train_0.7_D_number	5.900 ± 2.508	9.200 ± 3.709	7.800 ± 2.821	6.200 ± 3.156	4.600 ± 1.020	0.300 ± 0.458	0	0	0	0	0	0
Sim_0.9_train_0.7_E_number	5.700 ± 2.685	5.100 ± 3.176	7.600 ± 2.871	7.200 ± 1.720	5.800 ± 3.572	0.500 ± 0.671	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0

Table S7. The fine-tuning results on the VEGFR 10%-fine-tuning datasets with RDKit filtering.

VEGFR	CharRNN	AAE	VAE	Reinvent	ORGAN	GraphAF	RNNAttn			TransVAE		
							rand	high entropy	k high entropy	rand	high entropy	k high entropy
IntDiv	0.855 ± 0.003	0.863 ± 0.003	0.851 ± 0.004	0.853 ± 0.003	0.856 ± 0.003	0.894 ± 0.001	1	1	1	1	1	1
SNN/Gen_train	0.501 ± 0.026	0.500 ± 0.013	0.550 ± 0.018	0.511 ± 0.018	0.473 ± 0.013	0.256 ± 0.006	0.252 ± 0.006	0.310 ± 0.010	0.252 ± 0.006	0.241 ± 0.008	0.284 ± 0.008	0.240 ± 0.008
SNN/Gen_goal	0.470 ± 0.015	0.458 ± 0.007	0.494 ± 0.013	0.477 ± 0.012	0.439 ± 0.009	0.290 ± 0.003	0.285 ± 0.004	0.360 ± 0.005	0.286 ± 0.005	0.274 ± 0.008	0.325 ± 0.007	0.273 ± 0.008
IntDiv_ Rediscovery	0.813 ± 0.006	0.815 ± 0.009	0.812 ± 0.007	0.811 ± 0.007	0.791 ± 0.019	0.352 ± 0.063	/	/	/	/	/	/
SNN/Rediscovery_ train	0.775 ± 0.018	0.767 ± 0.016	0.770 ± 0.018	0.764 ± 0.020	0.761 ± 0.020	0.671 ± 0.044	/	0.306 ± 0.000	/	/	/	/
Rediscovery	0.003 ± 0.001	0.002 ± 0.000	0.003 ± 0.001	0.003 ± 0.001	0.002 ± 0.000	0	0	0	0	0	0	0
Rediscovery_ number	92.900 ± 15.978	74.800 ± 14.483	**97.400 ± 16.572**	90.000 ± 16.260	47.500 ± 9.922	0.900 ± 1.814	0	0.100 ± 0.300	0	0	0	0
Rediscovery_A_ number	0.400 ± 0.490	0.400 ± 0.490	0.400 ± 0.663	0.500 ± 0.500	0.100 ± 0.300	0	0	0	0	0	0	0
Rediscovery_B_ number	25.100 ± 4.592	16.300 ± 5.292	**25.300 ± 5.197**	25.200 ± 5.930	11.400 ± 4.055	0	0	0	0	0	0	0
Rediscovery_C_ number	33.900 ± 11.423	26.100 ± 10.894	35.800 ± 7.935	30.500 ± 8.958	17.600 ± 6.829	0.100 ± 0.300	0	0	0	0	0	0
Rediscovery_D_ number	20.500 ± 4.924	18.300 ± 4.627	22.000 ± 4.561	19.800 ± 5.231	11.100 ± 3.330	0.600 ± 1.200	0	0.100 ± 0.300	0	0	0	0

Rediscovery_E_number	13.000 ± 4.243	13.700 ± 3.900	13.900 ± 5.449	14.000 ± 4.050	7.300 ± 2.934	0.200 ± 0.400	0	0	0	0	0	0
Rediscovery_0.7	0.252 ± 0.078	0.288 ± 0.044	0.269 ± 0.058	0.283 ± 0.066	0.269 ± 0.097	0.217 ± 0.350	0	0.100 ± 0.300	0	0	0	0
Rediscovery_0.7_number	24.000 ± 9.737	21.500 ± 5.104	26.500 ± 8.201	25.600 ± 7.826	12.700 ± 5.100	0.600 ± 1.200	0	0.100 ± 0.300	0	0	0	0
Rediscovery_0.7_A_number	0	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0	0	0	0	0
Rediscovery_0.7_B_number	7.100 ± 4.571	4.700 ± 2.193	7.100 ± 4.346	7.600 ± 2.653	2.900 ± 2.300	0	0	0	0	0	0	0
Rediscovery_0.7_C_number	6.700 ± 2.722	6.000 ± 2.898	7.900 ± 3.048	6.400 ± 3.007	3.400 ± 2.728	0.100 ± 0.300	0	0	0	0	0	0
Rediscovery_0.7_D_number	5.200 ± 3.487	5.100 ± 2.166	6.100 ± 2.385	5.400 ± 3.169	2.900 ± 1.513	0.300 ± 0.640	0	0.100 ± 0.300	0	0	0	0
Rediscovery_0.7_E_number	5.000 ± 2.966	5.600 ± 2.835	5.300 ± 2.759	6.100 ± 3.646	3.500 ± 1.857	0.200 ± 0.400	0	0	0	0	0	0
Sim_0.7	0.084 ± 0.014	0.075 ± 0.008	0.098 ± 0.014	0.090 ± 0.013	0.044 ± 0.009	0.001 ± 0.000	0	0	0	0	0	0
Sim_0.7_number	2512.600 ± 410.561	2243.500 ± 249.008	2946.800 ± 418.307	2695.800 ± 393.457	1331.100 ± 262.491	19.500 ± 14.299	0.600 ± 0.800	5.000 ± 2.000	0.800 ± 0.872	0.500 ± 0.671	1.000 ± 0.775	0.300 ± 0.640
Sim_0.7_A_number	10.900 ± 8.631	10.200 ± 9.998	12.600 ± 10.082	10.900 ± 10.406	3.100 ± 2.700	0	0	0	0	0	0	0
Sim_0.7_B_number	554.700 ± 108.468	497.700 ± 111.232	652.500 ± 101.297	609.500 ± 111.608	255.000 ± 56.766	1.400 ± 1.020	0	0.400 ± 0.490	0	0	0	0
Sim_0.7_C_number	929.500 ± 264.471	797.100 ± 175.994	1095.600 ± 307.517	955.600 ± 257.867	500.800 ± 191.539	4.600 ± 4.271	0	0.200 ± 0.400	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0.300 ± 0.640

Sim_0.7_D_number	627.300 ± 189.377	539.200 ± 106.083	728.400 ± 163.833	684.800 ± 195.369	342.500 ± 106.856	6.600 ± 6.070	0	1.500 ± 1.118	0.300 ± 0.458	0	0	0
Sim_0.7_E_number	390.200 ± 66.068	399.300 ± 75.321	457.700 ± 96.218	435.000 ± 70.939	229.700 ± 71.599	6.900 ± 4.826	0.600 ± 0.800	2.900 ± 1.375	0.400 ± 0.490	0.400 ± 0.663	0.900 ± 0.700	0
Sim_0.7_train_0.7	0.192 ± 0.037	0.181 ± 0.031	0.181 ± 0.025	0.186 ± 0.040	0.263 ± 0.047	0.689 ± 0.153	0.400 ± 0.490	1.000 ± 0.000	0.500 ± 0.500	0.400 ± 0.490	0.700 ± 0.458	0.200 ± 0.400
Sim_0.7_train_0.7_number	471.400 ± 63.974	403.300 ± 65.774	530.600 ± 90.183	491.100 ± 80.248	343.000 ± 57.711	13.000 ± 9.940	0.600 ± 0.800	5.000 ± 2.000	0.800 ± 0.872	0.500 ± 0.671	1.000 ± 0.775	0.300 ± 0.640
Sim_0.7_train_0.7_A_number	4.400 ± 3.105	4.000 ± 4.266	5.400 ± 4.903	4.400 ± 3.499	2.000 ± 2.098	0	0	0	0	0	0	0
Sim_0.7_train_0.7_B_number	113.700 ± 19.678	96.400 ± 23.333	127.900 ± 29.039	129.000 ± 29.997	72.700 ± 25.116	0.900 ± 1.044	0	0.400 ± 0.490	0	0	0	0
Sim_0.7_train_0.7_C_number	158.900 ± 43.173	127.700 ± 39.110	180.400 ± 43.978	145.900 ± 32.862	111.500 ± 29.834	2.000 ± 2.324	0	0.200 ± 0.400	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0.300 ± 0.640
Sim_0.7_train_0.7_D_number	102.000 ± 22.365	77.400 ± 12.714	118.500 ± 28.083	108.500 ± 27.449	80.500 ± 25.610	4.400 ± 4.923	0	1.500 ± 1.118	0.300 ± 0.458	0	0	0
Sim_0.7_train_0.7_E_number	92.400 ± 23.290	97.800 ± 26.411	98.400 ± 25.641	103.300 ± 31.458	76.300 ± 29.042	5.700 ± 3.662	0.600 ± 0.800	2.900 ± 1.375	0.400 ± 0.490	0.400 ± 0.663	0.900 ± 0.700	0
Sim_0.8	0.018 ± 0.002	0.015 ± 0.002	0.021 ± 0.003	0.018 ± 0.002	0.009 ± 0.001	0	0	0	0	0	0	0
Sim_0.8_number	540.500 ± 66.587	451.800 ± 61.490	618.800 ± 86.284	550.700 ± 72.490	264.300 ± 39.865	2.800 ± 3.970	0.100 ± 0.300	0.200 ± 0.400	0.200 ± 0.400	0	0.100 ± 0.300	0
Sim_0.8_A_number	2.200 ± 2.272	2.300 ± 1.952	2.500 ± 1.857	2.200 ± 1.536	0.700 ± 1.269	0	0	0	0	0	0	0
Sim_0.8_B_number	129.900 ± 25.704	97.200 ± 20.371	141.400 ± 13.610	134.900 ± 20.530	57.800 ± 11.034	0.200 ± 0.400	0	0	0	0	0	0

Sim_0.8_C_number	203.800 ± 56.836	170.300 ± 33.490	233.800 ± 58.467	201.000 ± 39.540	105.600 ± 37.729	0.600 ± 0.800	0	0	0	0	0	0
Sim_0.8_D_number	128.900 ± 27.009	105.400 ± 28.444	150.700 ± 35.564	129.000 ± 29.840	62.700 ± 16.450	1.300 ± 2.283	0	0.100 ± 0.300	0.100 ± 0.300	0	0	0
Sim_0.8_E_number	75.700 ± 18.995	76.600 ± 15.259	90.400 ± 20.490	83.600 ± 17.031	37.500 ± 13.463	0.700 ± 1.269	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0
Sim_0.8_train_0.7	0.178 ± 0.035	0.182 ± 0.023	0.173 ± 0.031	0.182 ± 0.032	0.236 ± 0.052	0.289 ± 0.307	0.100 ± 0.300	0.200 ± 0.400	0.200 ± 0.400	0	0.100 ± 0.300	0
Sim_0.8_train_0.7_number	96.000 ± 19.437	82.500 ± 16.262	107.400 ± 25.613	100.600 ± 22.214	61.500 ± 14.066	1.600 ± 2.653	0.100 ± 0.300	0.200 ± 0.400	0.200 ± 0.400	0	0.100 ± 0.300	0
Sim_0.8_train_0.7_A_number	0.700 ± 1.005	1.000 ± 1.612	1.200 ± 1.166	1.000 ± 0.894	0.400 ± 0.663	0	0	0	0	0	0	0
Sim_0.8_train_0.7_B_number	26.300 ± 7.308	22.300 ± 6.017	26.800 ± 9.042	29.400 ± 7.940	14.000 ± 5.814	0.200 ± 0.400	0	0	0	0	0	0
Sim_0.8_train_0.7_C_number	30.300 ± 9.477	25.500 ± 9.636	35.000 ± 9.716	26.600 ± 6.606	21.300 ± 8.403	0.200 ± 0.400	0	0	0	0	0	0
Sim_0.8_train_0.7_D_number	20.500 ± 7.159	14.700 ± 4.496	24.000 ± 5.882	19.300 ± 7.550	14.100 ± 6.236	0.700 ± 1.487	0	0.100 ± 0.300	0.100 ± 0.300	0	0	0
Sim_0.8_train_0.7_E_number	18.200 ± 6.431	19.000 ± 5.814	20.400 ± 5.783	24.300 ± 9.809	11.700 ± 4.981	0.500 ± 1.025	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0	0.100 ± 0.300	0
Sim_0.9	0.005 ± 0.001	0.004 ± 0.001	0.005 ± 0.001	0.005 ± 0.001	0.002 ± 0.001	0	0	0	0	0	0	0
Sim_0.9_number	147.900 ± 20.964	114.400 ± 19.262	161.700 ± 26.575	143.500 ± 22.209	72.700 ± 16.100	1.100 ± 2.385	0	0.100 ± 0.300	0	0	0	0
Sim_0.9_A_number	0.400 ± 0.490	0.400 ± 0.490	0.400 ± 0.663	0.500 ± 0.500	0.200 ± 0.600	0	0	0	0	0	0	0

Sim_0.9_B_number	38.200 ± 7.139	23.700 ± 6.116	40.900 ± 5.629	38.200 ± 8.280	16.600 ± 5.122	0	0	0	0	0	0	0
Sim_0.9_C_number	56.000 ± 16.631	43.600 ± 13.720	60.400 ± 12.714	51.400 ± 10.052	27.700 ± 11.807	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.9_D_number	33.600 ± 8.499	27.600 ± 7.710	37.200 ± 12.295	31.900 ± 7.021	16.900 ± 5.467	0.700 ± 1.487	0	0.100 ± 0.300	0	0	0	0
Sim_0.9_E_number	19.700 ± 6.482	19.100 ± 5.108	22.800 ± 10.284	21.500 ± 6.249	11.300 ± 4.314	0.300 ± 0.640	0	0	0	0	0	0
Sim_0.9_train_0.7	0.196 ± 0.060	0.218 ± 0.037	0.202 ± 0.037	0.215 ± 0.050	0.249 ± 0.099	0.212 ± 0.345	0	0.100 ± 0.300	0	0	0	0
Sim_0.9_train_0.7_number	29.400 ± 10.735	24.800 ± 5.546	32.900 ± 8.960	31.100 ± 8.927	17.100 ± 5.924	0.700 ± 1.487	0	0.100 ± 0.300	0	0	0	0
Sim_0.9_train_0.7_A_number	0	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0.100 ± 0.300	0	0	0	0	0	0	0
Sim_0.9_train_0.7_B_number	8.200 ± 4.400	5.200 ± 2.638	8.200 ± 4.069	9.000 ± 3.493	4.000 ± 2.490	0	0	0	0	0	0	0
Sim_0.9_train_0.7_C_number	7.800 ± 3.370	7.300 ± 3.257	10.500 ± 3.294	7.800 ± 3.280	4.200 ± 2.891	0.100 ± 0.300	0	0	0	0	0	0
Sim_0.9_train_0.7_D_number	7.300 ± 3.689	5.700 ± 2.648	7.300 ± 3.164	6.200 ± 3.156	4.200 ± 2.676	0.300 ± 0.640	0	0.100 ± 0.300	0	0	0	0
Sim_0.9_train_0.7_E_number	6.100 ± 3.477	6.500 ± 2.941	6.800 ± 3.400	8.000 ± 3.975	4.600 ± 1.960	0.300 ± 0.640	0	0	0	0	0	0

Table S8. The fine-tuning results on the AR 10%-fine-tuning datasets with RDKit filtering.

AR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.860 ± 0.004	0.861 ± 0.007	0.854 ± 0.004	0.856 ± 0.003	0.855 ± 0.004
SNN/Gen_train	0.493 ± 0.025	0.500 ± 0.055	0.533 ± 0.016	0.516 ± 0.017	0.470 ± 0.019
SNN/Gen_goal	0.468 ± 0.014	0.468 ± 0.036	0.488 ± 0.011	0.484 ± 0.012	0.445 ± 0.012
IntDiv_Rediscovery	0.808 ± 0.009	0.810 ± 0.014	0.813 ± 0.007	0.808 ± 0.007	0.797 ± 0.020
SNN/Rediscovery_train	0.803 ± 0.022	0.777 ± 0.030	0.799 ± 0.015	0.793 ± 0.015	0.781 ± 0.022
Rediscovery	0.004 ± 0.000	0.003 ± 0.001	0.004 ± 0.001	0.003 ± 0.001	0.002 ± 0.000
Rediscovery_number	108.300 ± 11.731	82.200 ± 37.552	**114.400 ± 15.454**	99.400 ± 25.500	60.000 ± 11.207
Rediscovery_A_number	**1.700 ± 1.100**	0.800 ± 0.872	1.600 ± 1.200	**1.700 ± 1.345**	0.400 ± 0.490
Rediscovery_B_number	49.100 ± 8.757	34.100 ± 19.055	**51.400 ± 12.184**	45.500 ± 8.370	28.200 ± 4.833
Rediscovery_C_number	31.300 ± 6.230	25.700 ± 11.073	33.500 ± 5.971	28.500 ± 11.783	18.400 ± 4.200
Rediscovery_D_number	15.100 ± 4.763	12.700 ± 5.533	18.800 ± 6.554	14.100 ± 5.205	8.800 ± 4.400
Rediscovery_E_number	11.100 ± 4.460	8.900 ± 5.752	9.100 ± 4.700	9.600 ± 4.821	4.200 ± 2.182
Rediscovery_0.7	0.255 ± 0.074	0.325 ± 0.060	0.261 ± 0.062	0.273 ± 0.067	0.304 ± 0.072
Rediscovery_0.7_number	27.500 ± 8.041	24.900 ± 10.153	**30.200 ± 9.786**	27.600 ± 11.586	18.600 ± 6.484
Rediscovery_0.7_A_number	**1.500 ± 0.922**	0.800 ± 0.872	**1.500 ± 0.922**	**1.500 ± 1.118**	0.400 ± 0.490
Rediscovery_0.7_B_number	10.900 ± 3.448	8.500 ± 5.182	**12.000 ± 5.550**	9.700 ± 4.518	8.000 ± 2.646
Rediscovery_0.7_C_number	8.100 ± 4.482	7.500 ± 3.106	9.000 ± 3.578	8.900 ± 5.338	5.300 ± 1.418
Rediscovery_0.7_D_number	4.200 ± 1.661	5.200 ± 1.887	4.900 ± 1.814	4.400 ± 2.871	3.300 ± 2.369
Rediscovery_0.7_E_number	2.800 ± 1.833	2.900 ± 1.446	2.800 ± 1.990	3.100 ± 2.022	1.600 ± 1.114
Sim_0.7	0.079 ± 0.012	0.073 ± 0.032	0.088 ± 0.010	0.089 ± 0.014	0.046 ± 0.007
Sim_0.7_number	2383.800 ± 366.237	2182.200 ± 958.274	2630.500 ± 299.634	**2659.600 ± 407.090**	1380.400 ± 199.758
Sim_0.7_A_number	16.100 ± 9.148	16.100 ± 9.235	17.000 ± 9.960	17.400 ± 11.377	8.200 ± 5.582
Sim_0.7_B_number	1239.700 ± 355.929	1131.200 ± 608.121	1394.400 ± 326.484	1408.400 ± 382.496	706.400 ± 186.183
Sim_0.7_C_number	658.900 ± 117.426	578.700 ± 264.198	688.200 ± 114.678	698.300 ± 129.905	374.200 ± 94.084
Sim_0.7_D_number	290.900 ±	290.600 ±	341.000 ±	339.400 ±	186.500 ±

	38.806	113.460	50.218	65.672	44.032
Sim_0.7_E_number	178.200 ± 49.815	165.600 ± 98.191	189.900 ± 69.597	196.100 ± 70.317	105.100 ± 39.384
Sim_0.7_train_0.7	0.236 ± 0.044	0.293 ± 0.106	0.220 ± 0.030	0.230 ± 0.036	0.336 ± 0.062
Sim_0.7_train_0.7_number	553.300 ± 88.144	550.100 ± 187.976	571.000 ± 56.556	602.500 ± 80.240	455.600 ± 58.672
Sim_0.7_train_0.7_A_number	10.900 ± 6.188	13.800 ± 9.590	12.300 ± 8.186	12.700 ± 7.015	6.500 ± 4.129
Sim_0.7_train_0.7_B_number	226.300 ± 48.625	234.900 ± 103.575	235.700 ± 27.565	240.700 ± 37.720	183.200 ± 30.730
Sim_0.7_train_0.7_C_number	173.500 ± 31.123	152.600 ± 61.957	163.300 ± 15.146	178.300 ± 40.058	135.100 ± 36.037
Sim_0.7_train_0.7_D_number	97.600 ± 23.922	100.500 ± 36.081	111.500 ± 22.491	114.400 ± 32.343	86.800 ± 25.818
Sim_0.7_train_0.7_E_number	45.000 ± 15.511	48.300 ± 19.591	48.200 ± 18.643	56.400 ± 22.615	44.000 ± 21.804
Sim_0.8	0.021 ± 0.003	0.018 ± 0.009	0.023 ± 0.002	0.022 ± 0.004	0.011 ± 0.002
Sim_0.8_number	619.400 ± 92.901	536.800 ± 271.408	697.900 ± 66.203	658.300 ± 109.378	325.900 ± 56.732
Sim_0.8_A_number	5.300 ± 3.407	4.600 ± 3.292	6.200 ± 3.458	6.100 ± 4.158	2.400 ± 2.010
Sim_0.8_B_number	322.700 ± 94.417	281.000 ± 167.968	379.600 ± 88.284	350.000 ± 92.431	166.000 ± 45.646
Sim_0.8_C_number	167.700 ± 32.168	147.400 ± 78.158	181.000 ± 25.698	175.400 ± 45.060	92.500 ± 25.672
Sim_0.8_D_number	69.300 ± 12.570	61.800 ± 28.906	80.800 ± 19.600	76.800 ± 20.213	40.800 ± 14.211
Sim_0.8_E_number	54.400 ± 19.387	42.000 ± 24.548	50.300 ± 17.590	50.000 ± 14.717	24.200 ± 7.626
Sim_0.8_train_0.7	0.189 ± 0.028	0.256 ± 0.121	0.169 ± 0.030	0.179 ± 0.025	0.280 ± 0.059
Sim_0.8_train_0.7_number	115.800 ± 17.910	111.300 ± 47.206	117.100 ± 19.295	116.800 ± 21.623	88.800 ± 12.560
Sim_0.8_train_0.7_A_number	3.100 ± 1.814	3.800 ± 3.311	4.600 ± 3.499	4.600 ± 3.200	1.700 ± 1.187
Sim_0.8_train_0.7_B_number	49.200 ± 9.816	48.400 ± 28.168	49.100 ± 8.491	46.000 ± 8.683	38.900 ± 7.892
Sim_0.8_train_0.7_C_number	33.200 ± 6.337	29.600 ± 14.129	32.100 ± 6.549	35.400 ± 10.707	25.800 ± 5.016
Sim_0.8_train_0.7_D_number	19.900 ± 5.787	20.000 ± 6.957	21.600 ± 4.737	20.800 ± 6.210	13.800 ± 4.643
Sim_0.8_train_0.7_E_number	10.400 ± 4.821	9.500 ± 4.566	9.700 ± 5.041	10.000 ± 4.539	8.600 ± 4.005
Sim_0.9	0.007 ± 0.001	0.005 ± 0.003	0.007 ± 0.001	0.006 ± 0.001	0.003 ± 0.000
Sim_0.9_number	196.500 ± 30.117	161.800 ± 81.837	217.100 ± 32.922	194.600 ± 39.144	97.000 ± 14.248

Sim_0.9_A_number	3.100 ± 1.700	1.600 ± 1.281	3.900 ± 3.015	4.000 ± 2.569	1.700 ± 1.792
Sim_0.9_B_number	89.200 ± 21.577	73.700 ± 43.850	104.400 ± 26.238	94.400 ± 25.566	45.000 ± 12.853
Sim_0.9_C_number	57.000 ± 12.108	48.000 ± 24.731	60.500 ± 10.642	55.100 ± 14.618	30.200 ± 7.652
Sim_0.9_D_number	27.500 ± 9.244	22.600 ± 11.262	30.500 ± 8.936	25.100 ± 7.880	13.400 ± 5.783
Sim_0.9_E_number	19.700 ± 7.349	15.900 ± 12.062	17.800 ± 9.704	16.000 ± 7.912	6.700 ± 2.532
Sim_0.9_train_0.7	0.212 ± 0.047	0.274 ± 0.096	0.204 ± 0.059	0.204 ± 0.051	0.285 ± 0.061
Sim_0.9_train_0.7_number	40.900 ± 7.867	37.800 ± 14.455	44.200 ± 15.721	39.700 ± 13.726	27.400 ± 6.232
Sim_0.9_train_0.7_A_number	2.300 ± 1.100	1.200 ± 0.980	3.600 ± 2.973	3.400 ± 1.960	1.200 ± 1.327
Sim_0.9_train_0.7_B_number	15.300 ± 4.540	14.100 ± 7.635	15.300 ± 6.165	13.000 ± 5.292	10.200 ± 2.522
Sim_0.9_train_0.7_C_number	11.700 ± 4.880	11.000 ± 4.648	13.500 ± 5.005	12.200 ± 6.258	8.300 ± 2.052
Sim_0.9_train_0.7_D_number	6.800 ± 1.833	7.300 ± 2.452	7.200 ± 1.939	6.500 ± 3.138	5.400 ± 2.577
Sim_0.9_train_0.7_E_number	4.800 ± 2.482	4.200 ± 1.887	4.600 ± 3.262	4.600 ± 3.040	2.300 ± 1.100

Table S9. The fine-tuning results on the 5-HTR 10%-fine-tuning datasets with RDKit filtering.

5-HTR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.851 ± 0.004	0.855 ± 0.006	0.847 ± 0.007	0.846 ± 0.005	0.847 ± 0.005
SNN/Gen_train	0.565 ± 0.012	0.531 ± 0.031	0.572 ± 0.011	0.569 ± 0.010	0.523 ± 0.011
SNN/Gen_goal	0.526 ± 0.009	0.506 ± 0.021	0.528 ± 0.013	0.534 ± 0.008	0.499 ± 0.009
IntDiv_Rediscovery	0.803 ± 0.007	0.810 ± 0.015	0.806 ± 0.009	0.804 ± 0.008	0.801 ± 0.009
SNN/Rediscovery_train	0.809 ± 0.013	0.753 ± 0.028	0.797 ± 0.016	0.792 ± 0.010	0.788 ± 0.021
Rediscovery	0.010 ± 0.001	0.006 ± 0.002	0.009 ± 0.001	0.009 ± 0.001	0.006 ± 0.001
Rediscovery_number	299.700 ± 28.118	192.300 ± 57.228	257.400 ± 21.261	258.500 ± 30.342	181.300 ± 29.880
Rediscovery_A_number	0.700 ± 0.781	0.300 ± 0.458	0.500 ± 0.671	0.600 ± 0.490	0.400 ± 0.663
Rediscovery_B_number	135.700 ± 18.336	85.900 ± 30.051	112.200 ± 9.887	114.000 ± 15.395	80.000 ± 13.183
Rediscovery_C_number	99.800 ± 11.125	61.400 ± 21.082	86.600 ± 8.980	83.600 ± 10.229	59.600 ± 11.164
Rediscovery_D_number	50.300 ± 11.455	33.500 ± 10.112	42.600 ± 7.046	46.900 ± 12.405	31.900 ± 9.853
Rediscovery_E_number	13.200 ± 3.429	11.200 ± 4.167	15.500 ± 4.225	13.400 ± 2.728	9.400 ± 3.137
Rediscovery_0.7	0.268 ± 0.032	0.391 ± 0.076	0.293 ± 0.028	0.296 ± 0.025	0.309 ± 0.058
Rediscovery_0.7_number	80.500 ± 14.264	73.900 ± 21.380	75.500 ± 10.102	76.800 ± 12.844	57.100 ± 18.102
Rediscovery_0.7_A_number	0.400 ± 0.490	0.300 ± 0.458	0.300 ± 0.458	0.400 ± 0.490	0.200 ± 0.400
Rediscovery_0.7_B_number	35.000 ± 7.239	29.400 ± 10.210	28.100 ± 6.188	28.800 ± 6.508	21.800 ± 7.054
Rediscovery_0.7_C_number	26.200 ± 4.643	23.100 ± 8.549	25.200 ± 5.980	25.500 ± 3.612	19.400 ± 8.052
Rediscovery_0.7_D_number	13.800 ± 5.134	14.400 ± 4.883	14.600 ± 4.079	15.200 ± 4.534	11.700 ± 4.267
Rediscovery_0.7_E_number	5.100 ± 2.625	6.700 ± 3.348	7.300 ± 2.193	6.900 ± 3.048	4.000 ± 2.608
Sim_0.7	0.127 ± 0.011	0.092 ± 0.022	0.124 ± 0.013	0.130 ± 0.008	0.082 ± 0.008
Sim_0.7_number	3815.400 ± 317.804	2756.500 ± 653.001	3715.200 ± 389.191	3897.400 ± 253.045	2453.900 ± 244.850
Sim_0.7_A_number	8.000 ± 4.604	7.300 ± 6.357	8.100 ± 5.009	8.400 ± 4.499	5.500 ± 2.377
Sim_0.7_B_number	1756.600 ± 152.200	1286.300 ± 345.796	1694.000 ± 189.486	1790.400 ± 169.442	1086.200 ± 96.238
Sim_0.7_C_number	1220.800 ± 113.082	850.200 ± 229.445	1207.300 ± 140.791	1263.400 ± 97.735	803.900 ± 85.995
Sim_0.7_D_number	609.900 ± 107.240	450.600 ± 114.844	592.800 ± 111.559	616.600 ± 94.597	408.100 ± 85.764

Sim_0.7_E_number	220.100 ± 46.666	162.100 ± 50.268	213.000 ± 43.644	218.600 ± 38.790	150.200 ± 31.336
Sim_0.7_train_0.7	0.226 ± 0.020	0.302 ± 0.063	0.232 ± 0.014	0.239 ± 0.012	0.305 ± 0.031
Sim_0.7_train_0.7_number	858.300 ± 67.227	798.800 ± 93.433	861.300 ± 93.571	930.600 ± 63.880	750.100 ± 114.334
Sim_0.7_train_0.7_A_number	5.400 ± 2.417	4.400 ± 4.800	5.300 ± 4.562	4.200 ± 1.600	3.800 ± 2.713
Sim_0.7_train_0.7_B_number	384.700 ± 51.767	340.000 ± 66.491	365.100 ± 60.398	396.600 ± 46.265	308.000 ± 31.209
Sim_0.7_train_0.7_C_number	270.000 ± 19.955	266.200 ± 33.887	288.300 ± 30.725	312.200 ± 24.186	247.700 ± 41.442
Sim_0.7_train_0.7_D_number	139.900 ± 22.111	132.600 ± 36.876	141.300 ± 22.782	154.400 ± 28.664	138.800 ± 46.290
Sim_0.7_train_0.7_E_number	58.300 ± 14.677	55.600 ± 17.710	61.300 ± 15.113	63.200 ± 13.869	51.800 ± 14.858
Sim_0.8	0.039 ± 0.003	0.025 ± 0.007	0.037 ± 0.003	0.037 ± 0.003	0.024 ± 0.003
Sim_0.8_number	1163.000 ± 87.633	754.700 ± 200.661	1095.700 ± 93.599	1110.900 ± 81.510	716.200 ± 76.392
Sim_0.8_A_number	3.000 ± 1.732	1.500 ± 1.962	1.400 ± 1.200	2.800 ± 2.088	1.500 ± 1.360
Sim_0.8_B_number	539.800 ± 48.836	354.200 ± 112.098	492.700 ± 41.747	506.500 ± 43.348	314.500 ± 31.056
Sim_0.8_C_number	381.300 ± 40.787	235.600 ± 70.220	376.900 ± 48.057	375.100 ± 41.705	245.700 ± 33.562
Sim_0.8_D_number	180.200 ± 31.877	124.600 ± 33.847	166.300 ± 33.398	170.000 ± 39.782	116.100 ± 21.352
Sim_0.8_E_number	58.700 ± 18.363	38.800 ± 11.609	58.400 ± 17.523	56.500 ± 10.623	38.400 ± 12.249
Sim_0.8_train_0.7	0.190 ± 0.019	0.270 ± 0.081	0.196 ± 0.012	0.199 ± 0.021	0.248 ± 0.041
Sim_0.8_train_0.7_number	220.600 ± 28.479	195.600 ± 42.996	214.000 ± 20.189	221.100 ± 30.442	179.900 ± 44.734
Sim_0.8_train_0.7_A_number	2.000 ± 1.095	1.000 ± 1.549	0.700 ± 0.900	1.100 ± 0.831	1.000 ± 1.000
Sim_0.8_train_0.7_B_number	97.800 ± 15.445	80.200 ± 19.864	86.300 ± 10.565	92.000 ± 13.653	69.900 ± 15.978
Sim_0.8_train_0.7_C_number	67.800 ± 8.459	63.300 ± 16.255	72.700 ± 10.863	74.600 ± 12.492	63.400 ± 16.948
Sim_0.8_train_0.7_D_number	38.800 ± 9.887	36.900 ± 12.724	37.800 ± 7.718	37.800 ± 13.325	32.200 ± 14.260
Sim_0.8_train_0.7_E_number	14.200 ± 6.369	14.200 ± 5.564	16.500 ± 6.637	15.600 ± 4.758	13.400 ± 7.017
Sim_0.9	0.018 ± 0.002	0.011 ± 0.003	0.016 ± 0.002	0.016 ± 0.002	0.010 ± 0.001
Sim_0.9_number	526.700 ± 66.744	323.600 ± 96.719	481.400 ± 47.800	472.700 ± 52.202	311.200 ± 36.597
Sim_0.9_A_number	1.200 ± 0.872	0.600 ± 0.800	0.600 ± 0.663	0.700 ± 0.458	0.500 ± 0.806

Sim_0.9_B_number	242.700 ± 35.950	143.000 ± 53.149	215.400 ± 21.068	212.200 ± 26.713	139.600 ± 17.477
Sim_0.9_C_number	179.500 ± 32.515	106.800 ± 35.555	167.800 ± 22.418	160.600 ± 24.885	107.600 ± 20.309
Sim_0.9_D_number	82.100 ± 17.155	55.800 ± 14.183	72.800 ± 12.432	78.200 ± 18.093	49.400 ± 11.603
Sim_0.9_E_number	21.200 ± 6.911	17.400 ± 7.003	24.800 ± 7.820	21.000 ± 4.099	14.100 ± 3.360
Sim_0.9_train_0.7	0.208 ± 0.024	0.308 ± 0.073	0.219 ± 0.017	0.229 ± 0.025	0.263 ± 0.048
Sim_0.9_train_0.7_number	109.500 ± 19.320	96.800 ± 25.274	105.200 ± 12.632	108.100 ± 16.562	82.700 ± 21.991
Sim_0.9_train_0.7_A_number	0.800 ± 0.748	0.300 ± 0.458	0.300 ± 0.458	0.500 ± 0.500	0.300 ± 0.458
Sim_0.9_train_0.7_B_number	47.400 ± 9.468	37.500 ± 10.874	40.400 ± 8.237	41.400 ± 8.743	31.800 ± 10.255
Sim_0.9_train_0.7_C_number	35.900 ± 6.379	31.200 ± 10.815	36.700 ± 7.128	38.500 ± 6.233	30.100 ± 9.792
Sim_0.9_train_0.7_D_number	19.500 ± 6.281	20.000 ± 6.826	18.700 ± 5.515	19.300 ± 5.386	15.200 ± 5.134
Sim_0.9_train_0.7_E_number	5.900 ± 3.300	7.800 ± 3.628	9.100 ± 3.015	8.400 ± 2.973	5.300 ± 2.900

Table S10. The fine-tuning results on the DR 10%-fine-tuning datasets with RDKit filtering.

DR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.844 ± 0.003	0.842 ± 0.004	0.839 ± 0.004	0.834 ± 0.005	0.843 ± 0.004
SNN/Gen_train	0.550 ± 0.013	0.543 ± 0.015	0.570 ± 0.010	0.562 ± 0.015	0.504 ± 0.010
SNN/Gen_goal	0.520 ± 0.008	0.516 ± 0.017	0.530 ± 0.007	0.535 ± 0.009	0.484 ± 0.011
IntDiv_Rediscovery	0.795 ± 0.005	0.794 ± 0.006	0.790 ± 0.006	0.791 ± 0.007	0.784 ± 0.013
SNN/Rediscovery_train	0.789 ± 0.007	0.772 ± 0.013	0.781 ± 0.008	0.778 ± 0.011	0.767 ± 0.016
Rediscovery	0.008 ± 0.001	0.005 ± 0.002	0.008 ± 0.001	0.007 ± 0.000	0.004 ± 0.001
Rediscovery_number	227.900 ± 16.367	157.800 ± 63.234	**232.800 ± 26.728**	209.400 ± 14.151	129.700 ± 30.103
Rediscovery_A_number	1.400 ± 1.114	1.100 ± 0.831	**1.600 ± 0.663**	1.100 ± 0.831	1.100 ± 1.221
Rediscovery_B_number	101.700 ± 13.763	68.400 ± 29.540	**106.800 ± 17.674**	93.200 ± 8.750	60.000 ± 17.601
Rediscovery_C_number	66.300 ± 9.274	44.100 ± 17.552	65.700 ± 9.263	66.000 ± 11.323	34.800 ± 10.524
Rediscovery_D_number	39.800 ± 9.174	30.500 ± 14.080	41.000 ± 8.075	36.400 ± 4.363	22.900 ± 7.063
Rediscovery_E_number	18.700 ± 3.226	13.700 ± 5.967	17.700 ± 3.579	12.700 ± 1.900	10.900 ± 5.049
Rediscovery_0.7	0.275 ± 0.028	0.344 ± 0.071	0.317 ± 0.047	0.307 ± 0.036	0.340 ± 0.036
Rediscovery_0.7_number	62.500 ± 6.667	52.300 ± 19.267	**73.200 ± 9.938**	64.200 ± 7.718	44.200 ± 11.830
Rediscovery_0.7_A_number	0.700 ± 0.640	0.500 ± 0.500	**1.000 ± 1.000**	0.700 ± 0.640	0.800 ± 0.980
Rediscovery_0.7_B_number	27.000 ± 3.464	21.600 ± 7.552	**32.400 ± 4.779**	25.800 ± 5.706	18.600 ± 5.783
Rediscovery_0.7_C_number	17.700 ± 4.406	14.600 ± 5.333	19.600 ± 3.904	19.700 ± 4.691	12.700 ± 4.691
Rediscovery_0.7_D_number	12.000 ± 3.376	11.400 ± 6.359	14.000 ± 4.669	14.300 ± 3.494	8.700 ± 2.283
Rediscovery_0.7_E_number	5.100 ± 2.300	4.200 ± 1.939	6.200 ± 3.682	3.700 ± 1.100	3.400 ± 2.107
Sim_0.7	0.120 ± 0.014	0.100 ± 0.029	0.125 ± 0.013	0.126 ± 0.016	0.069 ± 0.014
Sim_0.7_number	3586.700 ± 405.591	3013.400 ± 857.735	3751.600 ± 392.995	**3774.700 ± 492.934**	2084.800 ± 431.788
Sim_0.7_A_number	11.300 ± 6.798	12.000 ± 6.033	12.800 ± 7.400	13.700 ± 9.540	8.800 ± 6.585
Sim_0.7_B_number	1773.800 ± 373.064	1440.300 ± 541.011	1883.000 ± 377.258	1938.500 ± 471.668	1055.400 ± 310.956
Sim_0.7_C_number	981.600 ± 89.293	810.100 ± 194.012	999.400 ± 88.768	1022.100 ± 93.326	545.800 ± 102.199
Sim_0.7_D_number	608.100 ± 64.230	547.300 ± 151.300	631.600 ± 70.670	602.000 ± 51.771	349.400 ± 56.013

Sim_0.7_E_number	211.900 ± 41.906	203.700 ± 62.584	224.900 ± 46.685	198.400 ± 33.248	125.400 ± 30.771
Sim_0.7_train_0.7	0.261 ± 0.025	0.290 ± 0.035	0.265 ± 0.024	0.278 ± 0.041	0.358 ± 0.027
Sim_0.7_train_0.7_number	930.100 ± 67.454	854.300 ± 201.838	988.800 ± 85.639	1034.300 ± 106.752	743.200 ± 141.137
Sim_0.7_train_0.7_A_number	5.700 ± 4.920	6.100 ± 2.625	6.700 ± 3.164	5.900 ± 2.468	6.000 ± 4.604
Sim_0.7_train_0.7_B_number	444.800 ± 65.221	366.400 ± 98.403	468.500 ± 70.669	505.000 ± 91.004	362.600 ± 76.697
Sim_0.7_train_0.7_C_number	263.300 ± 27.052	243.600 ± 70.786	278.500 ± 30.881	294.400 ± 51.521	200.000 ± 45.491
Sim_0.7_train_0.7_D_number	155.700 ± 15.265	170.000 ± 49.222	166.800 ± 19.312	171.300 ± 22.361	132.000 ± 28.189
Sim_0.7_train_0.7_E_number	60.600 ± 11.236	68.200 ± 18.862	68.300 ± 20.209	57.700 ± 11.082	42.600 ± 14.403
Sim_0.8	0.034 ± 0.002	0.026 ± 0.009	0.036 ± 0.003	0.034 ± 0.002	0.019 ± 0.004
Sim_0.8_number	1017.600 ± 62.450	782.300 ± 269.026	1070.200 ± 97.062	1011.600 ± 72.636	558.700 ± 107.393
Sim_0.8_A_number	3.800 ± 2.600	2.800 ± 1.939	4.500 ± 2.766	3.700 ± 3.466	3.300 ± 2.934
Sim_0.8_B_number	479.400 ± 60.810	369.800 ± 136.757	525.600 ± 64.293	496.900 ± 53.932	270.100 ± 62.641
Sim_0.8_C_number	296.400 ± 29.125	213.900 ± 66.129	294.900 ± 39.379	288.900 ± 25.403	155.000 ± 39.248
Sim_0.8_D_number	174.300 ± 24.389	137.600 ± 59.887	180.800 ± 32.056	166.400 ± 15.318	94.500 ± 14.445
Sim_0.8_E_number	63.700 ± 14.588	58.200 ± 22.063	64.400 ± 10.883	55.700 ± 11.296	35.800 ± 10.656
Sim_0.8_train_0.7	0.225 ± 0.019	0.249 ± 0.036	0.230 ± 0.019	0.234 ± 0.032	0.317 ± 0.037
Sim_0.8_train_0.7_number	228.000 ± 15.868	188.500 ± 54.879	245.900 ± 31.072	235.900 ± 28.560	176.200 ± 33.310
Sim_0.8_train_0.7_A_number	1.600 ± 1.356	1.300 ± 1.345	2.400 ± 1.855	1.700 ± 1.418	2.300 ± 2.193
Sim_0.8_train_0.7_B_number	98.500 ± 13.515	79.500 ± 20.319	113.700 ± 12.506	103.300 ± 19.945	80.100 ± 12.029
Sim_0.8_train_0.7_C_number	70.200 ± 10.235	55.100 ± 18.668	68.200 ± 11.418	67.800 ± 16.515	52.300 ± 17.493
Sim_0.8_train_0.7_D_number	41.100 ± 6.906	36.600 ± 15.743	46.100 ± 11.861	50.100 ± 6.625	31.300 ± 8.978
Sim_0.8_train_0.7_E_number	16.600 ± 5.817	16.000 ± 5.797	15.500 ± 5.390	13.000 ± 3.493	10.200 ± 4.956
Sim_0.9	0.014 ± 0.001	0.010 ± 0.004	0.014 ± 0.002	0.013 ± 0.001	0.007 ± 0.001
Sim_0.9_number	411.500 ± 28.504	293.200 ± 109.759	432.200 ± 50.519	380.800 ± 41.041	220.800 ± 42.572
Sim_0.9_A_number	2.400 ± 1.625	1.800 ± 1.166	3.000 ± 1.483	1.900 ± 1.044	2.000 ± 1.789

Sim_0.9_B_number	192.800 ± 20.148	135.200 ± 51.131	205.600 ± 21.974	183.400 ± 24.088	105.300 ± 23.753
Sim_0.9_C_number	120.500 ± 18.134	81.000 ± 26.952	122.400 ± 26.956	113.400 ± 17.107	61.100 ± 17.490
Sim_0.9_D_number	66.400 ± 12.330	51.800 ± 26.049	72.600 ± 15.461	62.200 ± 5.546	36.200 ± 8.704
Sim_0.9_E_number	29.400 ± 7.186	23.400 ± 10.298	28.600 ± 4.543	19.900 ± 4.369	16.200 ± 6.210
Sim_0.9_train_0.7	0.222 ± 0.021	0.269 ± 0.049	0.254 ± 0.027	0.248 ± 0.027	0.319 ± 0.037
Sim_0.9_train_0.7_number	91.400 ± 10.442	76.000 ± 25.072	109.600 ± 15.869	94.100 ± 12.692	71.000 ± 17.493
Sim_0.9_train_0.7_A_number	1.100 ± 1.221	0.800 ± 0.748	1.600 ± 1.200	1.200 ± 1.166	1.500 ± 1.360
Sim_0.9_train_0.7_B_number	40.700 ± 3.689	32.400 ± 9.178	51.300 ± 6.230	38.900 ± 9.823	29.600 ± 5.817
Sim_0.9_train_0.7_C_number	26.800 ± 6.226	21.200 ± 7.068	27.900 ± 6.935	28.000 ± 6.261	21.800 ± 8.340
Sim_0.9_train_0.7_D_number	15.600 ± 4.247	15.300 ± 8.533	20.600 ± 7.579	20.600 ± 3.904	12.800 ± 3.600
Sim_0.9_train_0.7_E_number	7.200 ± 3.458	6.300 ± 2.830	8.200 ± 4.045	5.400 ± 1.625	5.300 ± 3.689

Table S11. The fine-tuning results on the CDK 1%-fine-tuning datasets with RDKit filtering.

CDK	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.863 ± 0.009	0.864 ± 0.011	0.856 ± 0.014	0.861 ± 0.010	0.862 ± 0.009
SNN/Gen_train	0.348 ± 0.052	0.395 ± 0.048	0.419 ± 0.047	0.333 ± 0.049	0.346 ± 0.036
SNN/Gen_goal	0.378 ± 0.027	0.399 ± 0.028	0.416 ± 0.025	0.374 ± 0.024	0.363 ± 0.020
IntDiv_Rediscovery	0.725 ± 0.051	0.733 ± 0.047	0.739 ± 0.044	0.708 ± 0.058	0.652 ± 0.091
SNN/Rediscovery_train	0.734 ± 0.042	0.727 ± 0.039	0.728 ± 0.036	0.740 ± 0.043	0.710 ± 0.070
Rediscovery	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.000 ± 0.000
Rediscovery_number	28.820 ± 11.018	27.700 ± 10.065	**40.440 ± 13.528**	27.640 ± 11.435	13.320 ± 7.585
Rediscovery_A_number	0.090 ± 0.286	**0.180 ± 0.498**	0.160 ± 0.441	0.100 ± 0.300	0.080 ± 0.366
Rediscovery_B_number	3.820 ± 3.235	3.320 ± 2.881	**4.790 ± 4.134**	3.440 ± 3.471	1.470 ± 1.723
Rediscovery_C_number	12.540 ± 8.380	11.460 ± 7.104	16.550 ± 8.747	12.660 ± 8.251	5.440 ± 4.767
Rediscovery_D_number	7.960 ± 4.130	8.250 ± 3.887	12.020 ± 6.189	7.370 ± 3.820	3.910 ± 3.083
Rediscovery_E_number	4.410 ± 2.680	4.490 ± 2.851	6.920 ± 4.004	4.070 ± 2.933	2.420 ± 2.122
Rediscovery_0.7	0.400 ± 0.144	0.435 ± 0.138	0.422 ± 0.125	0.387 ± 0.152	0.488 ± 0.247
Rediscovery_0.7_number	11.220 ± 5.326	12.080 ± 5.700	**16.940 ± 7.182**	10.820 ± 6.249	6.610 ± 5.048
Rediscovery_0.7_A_number	0.010 ± 0.099	**0.080 ± 0.271**	0.050 ± 0.218	0.020 ± 0.140	0.030 ± 0.171
Rediscovery_0.7_B_number	1.540 ± 1.424	1.370 ± 1.222	**1.740 ± 1.383**	1.420 ± 1.218	0.690 ± 0.902
Rediscovery_0.7_C_number	5.050 ± 4.028	5.140 ± 4.145	7.240 ± 4.646	5.180 ± 4.276	2.970 ± 3.236
Rediscovery_0.7_D_number	2.900 ± 2.252	3.400 ± 2.227	4.610 ± 2.996	2.530 ± 1.873	1.810 ± 1.906
Rediscovery_0.7_E_number	1.720 ± 1.524	2.090 ± 1.795	3.300 ± 2.516	1.670 ± 1.721	1.110 ± 1.295
Sim_0.7	0.031 ± 0.014	0.041 ± 0.019	0.049 ± 0.018	0.030 ± 0.014	0.014 ± 0.008
Sim_0.7_number	935.180 ± 420.562	1236.570 ± 563.006	**1476.010 ± 553.351**	909.440 ± 424.861	430.420 ± 250.605
Sim_0.7_A_number	6.490 ± 30.206	11.430 ± 60.792	10.280 ± 44.479	9.890 ± 49.472	5.380 ± 31.423
Sim_0.7_B_number	124.370 ± 177.867	155.820 ± 207.164	184.040 ± 233.113	120.140 ± 192.571	46.470 ± 85.236
Sim_0.7_C_number	317.780 ± 198.251	409.380 ± 252.694	508.670 ± 260.339	313.490 ± 210.187	136.250 ± 98.004
Sim_0.7_D_number	314.280 ±	419.060 ±	502.680 ±	289.560 ±	153.800 ±

Sim_0.7_E_number	183.142 ± 147.269	246.067 ± 212.416	272.395 ± 228.998	165.820 ± 171.671	117.217 ± 109.369
Sim_0.7_train_0.7	0.273 ± 0.093	0.245 ± 0.076	0.275 ± 0.084	0.247 ± 0.090	0.392 ± 0.160
Sim_0.7_train_0.7_number	233.040 ± 88.346	287.180 ± 124.353	383.470 ± 133.998	209.350 ± 104.085	155.700 ± 92.095
Sim_0.7_train_0.7_A_number	0.650 ± 2.291	1.860 ± 5.552	1.660 ± 3.955	0.790 ± 2.696	0.570 ± 3.141
Sim_0.7_train_0.7_B_number	28.610 ± 24.376	33.680 ± 30.545	41.990 ± 33.311	23.510 ± 28.416	13.050 ± 18.634
Sim_0.7_train_0.7_C_number	88.480 ± 55.018	107.110 ± 66.974	149.910 ± 81.208	82.250 ± 59.083	55.550 ± 47.447
Sim_0.7_train_0.7_D_number	74.000 ± 34.646	89.470 ± 49.075	121.230 ± 59.399	65.170 ± 34.309	56.900 ± 49.175
Sim_0.7_train_0.7_E_number	41.300 ± 21.034	55.060 ± 32.237	68.680 ± 33.887	37.630 ± 25.858	29.630 ± 24.204
Sim_0.8	0.007 ± 0.004	0.009 ± 0.005	0.012 ± 0.006	0.007 ± 0.004	0.003 ± 0.003
Sim_0.8_number	224.360 ± 121.967	263.450 ± 145.181	357.650 ± 178.948	215.270 ± 115.912	94.540 ± 76.383
Sim_0.8_A_number	1.280 ± 4.658	1.840 ± 7.035	2.060 ± 7.597	1.700 ± 7.757	0.950 ± 4.904
Sim_0.8_B_number	29.460 ± 38.539	32.930 ± 41.816	44.140 ± 53.340	28.950 ± 41.668	10.560 ± 18.148
Sim_0.8_C_number	77.110 ± 49.912	88.390 ± 55.118	125.970 ± 70.972	79.760 ± 54.240	29.120 ± 24.090
Sim_0.8_D_number	81.600 ± 62.844	95.990 ± 75.771	128.620 ± 97.331	70.110 ± 47.581	35.620 ± 42.188
Sim_0.8_E_number	34.910 ± 42.989	44.300 ± 58.220	56.860 ± 75.951	34.750 ± 45.211	18.290 ± 35.324
Sim_0.8_train_0.7	0.243 ± 0.107	0.231 ± 0.101	0.249 ± 0.107	0.230 ± 0.111	0.347 ± 0.196
Sim_0.8_train_0.7_number	46.870 ± 20.807	53.300 ± 24.192	78.390 ± 32.439	44.160 ± 22.610	28.380 ± 20.384
Sim_0.8_train_0.7_A_number	0.200 ± 0.906	0.480 ± 1.507	0.430 ± 1.125	0.190 ± 0.484	0.230 ± 1.799
Sim_0.8_train_0.7_B_number	5.600 ± 5.678	6.210 ± 5.650	8.080 ± 7.247	4.990 ± 4.945	2.450 ± 3.330
Sim_0.8_train_0.7_C_number	19.180 ± 14.039	21.730 ± 14.105	33.560 ± 20.789	19.280 ± 14.661	10.990 ± 11.257
Sim_0.8_train_0.7_D_number	14.460 ± 8.211	16.140 ± 10.060	23.510 ± 13.248	13.150 ± 8.066	9.710 ± 10.220
Sim_0.8_train_0.7_E_number	7.430 ± 4.828	8.740 ± 5.624	12.810 ± 7.740	6.550 ± 4.951	5.000 ± 5.004
Sim_0.9	0.002 ± 0.001	0.002 ± 0.001	0.003 ± 0.001	0.002 ± 0.001	0.001 ± 0.000
Sim_0.9_number	53.240 ± 25.607	53.050 ± 24.058	77.340 ± 32.271	50.630 ± 22.797	23.190 ± 14.446

Sim_0.9_A_number	0.150 ± 0.497	0.320 ± 0.999	0.310 ± 0.902	0.130 ± 0.416	0.150 ± 0.712
Sim_0.9_B_number	7.570 ± 10.020	7.400 ± 9.122	10.540 ± 12.695	7.680 ± 11.062	2.860 ± 4.639
Sim_0.9_C_number	20.540 ± 13.407	19.840 ± 11.980	29.280 ± 16.310	20.370 ± 12.554	8.120 ± 6.526
Sim_0.9_D_number	16.500 ± 10.790	16.830 ± 10.298	24.170 ± 14.679	14.570 ± 8.087	7.360 ± 6.628
Sim_0.9_E_number	8.480 ± 6.768	8.660 ± 7.036	13.040 ± 10.098	7.880 ± 6.533	4.700 ± 5.560
Sim_0.9_train_0.7	0.308 ± 0.136	0.321 ± 0.125	0.322 ± 0.122	0.303 ± 0.142	0.415 ± 0.246
Sim_0.9_train_0.7_number	14.790 ± 6.930	16.020 ± 7.513	**23.110 ± 9.553**	14.450 ± 8.085	8.910 ± 6.782
Sim_0.9_train_0.7_A_number	0.020 ± 0.140	0.110 ± 0.371	0.110 ± 0.422	0.020 ± 0.140	0.050 ± 0.328
Sim_0.9_train_0.7_B_number	1.880 ± 1.716	1.830 ± 1.569	2.340 ± 1.840	1.780 ± 1.559	0.880 ± 1.219
Sim_0.9_train_0.7_C_number	6.110 ± 4.654	6.410 ± 4.936	9.490 ± 5.942	6.350 ± 5.271	3.560 ± 3.790
Sim_0.9_train_0.7_D_number	4.120 ± 3.185	4.760 ± 3.047	6.600 ± 3.942	3.790 ± 2.624	2.720 ± 2.815
Sim_0.9_train_0.7_E_number	2.660 ± 2.060	2.910 ± 2.200	4.570 ± 3.223	2.510 ± 2.443	1.700 ± 1.700

Table S12. The fine-tuning results on the EGFR 1%-fine-tuning datasets with RDKit filtering.

EGFR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.856 ± 0.017	0.857 ± 0.018	0.843 ± 0.022	0.857 ± 0.017	0.853 ± 0.014
SNN/Gen_train	0.357 ± 0.061	0.395 ± 0.049	0.436 ± 0.055	0.336 ± 0.058	0.354 ± 0.046
SNN/Gen_goal	0.386 ± 0.035	0.404 ± 0.031	0.434 ± 0.034	0.377 ± 0.032	0.376 ± 0.030
IntDiv_Rediscovery	0.663 ± 0.076	0.673 ± 0.061	0.688 ± 0.048	0.657 ± 0.084	0.580 ± 0.111
SNN/Rediscovery_train	0.723 ± 0.054	0.703 ± 0.056	0.710 ± 0.050	0.727 ± 0.062	0.677 ± 0.074
Rediscovery	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.000 ± 0.000
Rediscovery_number	23.650 ± 12.578	22.470 ± 11.798	**33.360 ± 14.974**	20.200 ± 10.963	10.380 ± 6.926
Rediscovery_A_number	0.100 ± 0.332	0.070 ± 0.292	**0.120 ± 0.354**	0.040 ± 0.196	0.050 ± 0.218
Rediscovery_B_number	6.280 ± 4.968	6.130 ± 4.372	**8.650 ± 6.059**	5.320 ± 4.249	2.550 ± 2.574
Rediscovery_C_number	8.160 ± 6.874	7.560 ± 5.487	11.820 ± 8.073	7.180 ± 6.180	3.640 ± 3.711
Rediscovery_D_number	4.480 ± 3.681	4.210 ± 3.213	6.190 ± 4.105	3.580 ± 3.163	1.720 ± 1.491
Rediscovery_E_number	4.630 ± 3.767	4.500 ± 3.791	6.580 ± 4.875	4.080 ± 3.155	2.420 ± 2.534
Rediscovery_0.7	0.434 ± 0.177	0.477 ± 0.183	0.466 ± 0.159	0.403 ± 0.206	0.523 ± 0.270
Rediscovery_0.7_number	10.570 ± 7.869	11.040 ± 7.639	**15.950 ± 9.780**	8.720 ± 7.050	6.020 ± 4.669
Rediscovery_0.7_A_number	0.020 ± 0.140	0.030 ± 0.171	**0.040 ± 0.196**	0.010 ± 0.099	0.030 ± 0.171
Rediscovery_0.7_B_number	2.930 ± 2.804	3.040 ± 2.596	**4.110 ± 3.677**	2.400 ± 2.600	1.550 ± 1.941
Rediscovery_0.7_C_number	3.080 ± 2.972	3.580 ± 3.030	5.490 ± 4.444	2.900 ± 3.071	2.180 ± 2.193
Rediscovery_0.7_D_number	2.310 ± 2.583	2.290 ± 2.160	3.000 ± 2.891	1.690 ± 2.185	0.980 ± 1.113
Rediscovery_0.7_E_number	2.230 ± 2.424	2.100 ± 2.119	3.310 ± 2.873	1.720 ± 1.524	1.280 ± 1.600
Sim_0.7	0.034 ± 0.018	0.044 ± 0.022	0.057 ± 0.022	0.033 ± 0.020	0.017 ± 0.011
Sim_0.7_number	1032.350 ± 543.647	1331.820 ± 661.001	**1705.640 ± 665.228**	1004.310 ± 586.360	524.570 ± 328.906
Sim_0.7_A_number	5.270 ± 17.687	4.470 ± 12.939	8.560 ± 21.977	4.860 ± 15.065	1.150 ± 3.427
Sim_0.7_B_number	309.280 ± 278.281	405.900 ± 398.987	509.140 ± 400.479	309.520 ± 355.006	136.280 ± 136.503
Sim_0.7_C_number	374.260 ± 303.647	449.810 ± 350.400	628.430 ± 405.665	364.770 ± 303.280	177.070 ± 185.371
Sim_0.7_D_number	167.960 ± 139.359	237.060 ± 187.724	280.330 ± 171.955	160.870 ± 135.804	87.820 ± 80.606

Sim_0.7_E_number	175.580 ± 144.283	234.580 ± 191.154	279.180 ± 175.927	164.290 ± 141.429	122.250 ± 134.840
Sim_0.7_train_0.7	0.289 ± 0.102	0.269 ± 0.096	0.314 ± 0.099	0.255 ± 0.109	0.446 ± 0.178
Sim_0.7_train_0.7_number	277.140 ± 146.518	345.200 ± 192.688	517.200 ± 226.272	238.640 ± 156.852	233.450 ± 174.892
Sim_0.7_train_0.7_A_number	1.260 ± 3.469	1.190 ± 2.810	3.500 ± 9.640	1.240 ± 3.353	0.780 ± 2.369
Sim_0.7_train_0.7_B_number	89.000 ± 70.930	110.650 ± 91.660	165.760 ± 105.380	79.970 ± 87.496	72.510 ± 92.636
Sim_0.7_train_0.7_C_number	92.450 ± 67.637	113.390 ± 84.111	181.070 ± 103.758	82.430 ± 64.477	74.860 ± 74.613
Sim_0.7_train_0.7_D_number	41.050 ± 23.875	57.060 ± 48.054	77.350 ± 42.493	33.030 ± 23.599	38.220 ± 39.121
Sim_0.7_train_0.7_E_number	53.380 ± 32.648	62.910 ± 46.953	89.520 ± 46.767	41.970 ± 26.609	47.080 ± 39.932
Sim_0.8	0.007 ± 0.004	0.008 ± 0.005	0.011 ± 0.005	0.007 ± 0.004	0.003 ± 0.002
Sim_0.8_number	210.760 ± 114.156	246.840 ± 144.622	344.810 ± 155.612	198.400 ± 115.345	93.920 ± 69.094
Sim_0.8_A_number	1.070 ± 3.938	1.240 ± 3.834	1.880 ± 4.934	1.090 ± 3.952	0.210 ± 0.668
Sim_0.8_B_number	61.930 ± 62.573	71.690 ± 83.853	100.530 ± 87.553	59.600 ± 67.812	24.220 ± 28.541
Sim_0.8_C_number	78.900 ± 62.362	86.030 ± 73.589	125.000 ± 87.658	72.180 ± 58.939	30.840 ± 35.293
Sim_0.8_D_number	33.070 ± 30.730	40.280 ± 39.223	57.260 ± 45.152	31.250 ± 32.459	14.850 ± 14.630
Sim_0.8_E_number	35.790 ± 29.659	47.600 ± 45.540	60.140 ± 43.990	34.280 ± 29.708	23.800 ± 30.680
Sim_0.8_train_0.7	0.273 ± 0.117	0.254 ± 0.118	0.285 ± 0.109	0.243 ± 0.125	0.422 ± 0.197
Sim_0.8_train_0.7_number	52.700 ± 29.575	54.960 ± 30.777	93.510 ± 46.025	43.890 ± 27.561	38.220 ± 32.431
Sim_0.8_train_0.7_A_number	0.180 ± 0.623	0.330 ± 1.040	0.640 ± 2.133	0.210 ± 0.791	0.120 ± 0.475
Sim_0.8_train_0.7_B_number	16.490 ± 13.214	16.530 ± 14.148	30.560 ± 21.791	13.950 ± 12.479	12.410 ± 19.096
Sim_0.8_train_0.7_C_number	17.500 ± 13.789	18.130 ± 12.899	31.680 ± 21.916	15.180 ± 12.114	12.020 ± 14.272
Sim_0.8_train_0.7_D_number	7.660 ± 5.818	8.270 ± 6.524	13.790 ± 9.399	6.110 ± 5.706	5.860 ± 7.969
Sim_0.8_train_0.7_E_number	10.870 ± 8.905	11.700 ± 9.821	16.840 ± 10.533	8.440 ± 6.080	7.810 ± 8.094
Sim_0.9	0.001 ± 0.001	0.001 ± 0.001	0.002 ± 0.001	0.001 ± 0.001	0.001 ± 0.000
Sim_0.9_number	40.750 ± 22.193	39.980 ± 21.509	61.720 ± 28.736	36.610 ± 20.822	17.400 ± 12.229
Sim_0.9_A_number	0.150 ± 0.477	0.210 ± 0.725	0.300 ± 0.933	0.180 ± 0.726	0.080 ± 0.306

Sim_0.9_B_number	10.730 ± 8.633	10.400 ± 7.846	16.350 ± 12.506	9.080 ± 7.400	4.170 ± 4.236
Sim_0.9_C_number	15.650 ± 15.071	14.510 ± 12.064	22.950 ± 17.716	14.240 ± 13.629	6.240 ± 7.257
Sim_0.9_D_number	6.980 ± 5.466	6.980 ± 6.015	10.820 ± 7.497	5.960 ± 5.342	2.780 ± 2.693
Sim_0.9_E_number	7.240 ± 6.030	7.880 ± 6.998	11.300 ± 8.271	7.150 ± 6.118	4.130 ± 4.507
Sim_0.9_train_0.7	0.345 ± 0.157	0.371 ± 0.175	0.366 ± 0.150	0.314 ± 0.182	0.428 ± 0.250
Sim_0.9_train_0.7_number	13.540 ± 9.209	14.070 ± 9.286	22.170 ± 13.018	11.430 ± 8.824	8.020 ± 6.542
Sim_0.9_train_0.7_A_number	0.050 ± 0.260	0.100 ± 0.458	0.150 ± 0.589	0.060 ± 0.276	0.040 ± 0.196
Sim_0.9_train_0.7_B_number	3.890 ± 3.647	4.120 ± 3.418	6.490 ± 5.754	3.370 ± 3.596	2.170 ± 2.926
Sim_0.9_train_0.7_C_number	4.270 ± 3.757	4.640 ± 3.913	7.690 ± 6.071	3.840 ± 3.770	2.880 ± 3.157
Sim_0.9_train_0.7_D_number	2.630 ± 2.862	2.670 ± 2.362	3.850 ± 3.454	2.020 ± 2.478	1.320 ± 1.690
Sim_0.9_train_0.7_E_number	2.700 ± 2.834	2.540 ± 2.343	3.990 ± 3.217	2.140 ± 1.913	1.610 ± 2.004

Table S13. The fine-tuning results on the JakA 1%-fine-tuning datasets with RDKit filtering.

JakA	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.856 ± 0.007	0.858 ± 0.010	0.849 ± 0.010	0.858 ± 0.009	0.857 ± 0.007
SNN/Gen_train	0.386 ± 0.046	0.437 ± 0.043	0.458 ± 0.039	0.365 ± 0.047	0.375 ± 0.030
SNN/Gen_goal	0.414 ± 0.030	0.442 ± 0.030	0.457 ± 0.026	0.401 ± 0.031	0.389 ± 0.021
IntDiv_Rediscovery	0.733 ± 0.036	0.739 ± 0.040	0.745 ± 0.035	0.721 ± 0.043	0.673 ± 0.097
SNN/Rediscovery_train	0.773 ± 0.030	0.764 ± 0.034	0.771 ± 0.028	0.778 ± 0.033	0.766 ± 0.047
Rediscovery	0.002 ± 0.001	0.002 ± 0.001	0.002 ± 0.001	0.002 ± 0.001	0.001 ± 0.000
Rediscovery_number	59.070 ± 23.090	56.160 ± 19.791	**71.610 ± 24.143**	49.760 ± 21.929	22.800 ± 12.407
Rediscovery_A_number	0.090 ± 0.319	0.080 ± 0.271	**0.100 ± 0.332**	0.050 ± 0.218	0.020 ± 0.140
Rediscovery_B_number	33.720 ± 14.471	30.760 ± 13.418	**41.160 ± 16.664**	28.340 ± 13.232	13.150 ± 8.595
Rediscovery_C_number	12.240 ± 9.475	13.440 ± 9.331	14.980 ± 9.806	11.150 ± 11.356	4.720 ± 4.739
Rediscovery_D_number	10.510 ± 7.513	9.220 ± 5.890	12.210 ± 8.496	7.950 ± 6.412	3.740 ± 3.384
Rediscovery_E_number	2.510 ± 2.978	2.660 ± 2.585	3.160 ± 3.492	2.270 ± 2.935	1.170 ± 1.738
Rediscovery_0.7	0.291 ± 0.096	0.319 ± 0.114	0.291 ± 0.099	0.271 ± 0.124	0.317 ± 0.173
Rediscovery_0.7_number	17.230 ± 9.872	18.230 ± 10.562	**21.050 ± 11.258**	13.780 ± 10.662	7.130 ± 5.471
Rediscovery_0.7_A_number	0.030 ± 0.171	0.030 ± 0.171	**0.030 ± 0.171**	0	0
Rediscovery_0.7_B_number	7.580 ± 4.966	7.360 ± 4.487	**9.710 ± 5.656**	6.020 ± 4.395	3.360 ± 2.897
Rediscovery_0.7_C_number	4.760 ± 4.574	6.200 ± 5.676	5.980 ± 5.107	4.460 ± 6.210	1.950 ± 2.621
Rediscovery_0.7_D_number	3.950 ± 3.067	3.710 ± 2.920	4.250 ± 3.392	2.450 ± 2.312	1.300 ± 1.735
Rediscovery_0.7_E_number	0.910 ± 1.078	0.930 ± 1.227	1.080 ± 1.419	0.850 ± 1.043	0.520 ± 0.741
Sim_0.7	0.080 ± 0.034	0.104 ± 0.038	0.113 ± 0.039	0.080 ± 0.042	0.036 ± 0.022
Sim_0.7_number	2396.280 ± 1029.843	3120.600 ± 1130.738	**3376.930 ± 1168.709**	2409.150 ± 1257.011	1094.180 ± 655.182
Sim_0.7_A_number	3.250 ± 4.046	2.050 ± 2.783	3.420 ± 5.026	1.740 ± 2.432	1.450 ± 3.031
Sim_0.7_B_number	1538.350 ± 900.823	2044.310 ± 1069.563	2165.570 ± 1114.809	1608.210 ± 1173.603	687.030 ± 561.939
Sim_0.7_C_number	411.220 ± 239.749	550.950 ± 278.095	580.600 ± 274.335	400.750 ± 261.765	193.410 ± 127.515
Sim_0.7_D_number	341.790 ± 320.558	401.210 ± 345.307	475.680 ± 432.845	305.640 ± 336.249	160.660 ± 159.757

Sim_0.7_E_number	101.670 ± 96.313	122.080 ± 116.161	151.660 ± 136.421	92.810 ± 108.763	51.630 ± 56.115
Sim_0.7_train_0.7	0.247 ± 0.083	0.244 ± 0.078	0.235 ± 0.079	0.219 ± 0.089	0.320 ± 0.125
Sim_0.7_train_0.7_number	545.860 ± 211.245	726.080 ± 280.762	753.310 ± 277.294	487.410 ± 278.557	307.690 ± 157.814
Sim_0.7_train_0.7_A_number	2.720 ± 3.507	1.660 ± 2.463	2.640 ± 3.764	1.240 ± 1.569	1.140 ± 2.554
Sim_0.7_train_0.7_B_number	310.890 ± 143.790	429.660 ± 209.707	421.010 ± 179.023	293.720 ± 196.606	167.470 ± 104.495
Sim_0.7_train_0.7_C_number	109.610 ± 77.330	163.210 ± 114.242	167.160 ± 108.871	106.540 ± 110.210	63.040 ± 44.025
Sim_0.7_train_0.7_D_number	93.930 ± 69.546	102.440 ± 74.998	126.160 ± 95.486	67.120 ± 59.325	61.870 ± 66.626
Sim_0.7_train_0.7_E_number	28.710 ± 21.567	29.110 ± 21.240	36.340 ± 28.655	18.790 ± 16.109	14.170 ± 12.499
Sim_0.8	0.021 ± 0.011	0.025 ± 0.012	0.029 ± 0.013	0.021 ± 0.013	0.009 ± 0.006
Sim_0.8_number	635.370 ± 325.365	748.010 ± 358.674	875.100 ± 389.819	631.950 ± 389.698	262.600 ± 186.433
Sim_0.8_A_number	0.400 ± 0.837	0.290 ± 0.621	0.440 ± 0.952	0.340 ± 1.041	0.090 ± 0.708
Sim_0.8_B_number	426.550 ± 298.725	520.160 ± 349.999	591.810 ± 371.166	444.580 ± 377.173	177.360 ± 170.860
Sim_0.8_C_number	97.090 ± 52.863	113.320 ± 60.051	131.820 ± 62.789	89.090 ± 54.002	40.290 ± 29.287
Sim_0.8_D_number	88.240 ± 82.358	88.440 ± 81.152	118.580 ± 114.746	77.180 ± 82.971	34.480 ± 36.834
Sim_0.8_E_number	23.090 ± 26.273	25.800 ± 30.352	32.450 ± 33.878	20.760 ± 26.808	10.380 ± 12.694
Sim_0.8_train_0.7	0.194 ± 0.085	0.189 ± 0.080	0.180 ± 0.080	0.171 ± 0.093	0.231 ± 0.122
Sim_0.8_train_0.7_number	109.080 ± 49.311	126.330 ± 53.712	142.310 ± 59.951	93.280 ± 59.543	51.110 ± 31.926
Sim_0.8_train_0.7_A_number	0.220 ± 0.460	0.130 ± 0.365	0.250 ± 0.555	0.170 ± 0.549	0.020 ± 0.140
Sim_0.8_train_0.7_B_number	61.150 ± 34.063	72.900 ± 38.046	77.390 ± 36.874	54.080 ± 39.975	28.000 ± 20.366
Sim_0.8_train_0.7_C_number	24.930 ± 17.874	31.740 ± 24.034	34.990 ± 21.474	22.970 ± 23.563	11.180 ± 9.543
Sim_0.8_train_0.7_D_number	17.290 ± 12.908	16.990 ± 12.431	23.200 ± 20.617	12.230 ± 11.271	9.550 ± 12.019
Sim_0.8_train_0.7_E_number	5.490 ± 5.065	4.570 ± 4.150	6.480 ± 6.039	3.830 ± 4.040	2.360 ± 2.492
Sim_0.9	0.005 ± 0.002	0.005 ± 0.002	0.006 ± 0.002	0.004 ± 0.002	0.002 ± 0.001
Sim_0.9_number	143.180 ± 61.954	139.520 ± 52.581	179.670 ± 66.893	127.440 ± 55.575	59.420 ± 44.137
Sim_0.9_A_number	0.110 ± 0.343	0.110 ± 0.313	0.160 ± 0.463	0.070 ± 0.255	0.010 ± 0.099

Sim_0.9_B_number	87.020 ± 47.474	85.040 ± 44.900	111.830 ± 57.396	80.860 ± 47.474	38.010 ± 40.042
Sim_0.9_C_number	28.150 ± 19.471	29.090 ± 17.683	34.910 ± 19.889	24.350 ± 18.012	10.700 ± 8.395
Sim_0.9_D_number	21.650 ± 16.005	19.460 ± 13.338	25.410 ± 18.552	17.150 ± 13.464	8.170 ± 7.309
Sim_0.9_E_number	6.250 ± 8.783	5.820 ± 7.115	7.360 ± 8.908	5.010 ± 7.257	2.530 ± 3.969
Sim_0.9_train_0.7	0.204 ± 0.084	0.221 ± 0.094	0.208 ± 0.088	0.190 ± 0.104	0.254 ± 0.145
Sim_0.9_train_0.7_number	27.780 ± 15.407	29.580 ± 15.392	35.690 ± 18.539	22.770 ± 16.162	13.170 ± 9.321
Sim_0.9_train_0.7_A_number	0.070 ± 0.255	0.050 ± 0.218	0.070 ± 0.292	0.030 ± 0.171	0
Sim_0.9_train_0.7_B_number	12.960 ± 8.692	13.630 ± 7.761	17.470 ± 10.325	10.990 ± 8.664	6.440 ± 5.281
Sim_0.9_train_0.7_C_number	8.130 ± 7.383	9.630 ± 7.807	10.360 ± 7.576	6.980 ± 7.773	3.470 ± 3.618
Sim_0.9_train_0.7_D_number	5.150 ± 3.930	4.870 ± 3.918	6.210 ± 5.113	3.590 ± 3.622	2.530 ± 3.463
Sim_0.9_train_0.7_E_number	1.470 ± 1.706	1.400 ± 1.726	1.580 ± 1.856	1.180 ± 1.452	0.730 ± 0.958

Table S14. The fine-tuning results on the PDGFR 1%-fine-tuning datasets with RDKit filtering.

PDGFR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.857 ± 0.009	0.859 ± 0.009	0.847 ± 0.011	0.854 ± 0.008	0.854 ± 0.007
SNN/Gen_train	0.359 ± 0.048	0.400 ± 0.042	0.435 ± 0.041	0.347 ± 0.041	0.358 ± 0.028
SNN/Gen_goal	0.389 ± 0.028	0.406 ± 0.027	0.429 ± 0.027	0.387 ± 0.024	0.377 ± 0.019
IntDiv_Rediscovery	0.704 ± 0.057	0.728 ± 0.043	0.731 ± 0.042	0.696 ± 0.055	0.615 ± 0.097
SNN/Rediscovery_train	0.756 ± 0.048	0.745 ± 0.045	0.751 ± 0.035	0.766 ± 0.041	0.736 ± 0.064
Rediscovery	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.000 ± 0.000
Rediscovery_number	21.210 ± 9.820	21.980 ± 8.962	31.470 ± 13.379	20.460 ± 10.320	9.390 ± 5.616
Rediscovery_A_number	0.110 ± 0.313	0.180 ± 0.409	0.250 ± 0.456	0.140 ± 0.347	0.070 ± 0.255
Rediscovery_B_number	6.350 ± 5.636	6.670 ± 5.162	10.020 ± 7.201	6.970 ± 5.733	2.680 ± 2.576
Rediscovery_C_number	5.670 ± 3.704	5.450 ± 3.103	8.440 ± 5.452	5.250 ± 3.505	2.540 ± 2.443
Rediscovery_D_number	4.960 ± 3.916	5.580 ± 3.943	7.400 ± 5.292	4.490 ± 3.651	2.140 ± 2.030
Rediscovery_E_number	4.120 ± 2.947	4.100 ± 2.689	5.360 ± 3.239	3.610 ± 3.079	1.960 ± 2.064
Rediscovery_0.7	0.316 ± 0.165	0.356 ± 0.175	0.313 ± 0.129	0.275 ± 0.148	0.365 ± 0.216
Rediscovery_0.7_number	6.600 ± 4.268	7.730 ± 4.445	9.790 ± 5.415	5.770 ± 4.345	3.510 ± 2.632
Rediscovery_0.7_A_number	0.020 ± 0.140	0.020 ± 0.140	0.030 ± 0.171	0.010 ± 0.099	0.010 ± 0.099
Rediscovery_0.7_B_number	1.420 ± 1.662	1.750 ± 1.997	2.490 ± 2.540	1.620 ± 2.048	0.830 ± 1.114
Rediscovery_0.7_C_number	1.430 ± 1.444	1.650 ± 1.602	2.140 ± 2.177	1.320 ± 1.476	0.820 ± 1.135
Rediscovery_0.7_D_number	1.880 ± 2.011	2.280 ± 2.108	2.610 ± 2.073	1.310 ± 1.440	0.870 ± 1.064
Rediscovery_0.7_E_number	1.850 ± 1.841	2.030 ± 1.931	2.520 ± 2.343	1.510 ± 1.873	0.980 ± 1.273
Sim_0.7	0.033 ± 0.019	0.040 ± 0.019	0.054 ± 0.026	0.034 ± 0.022	0.015 ± 0.009
Sim_0.7_number	988.750 ± 584.568	1211.120 ± 581.171	1621.050 ± 782.743	1015.730 ± 656.698	451.070 ± 284.393
Sim_0.7_A_number	6.700 ± 13.680	9.560 ± 17.979	13.280 ± 25.459	6.300 ± 12.916	3.360 ± 11.127
Sim_0.7_B_number	329.680 ± 365.185	387.760 ± 329.640	558.490 ± 471.870	348.870 ± 400.915	137.620 ± 158.048
Sim_0.7_C_number	276.600 ± 242.495	348.670 ± 272.580	455.100 ± 347.595	282.760 ± 313.526	133.060 ± 125.342
Sim_0.7_D_number	214.860 ±	258.510 ±	346.260 ±	215.960 ±	101.480 ±

Sim_0.7_E_number	174.800 160.910 ± 133.193	193.079 206.620 ± 169.618	257.334 247.920 ± 185.033	181.160 161.840 ± 145.218	110.885 75.550 ± 63.927
Sim_0.7_train_0.7	0.239 ± 0.084	0.224 ± 0.075	0.241 ± 0.084	0.207 ± 0.086	0.377 ± 0.153
Sim_0.7_train_0.7_number	206.960 ± 82.628	248.830 ± 98.644	358.120 ± 139.569	180.590 ± 79.697	157.270 ± 108.084
Sim_0.7_train_0.7_A_number	1.560 ± 3.790	2.600 ± 7.386	3.240 ± 7.342	1.230 ± 4.005	1.400 ± 6.744
Sim_0.7_train_0.7_B_number	63.290 ± 47.033	75.360 ± 46.249	118.350 ± 68.259	60.860 ± 47.765	44.850 ± 45.543
Sim_0.7_train_0.7_C_number	49.590 ± 29.483	60.700 ± 35.697	86.130 ± 53.797	42.200 ± 31.309	43.020 ± 41.706
Sim_0.7_train_0.7_D_number	49.950 ± 29.117	57.390 ± 30.440	83.990 ± 55.178	40.780 ± 24.225	36.940 ± 56.024
Sim_0.7_train_0.7_E_number	42.570 ± 22.935	52.780 ± 35.115	66.410 ± 39.581	35.520 ± 22.327	31.060 ± 29.389
Sim_0.8	0.007 ± 0.004	0.008 ± 0.004	0.011 ± 0.006	0.007 ± 0.005	0.003 ± 0.002
Sim_0.8_number	205.860 ± 121.628	230.100 ± 124.246	342.160 ± 176.553	208.240 ± 138.317	85.620 ± 62.957
Sim_0.8_A_number	1.320 ± 3.671	1.890 ± 5.167	2.900 ± 8.900	1.350 ± 4.080	0.560 ± 1.608
Sim_0.8_B_number	69.940 ± 70.642	77.140 ± 66.287	123.830 ± 102.463	74.910 ± 76.996	26.450 ± 31.320
Sim_0.8_C_number	57.180 ± 46.410	63.140 ± 48.524	92.710 ± 68.846	56.170 ± 57.350	26.070 ± 29.667
Sim_0.8_D_number	46.360 ± 47.381	51.430 ± 52.012	74.950 ± 69.039	45.690 ± 50.771	18.830 ± 22.904
Sim_0.8_E_number	31.060 ± 28.850	36.500 ± 32.795	47.770 ± 39.948	30.120 ± 32.150	13.710 ± 14.967
Sim_0.8_train_0.7	0.203 ± 0.089	0.209 ± 0.090	0.204 ± 0.080	0.179 ± 0.091	0.314 ± 0.163
Sim_0.8_train_0.7_number	36.990 ± 17.630	43.020 ± 18.439	63.630 ± 27.150	32.530 ± 17.193	23.100 ± 15.256
Sim_0.8_train_0.7_A_number	0.210 ± 0.697	0.390 ± 1.248	0.430 ± 1.194	0.140 ± 0.425	0.110 ± 0.397
Sim_0.8_train_0.7_B_number	11.210 ± 8.896	13.060 ± 9.957	20.650 ± 14.907	11.660 ± 11.156	6.980 ± 7.570
Sim_0.8_train_0.7_C_number	8.410 ± 6.430	10.040 ± 5.987	14.400 ± 9.890	6.530 ± 5.220	6.250 ± 7.770
Sim_0.8_train_0.7_D_number	9.250 ± 6.862	10.490 ± 6.713	15.710 ± 10.977	7.210 ± 5.023	4.970 ± 5.542
Sim_0.8_train_0.7_E_number	7.910 ± 5.952	9.040 ± 7.389	12.440 ± 9.136	6.990 ± 5.857	4.790 ± 5.470
Sim_0.9	0.001 ± 0.001	0.001 ± 0.001	0.002 ± 0.001	0.001 ± 0.001	0.001 ± 0.000
Sim_0.9_number	40.080 ± 21.504	41.090 ± 18.012	61.450 ± 28.271	38.710 ± 21.183	17.180 ± 12.111

Sim_0.9_A_number	0.380 ± 1.056	0.440 ± 1.033	0.550 ± 1.186	0.430 ± 1.358	0.140 ± 0.469
Sim_0.9_B_number	12.680 ± 12.393	13.170 ± 10.152	21.410 ± 15.719	13.730 ± 12.738	4.810 ± 5.486
Sim_0.9_C_number	11.690 ± 9.872	10.800 ± 6.456	16.910 ± 12.374	10.550 ± 8.133	5.240 ± 6.104
Sim_0.9_D_number	8.360 ± 7.203	9.080 ± 7.089	12.720 ± 9.679	7.630 ± 7.181	3.680 ± 4.382
Sim_0.9_E_number	6.970 ± 4.884	7.600 ± 4.750	9.860 ± 6.269	6.370 ± 5.302	3.310 ± 3.690
Sim_0.9_train_0.7	0.237 ± 0.152	0.276 ± 0.154	0.236 ± 0.109	0.210 ± 0.124	0.304 ± 0.205
Sim_0.9_train_0.7_number	8.600 ± 5.550	10.520 ± 5.633	13.690 ± 7.110	7.640 ± 5.311	4.770 ± 3.660
Sim_0.9_train_0.7_A_number	0.030 ± 0.171	0.060 ± 0.420	0.040 ± 0.196	0.020 ± 0.199	0.010 ± 0.099
Sim_0.9_train_0.7_B_number	1.980 ± 2.306	2.560 ± 2.613	3.690 ± 3.711	2.300 ± 2.820	1.170 ± 1.456
Sim_0.9_train_0.7_C_number	1.880 ± 1.872	2.170 ± 1.924	3.050 ± 3.074	1.630 ± 1.804	1.140 ± 1.715
Sim_0.9_train_0.7_D_number	2.470 ± 2.480	2.850 ± 2.459	3.460 ± 2.590	1.650 ± 1.699	1.120 ± 1.267
Sim_0.9_train_0.7_E_number	2.240 ± 2.173	2.880 ± 2.422	3.450 ± 3.001	2.040 ± 2.315	1.330 ± 1.817

Table S15. The fine-tuning results on the VEGFR 1%-fine-tuning datasets with RDKit filtering.

VEGFR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.856 ± 0.010	0.859 ± 0.012	0.848 ± 0.012	0.853 ± 0.011	0.853 ± 0.012
SNN/Gen_train	0.362 ± 0.046	0.403 ± 0.046	0.434 ± 0.038	0.345 ± 0.044	0.358 ± 0.035
SNN/Gen_goal	0.394 ± 0.028	0.411 ± 0.031	0.433 ± 0.026	0.389 ± 0.027	0.378 ± 0.024
IntDiv_Rediscovery	0.711 ± 0.047	0.717 ± 0.042	0.733 ± 0.038	0.695 ± 0.060	0.587 ± 0.129
SNN/Rediscovery_train	0.758 ± 0.044	0.756 ± 0.040	0.747 ± 0.039	0.760 ± 0.043	0.735 ± 0.079
Rediscovery	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.000 ± 0.000
Rediscovery_number	23.190 ± 11.532	21.970 ± 10.979	32.340 ± 12.662	20.870 ± 11.185	8.620 ± 6.709
Rediscovery_A_number	0.090 ± 0.286	0.050 ± 0.218	0.070 ± 0.255	0.010 ± 0.099	0.030 ± 0.171
Rediscovery_B_number	4.720 ± 4.325	4.690 ± 3.994	6.470 ± 4.627	4.890 ± 5.225	1.370 ± 2.382
Rediscovery_C_number	8.190 ± 6.326	7.340 ± 5.776	11.730 ± 7.384	7.010 ± 6.110	2.860 ± 2.936
Rediscovery_D_number	5.380 ± 3.977	5.000 ± 4.200	7.800 ± 4.968	4.660 ± 3.848	2.260 ± 3.081
Rediscovery_E_number	4.810 ± 3.635	4.890 ± 4.149	6.270 ± 4.071	4.300 ± 3.041	2.100 ± 1.952
Rediscovery_0.7	0.320 ± 0.146	0.333 ± 0.124	0.361 ± 0.134	0.302 ± 0.137	0.411 ± 0.256
Rediscovery_0.7_number	7.250 ± 4.371	7.110 ± 3.776	11.560 ± 5.783	6.260 ± 4.100	3.390 ± 3.289
Rediscovery_0.7_A_number	0.030 ± 0.171	0.020 ± 0.140	0.030 ± 0.171	0	0.010 ± 0.099
Rediscovery_0.7_B_number	1.150 ± 1.609	1.380 ± 1.362	2.000 ± 1.918	1.390 ± 1.913	0.430 ± 1.003
Rediscovery_0.7_C_number	2.210 ± 2.001	2.160 ± 1.821	3.830 ± 2.963	1.750 ± 1.705	1.010 ± 1.292
Rediscovery_0.7_D_number	2.060 ± 1.923	1.960 ± 1.854	3.290 ± 2.503	1.490 ± 1.533	1.040 ± 1.599
Rediscovery_0.7_E_number	1.800 ± 1.536	1.590 ± 1.569	2.410 ± 1.828	1.630 ± 1.635	0.900 ± 1.054
Sim_0.7	0.033 ± 0.018	0.041 ± 0.022	0.051 ± 0.021	0.033 ± 0.019	0.013 ± 0.009
Sim_0.7_number	989.380 ± 554.233	1226.600 ± 660.421	1529.240 ± 623.804	976.110 ± 581.677	401.410 ± 274.805
Sim_0.7_A_number	3.010 ± 7.570	3.190 ± 9.494	5.960 ± 17.210	2.800 ± 8.744	1.280 ± 5.419
Sim_0.7_B_number	201.780 ± 185.872	271.170 ± 266.274	303.800 ± 252.161	213.820 ± 233.825	66.960 ± 71.412
Sim_0.7_C_number	343.110 ± 365.455	415.260 ± 416.277	539.920 ± 427.664	324.710 ± 372.884	127.200 ± 173.813
Sim_0.7_D_number	236.950 ± 160.398	272.570 ± 193.512	372.720 ± 220.272	237.840 ± 174.385	104.980 ± 95.277

Sim_0.7_E_number	204.530 ± 146.860	264.410 ± 189.089	306.840 ± 192.063	196.940 ± 153.195	100.990 ± 97.120
Sim_0.7_train_0.7	0.261 ± 0.091	0.247 ± 0.087	0.273 ± 0.088	0.231 ± 0.089	0.388 ± 0.139
Sim_0.7_train_0.7_number	241.850 ± 133.491	293.040 ± 180.159	397.340 ± 161.785	214.980 ± 158.379	145.830 ± 93.011
Sim_0.7_train_0.7_A_number	1.300 ± 3.494	1.150 ± 3.087	2.300 ± 7.908	0.860 ± 2.379	0.930 ± 4.780
Sim_0.7_train_0.7_B_number	53.070 ± 60.415	71.200 ± 99.870	78.910 ± 72.597	50.170 ± 77.799	23.460 ± 30.905
Sim_0.7_train_0.7_C_number	75.110 ± 49.633	92.110 ± 69.641	131.860 ± 73.139	64.360 ± 56.564	42.150 ± 36.741
Sim_0.7_train_0.7_D_number	52.270 ± 30.455	61.180 ± 35.609	99.110 ± 49.647	51.460 ± 36.270	38.590 ± 40.075
Sim_0.7_train_0.7_E_number	60.100 ± 44.788	67.400 ± 47.070	85.160 ± 48.241	48.130 ± 48.073	40.700 ± 41.801
Sim_0.8	0.007 ± 0.004	0.007 ± 0.004	0.010 ± 0.004	0.006 ± 0.004	0.003 ± 0.002
Sim_0.8_number	198.100 ± 105.418	217.330 ± 115.843	305.540 ± 129.172	192.770 ± 112.135	77.040 ± 55.208
Sim_0.8_A_number	0.580 ± 1.491	0.670 ± 1.898	0.740 ± 2.048	0.360 ± 1.091	0.260 ± 1.205
Sim_0.8_B_number	39.070 ± 41.318	47.120 ± 51.542	60.290 ± 60.447	44.680 ± 57.844	13.370 ± 17.310
Sim_0.8_C_number	74.680 ± 69.560	79.700 ± 71.218	115.850 ± 81.840	67.580 ± 67.142	25.470 ± 31.699
Sim_0.8_D_number	45.020 ± 32.835	46.120 ± 35.803	71.750 ± 47.503	43.940 ± 31.410	20.960 ± 23.548
Sim_0.8_E_number	38.750 ± 26.714	43.720 ± 31.493	56.910 ± 42.359	36.210 ± 28.749	16.980 ± 16.952
Sim_0.8_train_0.7	0.229 ± 0.094	0.222 ± 0.090	0.235 ± 0.088	0.211 ± 0.087	0.315 ± 0.157
Sim_0.8_train_0.7_number	41.560 ± 20.388	45.050 ± 24.637	67.820 ± 29.321	37.280 ± 21.734	22.740 ± 18.108
Sim_0.8_train_0.7_A_number	0.300 ± 0.995	0.200 ± 0.735	0.400 ± 1.476	0.180 ± 0.726	0.170 ± 1.059
Sim_0.8_train_0.7_B_number	8.380 ± 8.148	10.290 ± 11.915	12.610 ± 10.876	8.240 ± 10.547	3.150 ± 5.078
Sim_0.8_train_0.7_C_number	13.340 ± 9.019	14.770 ± 9.895	24.370 ± 15.250	11.010 ± 8.594	7.360 ± 8.186
Sim_0.8_train_0.7_D_number	9.360 ± 6.511	9.910 ± 7.792	17.630 ± 11.650	9.240 ± 7.840	6.550 ± 8.090
Sim_0.8_train_0.7_E_number	10.180 ± 8.010	9.880 ± 8.495	12.810 ± 8.035	8.610 ± 7.730	5.510 ± 6.294
Sim_0.9	0.001 ± 0.001	0.001 ± 0.001	0.002 ± 0.001	0.001 ± 0.001	0
Sim_0.9_number	41.640 ± 21.722	39.550 ± 20.876	61.560 ± 26.564	38.740 ± 22.464	14.980 ± 12.327
Sim_0.9_A_number	0.110 ± 0.343	0.070 ± 0.255	0.100 ± 0.361	0.030 ± 0.171	0.060 ± 0.310

Sim_0.9_B_number	8.390 ± 7.806	8.000 ± 6.731	11.710 ± 9.912	8.790 ± 9.948	2.300 ± 3.692
Sim_0.9_C_number	16.030 ± 13.354	14.730 ± 11.700	23.460 ± 15.760	14.140 ± 13.516	5.160 ± 6.548
Sim_0.9_D_number	9.220 ± 7.405	8.660 ± 7.462	14.410 ± 9.851	8.320 ± 6.025	3.940 ± 5.067
Sim_0.9_E_number	7.890 ± 5.718	8.090 ± 6.271	11.880 ± 9.896	7.460 ± 6.894	3.520 ± 3.375
Sim_0.9_train_0.7	0.241 ± 0.127	0.251 ± 0.105	0.274 ± 0.112	0.229 ± 0.119	0.334 ± 0.190
Sim_0.9_train_0.7_number	9.330 ± 5.367	9.370 ± 5.087	15.990 ± 7.763	8.300 ± 5.047	4.540 ± 4.390
Sim_0.9_train_0.7_A_number	0.050 ± 0.260	0.020 ± 0.140	0.040 ± 0.196	0	0.030 ± 0.222
Sim_0.9_train_0.7_B_number	1.560 ± 1.835	1.790 ± 1.745	2.740 ± 2.700	1.750 ± 2.334	0.570 ± 1.219
Sim_0.9_train_0.7_C_number	2.890 ± 2.477	3.000 ± 2.289	5.470 ± 3.864	2.370 ± 2.274	1.410 ± 1.773
Sim_0.9_train_0.7_D_number	2.560 ± 2.282	2.510 ± 2.304	4.670 ± 3.502	2.010 ± 1.884	1.280 ± 2.069
Sim_0.9_train_0.7_E_number	2.270 ± 1.928	2.050 ± 2.022	3.070 ± 2.389	2.170 ± 2.079	1.250 ± 1.388

Table S16. The fine-tuning results on the AR 1%-fine-tuning datasets with RDKit filtering.

AR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.856 ± 0.012	0.852 ± 0.014	0.846 ± 0.015	0.855 ± 0.014	0.850 ± 0.013
SNN/Gen_train	0.377 ± 0.050	0.441 ± 0.044	0.438 ± 0.043	0.358 ± 0.049	0.373 ± 0.036
SNN/Gen_goal	0.406 ± 0.032	0.440 ± 0.031	0.438 ± 0.030	0.401 ± 0.032	0.390 ± 0.024
IntDiv_Rediscovery	0.717 ± 0.054	0.729 ± 0.043	0.732 ± 0.036	0.714 ± 0.047	0.621 ± 0.132
SNN/Rediscovery_train	0.780 ± 0.053	0.757 ± 0.052	0.769 ± 0.048	0.762 ± 0.052	0.759 ± 0.077
Rediscovery	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.001 ± 0.000	0.000 ± 0.000
Rediscovery_number	30.190 ± 12.110	27.800 ± 12.785	**36.100 ± 12.905**	26.430 ± 11.529	11.770 ± 6.588
Rediscovery_A_number	**0.390 ± 0.747**	0.360 ± 0.794	0.310 ± 0.703	0.290 ± 0.697	0.100 ± 0.361
Rediscovery_B_number	13.550 ± 6.360	11.510 ± 6.388	**16.710 ± 7.203**	11.420 ± 5.793	5.160 ± 4.110
Rediscovery_C_number	8.290 ± 5.307	8.280 ± 5.738	10.310 ± 6.236	7.950 ± 5.371	3.530 ± 2.903
Rediscovery_D_number	4.840 ± 3.440	4.920 ± 3.399	5.370 ± 3.300	4.130 ± 3.097	1.930 ± 2.031
Rediscovery_E_number	3.120 ± 3.476	2.730 ± 2.638	3.400 ± 3.243	2.640 ± 2.755	1.050 ± 1.552
Rediscovery_0.7	0.299 ± 0.127	0.357 ± 0.135	0.327 ± 0.118	0.342 ± 0.138	0.358 ± 0.226
Rediscovery_0.7_number	9.160 ± 5.704	9.890 ± 5.641	**11.900 ± 6.283**	9.170 ± 6.155	4.050 ± 3.021
Rediscovery_0.7_A_number	**0.380 ± 0.745**	0.340 ± 0.790	0.290 ± 0.697	0.270 ± 0.676	0.080 ± 0.337
Rediscovery_0.7_B_number	3.240 ± 2.673	3.120 ± 2.601	**4.500 ± 3.339**	2.840 ± 2.501	1.350 ± 1.584
Rediscovery_0.7_C_number	2.540 ± 2.242	2.950 ± 2.360	3.270 ± 2.619	2.970 ± 2.651	1.370 ± 1.566
Rediscovery_0.7_D_number	1.960 ± 1.933	2.300 ± 1.947	2.300 ± 1.819	1.960 ± 2.083	0.880 ± 1.336
Rediscovery_0.7_E_number	1.040 ± 1.428	1.180 ± 1.135	1.540 ± 1.694	1.130 ± 1.405	0.370 ± 0.783
Sim_0.7	0.037 ± 0.020	0.050 ± 0.024	0.050 ± 0.022	0.035 ± 0.020	0.017 ± 0.010
Sim_0.7_number	1124.290 ± 605.472	1487.650 ± 721.985	**1496.370 ± 648.527**	1041.600 ± 605.231	500.090 ± 306.275
Sim_0.7_A_number	5.830 ± 9.296	12.650 ± 25.640	9.760 ± 20.173	5.720 ± 14.164	4.230 ± 9.779
Sim_0.7_B_number	595.630 ± 399.855	812.050 ± 494.449	813.450 ± 439.053	558.610 ± 397.305	248.570 ± 175.199
Sim_0.7_C_number	289.570 ± 189.732	364.660 ± 240.382	380.270 ± 224.737	264.030 ± 182.068	141.020 ± 124.503
Sim_0.7_D_number	133.940 ±	180.590 ±	179.990 ±	116.680 ±	62.370 ±

Sim_0.7_E_number	96.881 99.320 ± 144.310	129.325 117.700 ± 161.794	124.161 112.900 ± 146.574	94.656 96.560 ± 191.230	49.922 43.900 ± 73.750
Sim_0.7_train_0.7	0.270 ± 0.105	0.291 ± 0.100	0.277 ± 0.102	0.268 ± 0.127	0.381 ± 0.155
Sim_0.7_train_0.7_number	285.060 ± 165.470	403.990 ± 183.319	399.320 ± 199.223	261.440 ± 190.138	175.750 ± 106.664
Sim_0.7_train_0.7_A_number	4.430 ± 6.422	9.910 ± 20.825	7.340 ± 16.996	3.670 ± 8.111	3.410 ± 8.736
Sim_0.7_train_0.7_B_number	108.470 ± 63.375	157.660 ± 93.894	157.920 ± 82.174	91.000 ± 61.729	67.300 ± 44.432
Sim_0.7_train_0.7_C_number	84.050 ± 58.894	110.530 ± 65.764	117.930 ± 78.255	78.030 ± 56.646	54.340 ± 46.887
Sim_0.7_train_0.7_D_number	58.150 ± 46.306	80.710 ± 48.227	77.480 ± 54.633	52.600 ± 51.358	32.750 ± 30.467
Sim_0.7_train_0.7_E_number	29.960 ± 46.267	45.180 ± 69.489	38.650 ± 54.582	36.140 ± 83.414	17.950 ± 26.602
Sim_0.8	0.009 ± 0.004	0.010 ± 0.005	0.011 ± 0.005	0.008 ± 0.004	0.003 ± 0.002
Sim_0.8_number	256.250 ± 122.585	303.480 ± 152.162	340.610 ± 142.806	230.810 ± 119.935	98.410 ± 58.935
Sim_0.8_A_number	1.630 ± 2.788	2.850 ± 6.131	2.530 ± 6.154	1.640 ± 5.469	1.000 ± 3.473
Sim_0.8_B_number	134.630 ± 81.715	168.830 ± 117.390	187.880 ± 99.580	123.600 ± 79.878	49.370 ± 37.388
Sim_0.8_C_number	67.330 ± 47.269	74.950 ± 50.175	87.770 ± 56.055	61.500 ± 48.443	27.180 ± 24.968
Sim_0.8_D_number	27.770 ± 20.521	33.270 ± 24.128	36.710 ± 29.325	23.960 ± 18.853	12.090 ± 10.403
Sim_0.8_E_number	24.890 ± 39.538	23.580 ± 32.835	25.720 ± 35.572	20.110 ± 34.993	8.770 ± 15.994
Sim_0.8_train_0.7	0.210 ± 0.100	0.242 ± 0.100	0.214 ± 0.102	0.222 ± 0.122	0.297 ± 0.163
Sim_0.8_train_0.7_number	50.200 ± 28.037	66.940 ± 33.357	68.280 ± 35.804	45.940 ± 27.627	27.060 ± 19.361
Sim_0.8_train_0.7_A_number	1.190 ± 1.853	2.180 ± 3.681	2.020 ± 5.826	0.870 ± 1.880	0.760 ± 3.265
Sim_0.8_train_0.7_B_number	19.550 ± 13.358	24.710 ± 17.390	26.200 ± 15.074	16.210 ± 10.102	9.450 ± 7.953
Sim_0.8_train_0.7_C_number	14.190 ± 10.959	18.810 ± 12.794	20.440 ± 15.279	13.920 ± 10.815	8.510 ± 9.097
Sim_0.8_train_0.7_D_number	10.480 ± 8.201	14.420 ± 9.754	13.580 ± 10.487	10.140 ± 9.388	5.380 ± 6.221
Sim_0.8_train_0.7_E_number	4.790 ± 5.689	6.820 ± 9.494	6.040 ± 6.720	4.800 ± 6.863	2.960 ± 5.836
Sim_0.9	0.002 ± 0.001	0.002 ± 0.001	0.003 ± 0.001	0.002 ± 0.001	0.001 ± 0.000
Sim_0.9_number	61.930 ± 26.679	63.180 ± 29.951	82.200 ± 36.071	54.250 ± 24.834	24.070 ± 12.925

Sim_0.9_A_number	0.890 ± 1.413	0.990 ± 1.847	1.390 ± 3.731	0.600 ± 1.114	0.510 ± 2.322
Sim_0.9_B_number	28.220 ± 14.703	29.510 ± 19.531	40.430 ± 23.871	24.880 ± 14.844	10.580 ± 8.594
Sim_0.9_C_number	18.680 ± 13.016	18.050 ± 13.109	23.620 ± 17.017	16.580 ± 12.365	7.230 ± 6.313
Sim_0.9_D_number	8.720 ± 6.342	9.520 ± 7.180	10.420 ± 7.798	7.750 ± 6.709	3.450 ± 3.031
Sim_0.9_E_number	5.420 ± 6.022	5.110 ± 6.091	6.340 ± 6.668	4.440 ± 5.092	2.300 ± 4.431
Sim_0.9_train_0.7	0.241 ± 0.107	0.287 ± 0.128	0.265 ± 0.121	0.267 ± 0.129	0.314 ± 0.200
Sim_0.9_train_0.7_number	14.530 ± 8.281	16.860 ± 9.066	20.060 ± 10.587	13.770 ± 8.778	7.200 ± 6.185
Sim_0.9_train_0.7_A_number	0.820 ± 1.330	0.900 ± 1.797	1.260 ± 3.703	0.540 ± 1.004	0.480 ± 2.300
Sim_0.9_train_0.7_B_number	4.610 ± 3.355	5.150 ± 4.234	6.740 ± 4.500	3.850 ± 3.211	2.150 ± 2.455
Sim_0.9_train_0.7_C_number	4.090 ± 3.172	4.730 ± 3.301	5.750 ± 3.634	4.420 ± 3.163	2.390 ± 2.549
Sim_0.9_train_0.7_D_number	3.270 ± 2.698	4.220 ± 4.161	4.110 ± 3.228	3.340 ± 3.705	1.360 ± 1.591
Sim_0.9_train_0.7_E_number	1.740 ± 2.614	1.860 ± 1.844	2.200 ± 2.498	1.620 ± 2.310	0.820 ± 2.511

Table S17. The fine-tuning results on the 5-HTR 1%-fine-tuning datasets with RDKit filtering.

5-HTR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.850 ± 0.010	0.847 ± 0.012	0.843 ± 0.012	0.844 ± 0.013	0.846 ± 0.010
SNN/Gen_train	0.438 ± 0.034	0.478 ± 0.041	0.476 ± 0.028	0.421 ± 0.036	0.415 ± 0.024
SNN/Gen_goal	0.463 ± 0.021	0.482 ± 0.028	0.483 ± 0.021	0.464 ± 0.023	0.438 ± 0.018
IntDiv_Rediscovery	0.783 ± 0.020	0.784 ± 0.021	0.784 ± 0.021	0.773 ± 0.025	0.760 ± 0.028
SNN/Rediscovery_train	0.758 ± 0.032	0.723 ± 0.046	0.743 ± 0.039	0.736 ± 0.037	0.741 ± 0.046
Rediscovery	0.003 ± 0.001	0.003 ± 0.001	0.003 ± 0.001	0.002 ± 0.001	0.001 ± 0.000
Rediscovery_number	84.920 ± 19.353	76.000 ± 21.576	**91.730 ± 20.555**	73.470 ± 17.791	42.480 ± 14.111
Rediscovery_A_number	0.110 ± 0.313	0.150 ± 0.409	**0.260 ± 0.594**	0.100 ± 0.300	0.110 ± 0.422
Rediscovery_B_number	36.140 ± 11.422	31.190 ± 11.294	**37.710 ± 11.698**	30.500 ± 10.564	17.090 ± 7.053
Rediscovery_C_number	28.150 ± 9.376	24.610 ± 9.241	30.930 ± 10.077	24.140 ± 8.555	14.200 ± 6.366
Rediscovery_D_number	15.580 ± 6.561	14.940 ± 6.015	17.240 ± 6.470	14.070 ± 5.494	8.340 ± 4.570
Rediscovery_E_number	4.940 ± 2.709	5.110 ± 2.541	5.590 ± 2.804	4.660 ± 2.471	2.740 ± 2.086
Rediscovery_0.7	0.371 ± 0.088	0.452 ± 0.118	0.409 ± 0.100	0.420 ± 0.096	0.410 ± 0.124
Rediscovery_0.7_number	31.240 ± 9.470	33.360 ± 10.177	**37.020 ± 10.617**	30.450 ± 8.955	17.270 ± 7.879
Rediscovery_0.7_A_number	0.090 ± 0.286	0.130 ± 0.391	**0.200 ± 0.469**	0.050 ± 0.218	0.070 ± 0.255
Rediscovery_0.7_B_number	12.100 ± 4.588	12.030 ± 4.934	**13.590 ± 5.058**	10.870 ± 4.139	6.170 ± 3.181
Rediscovery_0.7_C_number	10.080 ± 4.599	10.640 ± 5.015	12.090 ± 5.622	9.620 ± 4.651	5.430 ± 3.564
Rediscovery_0.7_D_number	6.410 ± 3.750	7.320 ± 3.328	8.080 ± 3.979	6.940 ± 3.056	4.140 ± 3.010
Rediscovery_0.7_E_number	2.560 ± 1.751	3.240 ± 2.178	3.060 ± 2.190	2.970 ± 2.027	1.460 ± 1.513
Sim_0.7	0.060 ± 0.019	0.065 ± 0.025	0.070 ± 0.019	0.056 ± 0.020	0.034 ± 0.012
Sim_0.7_number	1790.590 ± 557.284	1943.740 ± 735.696	**2098.580 ± 579.483**	1675.910 ± 610.011	1007.500 ± 359.015
Sim_0.7_A_number	3.540 ± 5.428	3.930 ± 5.243	4.790 ± 7.415	3.220 ± 7.123	2.280 ± 5.115
Sim_0.7_B_number	847.850 ± 369.246	921.350 ± 464.917	995.590 ± 387.849	809.570 ± 414.331	456.310 ± 222.363
Sim_0.7_C_number	560.900 ± 254.991	601.200 ± 282.992	654.700 ± 251.496	525.050 ± 264.934	321.970 ± 167.797
Sim_0.7_D_number	275.950 ± 105.472	310.590 ± 152.490	324.480 ± 129.247	251.260 ± 112.305	167.530 ± 72.741

Sim_0.7_E_number	102.350 ± 79.768	106.670 ± 73.135	119.020 ± 86.634	86.810 ± 81.014	59.410 ± 50.407
Sim_0.7_train_0.7	0.311 ± 0.072	0.333 ± 0.081	0.317 ± 0.070	0.347 ± 0.081	0.403 ± 0.100
Sim_0.7_train_0.7_number	540.570 ± 161.571	616.420 ± 212.535	650.050 ± 185.183	564.970 ± 205.404	391.740 ± 129.592
Sim_0.7_train_0.7_A_number	1.910 ± 2.546	2.290 ± 2.947	2.920 ± 4.372	1.530 ± 1.977	1.600 ± 3.575
Sim_0.7_train_0.7_B_number	230.130 ± 86.717	254.610 ± 106.882	273.450 ± 99.502	241.020 ± 105.344	160.040 ± 66.697
Sim_0.7_train_0.7_C_number	177.680 ± 68.801	208.620 ± 93.290	216.120 ± 77.656	193.470 ± 96.314	131.870 ± 59.527
Sim_0.7_train_0.7_D_number	89.630 ± 32.005	105.530 ± 42.153	111.050 ± 42.978	92.740 ± 32.322	71.950 ± 34.528
Sim_0.7_train_0.7_E_number	41.220 ± 23.920	45.370 ± 24.134	46.510 ± 24.958	36.210 ± 19.233	26.280 ± 17.831
Sim_0.8	0.015 ± 0.004	0.014 ± 0.005	0.017 ± 0.005	0.013 ± 0.004	0.008 ± 0.003
Sim_0.8_number	438.910 ± 117.793	433.540 ± 160.584	522.170 ± 138.786	391.080 ± 115.408	230.270 ± 75.369
Sim_0.8_A_number	0.910 ± 1.955	0.690 ± 1.294	0.980 ± 2.107	0.910 ± 4.000	0.440 ± 1.203
Sim_0.8_B_number	201.300 ± 71.681	198.930 ± 93.746	244.310 ± 91.788	183.050 ± 68.576	102.820 ± 45.831
Sim_0.8_C_number	143.780 ± 56.883	140.760 ± 73.204	169.440 ± 66.049	128.100 ± 60.198	76.530 ± 37.741
Sim_0.8_D_number	67.810 ± 27.558	67.510 ± 35.436	78.370 ± 35.244	58.340 ± 25.019	36.740 ± 17.427
Sim_0.8_E_number	25.110 ± 16.993	25.650 ± 19.361	29.070 ± 20.601	20.680 ± 14.293	13.740 ± 12.324
Sim_0.8_train_0.7	0.268 ± 0.072	0.308 ± 0.097	0.272 ± 0.077	0.305 ± 0.082	0.334 ± 0.104
Sim_0.8_train_0.7_number	113.400 ± 30.020	123.980 ± 41.352	136.620 ± 37.666	114.940 ± 31.693	73.920 ± 26.801
Sim_0.8_train_0.7_A_number	0.410 ± 0.826	0.380 ± 0.772	0.550 ± 1.228	0.280 ± 0.584	0.290 ± 0.898
Sim_0.8_train_0.7_B_number	47.320 ± 16.652	48.860 ± 19.456	56.160 ± 20.780	48.330 ± 15.966	29.490 ± 13.444
Sim_0.8_train_0.7_C_number	37.010 ± 11.622	42.020 ± 18.168	45.280 ± 15.708	37.580 ± 13.039	24.230 ± 11.520
Sim_0.8_train_0.7_D_number	20.100 ± 9.611	23.150 ± 10.688	24.840 ± 11.210	20.680 ± 8.643	14.810 ± 9.225
Sim_0.8_train_0.7_E_number	8.560 ± 4.938	9.570 ± 5.094	9.790 ± 5.311	8.070 ± 4.763	5.100 ± 3.867
Sim_0.9	0.005 ± 0.001	0.005 ± 0.002	0.006 ± 0.002	0.005 ± 0.001	0.003 ± 0.001
Sim_0.9_number	155.060 ± 39.798	143.080 ± 48.632	182.140 ± 47.389	139.830 ± 39.135	79.240 ± 26.431
Sim_0.9_A_number	0.270 ± 0.661	0.260 ± 0.577	0.360 ± 0.781	0.190 ± 0.462	0.200 ± 0.632

Sim_0.9_B_number	67.930 ± 24.620	60.510 ± 26.541	81.540 ± 32.683	61.370 ± 23.421	32.820 ± 14.675
Sim_0.9_C_number	52.400 ± 20.129	48.320 ± 22.535	61.700 ± 22.818	46.910 ± 19.506	27.760 ± 13.250
Sim_0.9_D_number	25.100 ± 10.323	24.590 ± 12.369	28.230 ± 10.922	23.260 ± 9.831	13.720 ± 7.347
Sim_0.9_E_number	9.360 ± 6.856	9.400 ± 7.129	10.310 ± 6.864	8.100 ± 5.187	4.740 ± 3.810
Sim_0.9_train_0.7	0.293 ± 0.080	0.363 ± 0.114	0.308 ± 0.092	0.337 ± 0.092	0.352 ± 0.115
Sim_0.9_train_0.7_number	43.910 ± 12.166	48.730 ± 15.444	54.090 ± 15.978	45.710 ± 13.979	27.070 ± 11.489
Sim_0.9_train_0.7_A_number	0.220 ± 0.593	0.230 ± 0.563	0.280 ± 0.634	0.110 ± 0.343	0.160 ± 0.543
Sim_0.9_train_0.7_B_number	17.310 ± 6.311	18.010 ± 7.137	20.690 ± 7.996	17.660 ± 6.377	9.990 ± 4.789
Sim_0.9_train_0.7_C_number	14.900 ± 6.210	16.620 ± 7.522	18.840 ± 8.291	15.360 ± 6.563	9.120 ± 5.373
Sim_0.9_train_0.7_D_number	8.290 ± 4.339	9.990 ± 4.698	10.510 ± 4.734	8.970 ± 4.046	5.890 ± 4.411
Sim_0.9_train_0.7_E_number	3.190 ± 1.906	3.880 ± 2.384	3.770 ± 2.310	3.610 ± 2.358	1.910 ± 1.715

Table S18. The fine-tuning results on the DR 1%-fine-tuning datasets with RDKit filtering.

DR	CharRNN	AAE	VAE	Reinvent	ORGAN
IntDiv	0.847 ± 0.009	0.844 ± 0.010	0.839 ± 0.009	0.843 ± 0.010	0.843 ± 0.011
SNN/Gen_train	0.411 ± 0.043	0.464 ± 0.035	0.458 ± 0.037	0.395 ± 0.042	0.398 ± 0.036
SNN/Gen_goal	0.447 ± 0.023	0.476 ± 0.021	0.473 ± 0.020	0.447 ± 0.023	0.425 ± 0.024
IntDiv_Rediscovery	0.751 ± 0.021	0.750 ± 0.024	0.749 ± 0.022	0.740 ± 0.024	0.710 ± 0.037
SNN/Rediscovery_train	0.745 ± 0.035	0.717 ± 0.041	0.736 ± 0.033	0.724 ± 0.039	0.731 ± 0.052
Rediscovery	0.002 ± 0.001	0.002 ± 0.001	0.002 ± 0.001	0.002 ± 0.001	0.001 ± 0.000
Rediscovery_number	57.430 ± 16.921	57.220 ± 17.993	**66.540 ± 20.585**	52.020 ± 16.123	25.750 ± 11.121
Rediscovery_A_number	0.220 ± 0.460	**0.320 ± 0.631**	0.270 ± 0.545	0.180 ± 0.384	0.160 ± 0.418
Rediscovery_B_number	23.190 ± 11.166	23.550 ± 11.396	**28.570 ± 13.114**	20.970 ± 10.534	10.290 ± 6.489
Rediscovery_C_number	16.440 ± 6.923	16.160 ± 6.069	18.970 ± 6.968	15.220 ± 5.881	7.330 ± 4.055
Rediscovery_D_number	12.220 ± 4.491	12.160 ± 5.240	13.090 ± 5.115	11.460 ± 4.928	5.730 ± 3.638
Rediscovery_E_number	5.360 ± 3.604	5.030 ± 2.975	5.650 ± 3.505	4.190 ± 2.876	2.240 ± 2.367
Rediscovery_0.7	0.397 ± 0.105	0.462 ± 0.111	0.419 ± 0.094	0.441 ± 0.115	0.435 ± 0.148
Rediscovery_0.7_number	22.740 ± 8.778	25.870 ± 8.991	**28.020 ± 11.220**	23.050 ± 9.751	11.170 ± 6.744
Rediscovery_0.7_A_number	0.190 ± 0.417	**0.270 ± 0.581**	0.210 ± 0.475	0.140 ± 0.347	0.130 ± 0.365
Rediscovery_0.7_B_number	9.400 ± 5.607	10.860 ± 6.215	**12.460 ± 7.250**	9.220 ± 6.090	4.700 ± 4.063
Rediscovery_0.7_C_number	5.940 ± 3.243	6.940 ± 3.171	7.580 ± 3.488	6.300 ± 3.407	3.120 ± 2.325
Rediscovery_0.7_D_number	4.980 ± 2.518	5.620 ± 2.863	5.620 ± 3.055	5.740 ± 3.189	2.430 ± 2.165
Rediscovery_0.7_E_number	2.230 ± 1.702	2.180 ± 1.526	2.150 ± 1.746	1.650 ± 1.479	0.790 ± 0.973
Sim_0.7	0.049 ± 0.019	0.059 ± 0.020	0.061 ± 0.019	0.047 ± 0.019	0.024 ± 0.010
Sim_0.7_number	1477.540 ± 559.394	1782.190 ± 605.383	**1835.910 ± 563.612**	1414.780 ± 577.227	723.280 ± 308.095
Sim_0.7_A_number	3.700 ± 5.839	4.540 ± 6.231	4.970 ± 7.883	3.250 ± 8.645	3.000 ± 7.509
Sim_0.7_B_number	648.530 ± 451.821	776.490 ± 450.045	817.690 ± 471.305	620.090 ± 464.982	311.820 ± 227.316
Sim_0.7_C_number	437.000 ± 170.432	530.550 ± 205.280	542.820 ± 193.134	432.270 ± 187.849	214.030 ± 106.075
Sim_0.7_D_number	281.750 ± 136.505	346.060 ± 173.306	345.210 ± 152.989	267.590 ± 147.880	137.910 ± 69.783

Sim_0.7_E_number	106.560 ± 55.952	124.550 ± 70.170	125.220 ± 66.404	91.580 ± 54.369	56.520 ± 44.442
Sim_0.7_train_0.7	0.351 ± 0.073	0.362 ± 0.090	0.358 ± 0.073	0.392 ± 0.090	0.455 ± 0.097
Sim_0.7_train_0.7_number	498.990 ± 161.567	616.150 ± 185.732	636.370 ± 167.507	530.160 ± 184.791	317.540 ± 117.505
Sim_0.7_train_0.7_A_number	2.220 ± 3.540	2.620 ± 3.481	2.910 ± 5.246	1.340 ± 2.201	2.140 ± 6.609
Sim_0.7_train_0.7_B_number	222.540 ± 127.849	270.650 ± 130.828	290.730 ± 129.734	230.940 ± 127.554	137.800 ± 79.018
Sim_0.7_train_0.7_C_number	140.840 ± 49.251	176.880 ± 60.420	180.510 ± 56.147	159.230 ± 63.160	92.730 ± 44.138
Sim_0.7_train_0.7_D_number	95.080 ± 32.102	122.310 ± 44.077	117.450 ± 40.599	101.510 ± 43.503	59.650 ± 31.264
Sim_0.7_train_0.7_E_number	38.310 ± 15.631	43.690 ± 19.945	44.770 ± 20.702	37.140 ± 18.150	25.220 ± 26.227
Sim_0.8	0.012 ± 0.004	0.013 ± 0.005	0.015 ± 0.005	0.011 ± 0.004	0.005 ± 0.002
Sim_0.8_number	348.630 ± 121.957	387.530 ± 147.194	438.650 ± 149.340	331.140 ± 127.557	157.870 ± 66.899
Sim_0.8_A_number	0.870 ± 1.718	1.030 ± 1.763	0.940 ± 1.912	0.820 ± 2.647	0.430 ± 1.089
Sim_0.8_B_number	149.630 ± 92.616	171.210 ± 102.601	195.900 ± 114.588	143.290 ± 93.890	66.760 ± 47.234
Sim_0.8_C_number	104.070 ± 41.543	113.120 ± 52.931	129.000 ± 50.130	101.450 ± 47.638	45.920 ± 21.295
Sim_0.8_D_number	67.480 ± 31.149	72.890 ± 37.222	81.470 ± 37.739	64.310 ± 31.415	32.460 ± 18.052
Sim_0.8_E_number	26.580 ± 17.829	29.280 ± 20.774	31.340 ± 21.525	21.270 ± 14.784	12.300 ± 11.608
Sim_0.8_train_0.7	0.310 ± 0.086	0.339 ± 0.117	0.326 ± 0.082	0.354 ± 0.100	0.398 ± 0.115
Sim_0.8_train_0.7_number	102.970 ± 30.230	121.340 ± 39.433	137.280 ± 40.996	111.110 ± 36.760	60.460 ± 25.365
Sim_0.8_train_0.7_A_number	0.650 ± 1.314	0.780 ± 1.432	0.630 ± 1.197	0.360 ± 0.806	0.350 ± 0.973
Sim_0.8_train_0.7_B_number	44.640 ± 19.792	54.490 ± 26.623	63.060 ± 27.818	48.670 ± 21.518	25.860 ± 14.665
Sim_0.8_train_0.7_C_number	28.870 ± 11.298	31.780 ± 12.811	37.370 ± 13.574	31.620 ± 13.884	17.480 ± 10.214
Sim_0.8_train_0.7_D_number	20.730 ± 8.787	25.330 ± 11.035	26.620 ± 11.440	23.400 ± 10.153	12.240 ± 7.660
Sim_0.8_train_0.7_E_number	8.080 ± 4.489	8.960 ± 4.923	9.600 ± 5.720	7.060 ± 4.347	4.530 ± 5.423
Sim_0.9	0.004 ± 0.001	0.004 ± 0.001	0.005 ± 0.002	0.004 ± 0.001	0.002 ± 0.001
Sim_0.9_number	113.180 ± 38.030	116.700 ± 41.767	139.840 ± 47.539	106.070 ± 37.011	50.370 ± 22.262
Sim_0.9_A_number	0.400 ± 0.849	0.590 ± 1.050	0.520 ± 1.109	0.330 ± 0.813	0.210 ± 0.535

Sim_0.9_B_number	49.160 ± 26.370	51.060 ± 27.430	62.810 ± 32.313	44.880 ± 25.341	21.070 ± 12.879
Sim_0.9_C_number	32.670 ± 13.754	33.030 ± 14.153	39.830 ± 16.287	32.370 ± 15.577	14.140 ± 7.981
Sim_0.9_D_number	21.380 ± 10.202	22.230 ± 10.994	25.490 ± 12.777	20.490 ± 9.784	10.430 ± 6.677
Sim_0.9_E_number	9.570 ± 7.689	9.790 ± 8.359	11.190 ± 8.812	8.000 ± 6.256	4.520 ± 4.947
Sim_0.9_train_0.7	0.329 ± 0.098	0.380 ± 0.115	0.345 ± 0.091	0.374 ± 0.108	0.393 ± 0.138
Sim_0.9_train_0.7_number	36.060 ± 13.074	42.040 ± 14.264	47.210 ± 17.107	38.320 ± 14.445	19.280 ± 10.690
Sim_0.9_train_0.7_A_number	0.320 ± 0.691	0.490 ± 0.943	0.370 ± 0.783	0.220 ± 0.576	0.170 ± 0.470
Sim_0.9_train_0.7_B_number	15.670 ± 7.669	18.380 ± 9.748	22.070 ± 11.196	16.160 ± 8.378	8.320 ± 5.767
Sim_0.9_train_0.7_C_number	9.740 ± 5.009	11.240 ± 4.723	12.650 ± 5.601	10.870 ± 5.289	5.240 ± 3.808
Sim_0.9_train_0.7_D_number	7.210 ± 3.769	8.630 ± 4.279	8.690 ± 5.213	8.280 ± 4.402	4.010 ± 3.534
Sim_0.9_train_0.7_E_number	3.120 ± 2.273	3.300 ± 2.632	3.430 ± 2.650	2.790 ± 2.104	1.540 ± 1.878

Table S19. Class A compounds reproduced by VAE and their corresponding nearest neighbor molecules in the AR fine-tuning dataset.

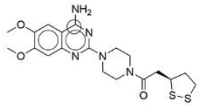
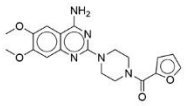
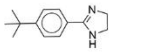
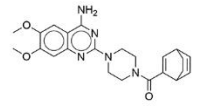
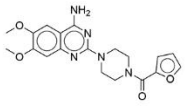
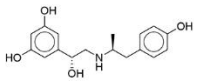
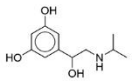
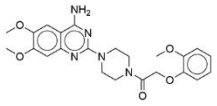
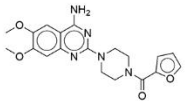
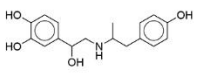
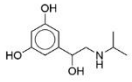
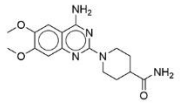
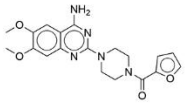
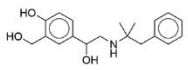
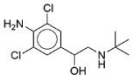
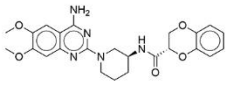
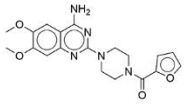
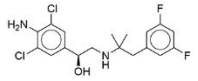
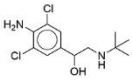
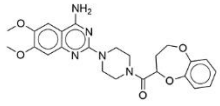
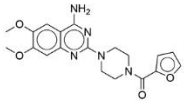
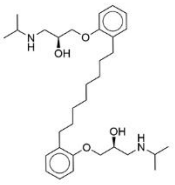
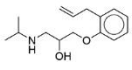
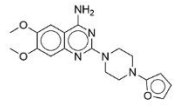
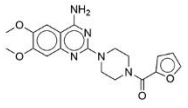
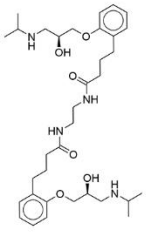
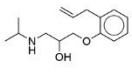
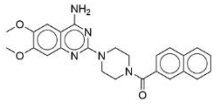
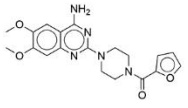
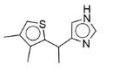
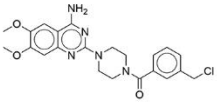
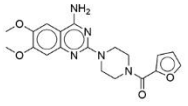
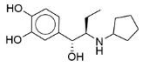
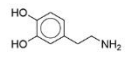
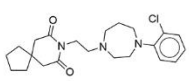
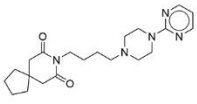
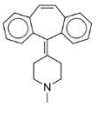
Target	The most similar compounds in training set	Level	Reproduced class A compounds	Target	The most similar compounds in training set	Level	Reproduced class A compounds
AR		B		AR		C	
AR		B		AR		B	
AR		B		AR		D	
AR		B		AR		C	
AR		B		AR		D	
AR		B		AR		B	
AR		B		AR		B	
AR		C		AR		B	
AR		C		AR		B	
AR		B		AR		C	
AR		B		AR		B	
AR		D		AR		A	
AR		B		AR		A	
AR		C		AR		C	

Table S20. Class A compounds reproduced by VAE and their corresponding nearest neighbor molecules in the 5-HTR fine-tuning dataset.

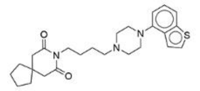
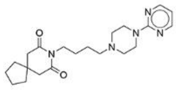
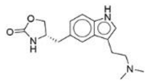
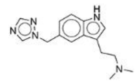
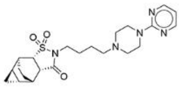
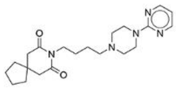
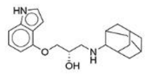
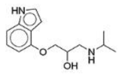
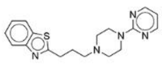
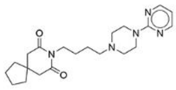
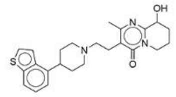
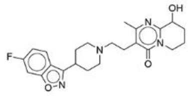
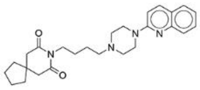
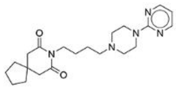
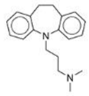
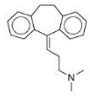
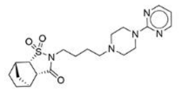
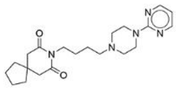
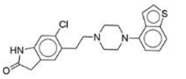
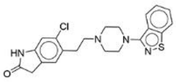
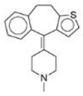
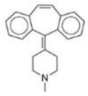
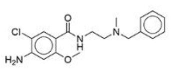
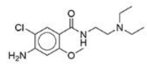
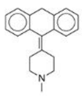
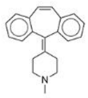
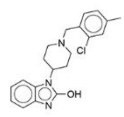
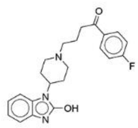
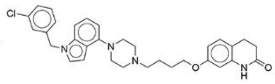
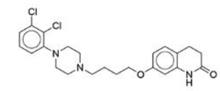
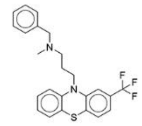
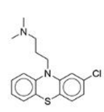
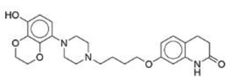
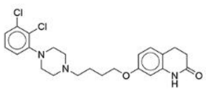
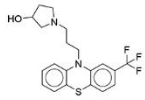
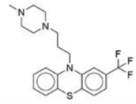
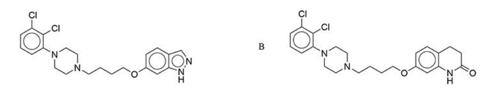
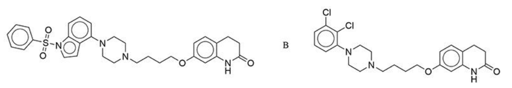
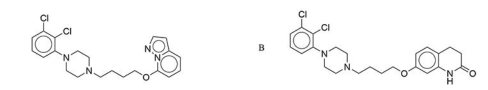
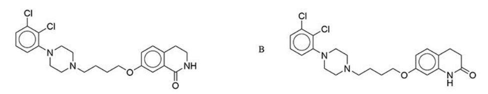
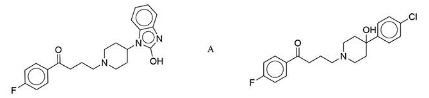
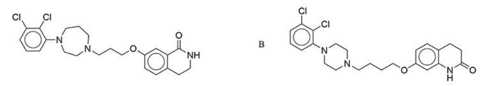
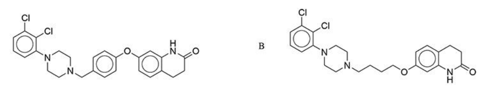
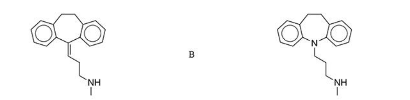
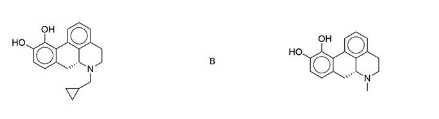
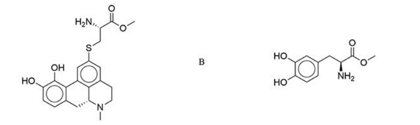
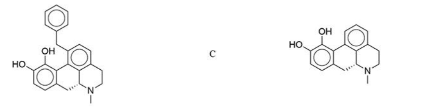
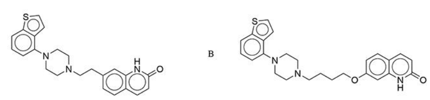
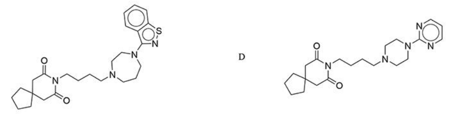
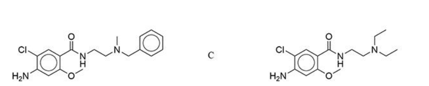
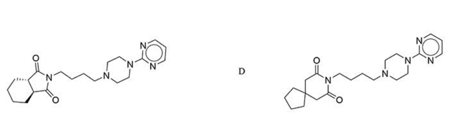
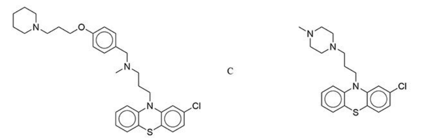
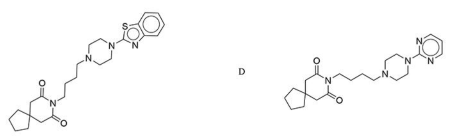
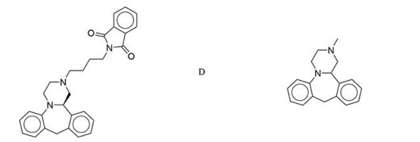
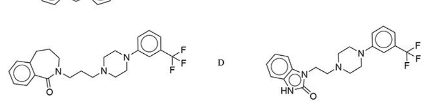
Target	The most similar compounds in training set	Level	Reproduced class A compounds	Target	The most similar compounds in training set	Level	Reproduced class A compounds
5-HTR		B		5-HTR		A	
5-HTR		C		5-HTR		B	
5-HTR		C		5-HTR		B	
5-HTR		C		5-HTR		B	
5-HTR		C		5-HTR		B	
5-HTR		B		5-HTR		D	
5-HTR		B		5-HTR		D	
5-HTR		B		5-HTR		E	
5-HTR		D		5-HTR		E	

Table S21. Class A compounds reproduced by VAE and their corresponding nearest neighbor molecules in the DR fine-tuning dataset.

Target	The most similar compounds in training set	Level	Reproduced class A compounds	Target	The most similar compounds in training set	Level	Reproduced class A compounds
DR		B		DR		D	
DR		B		DR		D	
DR		B		DR		E	
DR		B		DR		A	
DR		B		DR		B	
DR		B		DR		B	
DR		B		DR		B	
DR		B		DR		B	
DR		C		DR		B	
DR		D		DR		C	
DR		D		DR		C	
DR		D		DR		D	
DR		B		DR		D	