Supplementary materials

Deep active learning with high structural discriminability for molecular mutagenicity prediction

Huiyan Xu [a, 1], Yanpeng Zhao [b, 1], Junshan Han [b], Peng Zan [a, *], Song He [b, *], Xiaochen Bo [b, *]

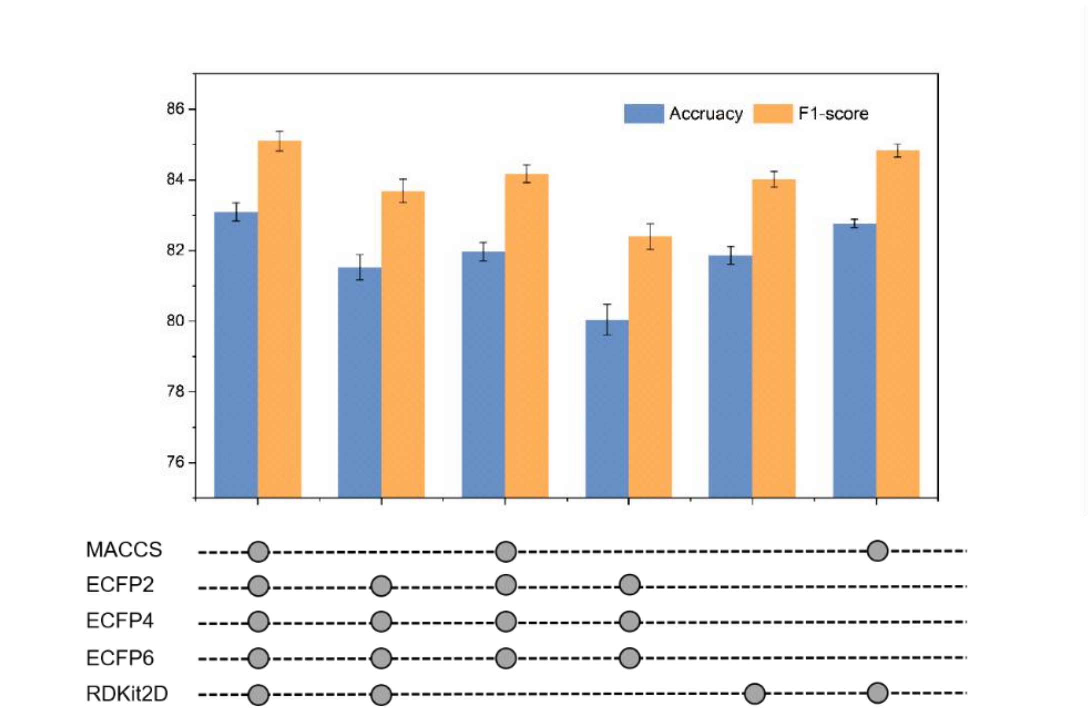
a Shanghai Key Laboratory of Power Station Automation Technology, School of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200444, China

b Department of Biotechnology, Institute of Health Service and Transfusion Medicine, Beijing, 100850, China

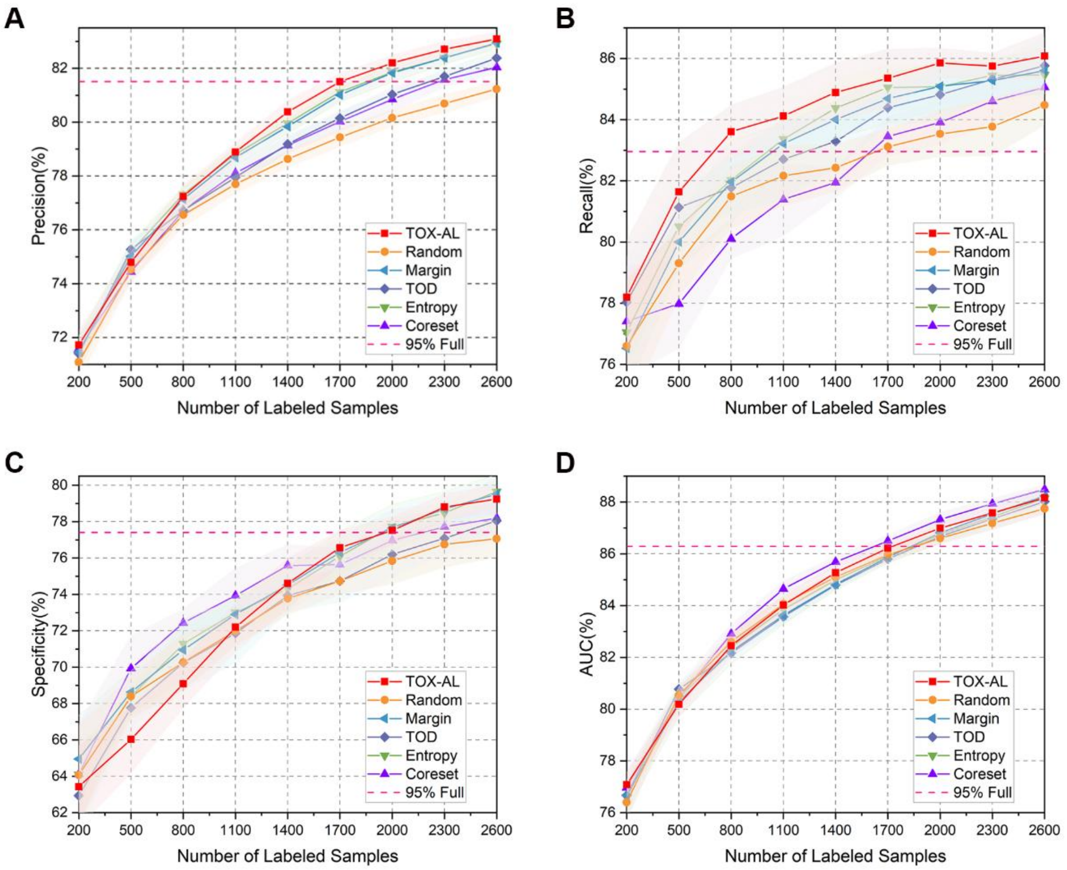
Multi-feature fusion makes performance better.

In TOX-AL, we use four molecular fingerprints MACCS, ECFP2, ECFP4, ECFP6 and a molecular descriptor RDKit2D as inputs to the backbone module. To investigate the effectiveness of different feature fusions on the overall classification performance, we design five ablation experiments using different features. The accuracy and F1-score of the last active learning cycle are compared, and the results are shown in Supplementary Figure 1.

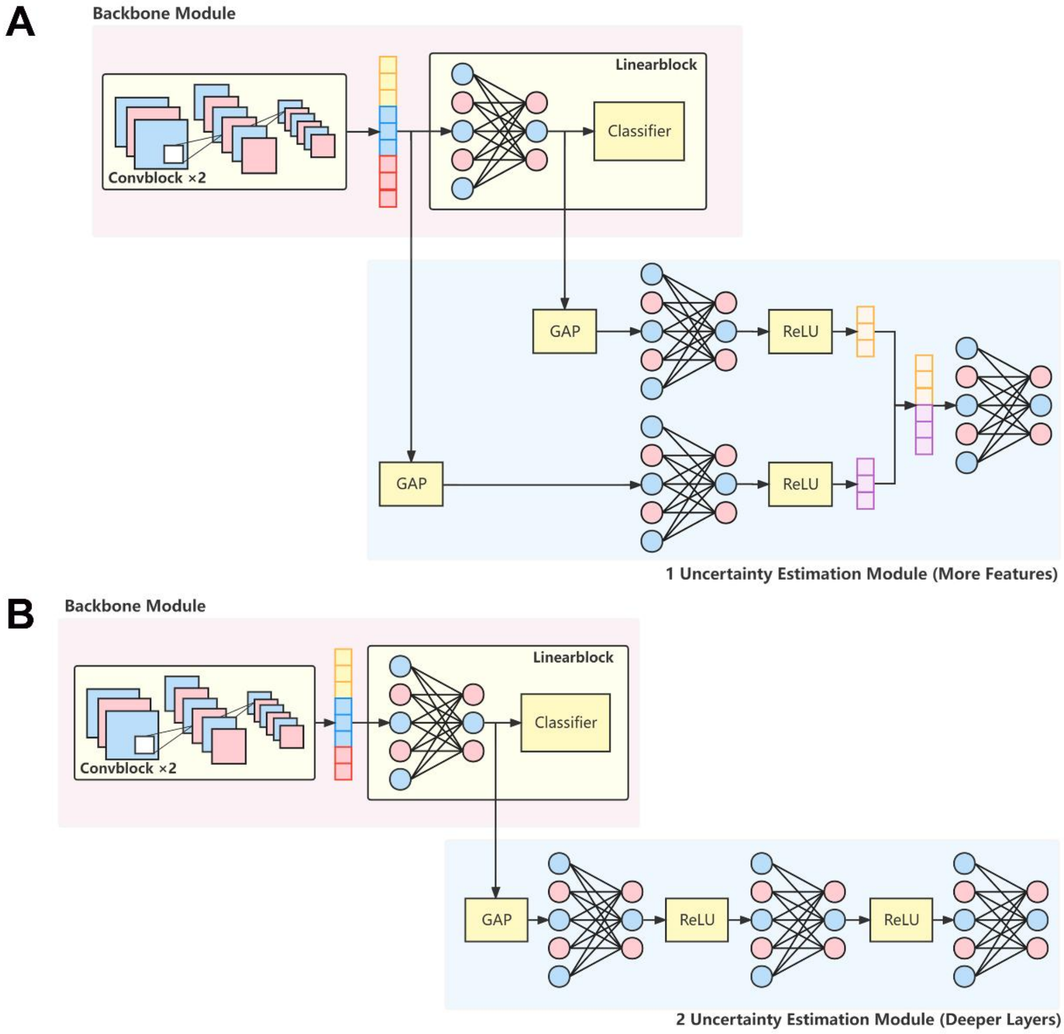
We can see from Supplementary Figure 1 that the best performance is achieved when all features are used, indicating that each type of feature is essential in the mutagenicity prediction task. The worst performance is achieved using only three fusion ECFPs, meaning the lowest contribution. Better results are obtained using only RDKit2D and RDKit2D with MACCS, demonstrating their higher importance. Furthermore, the two metrics decrease to varying degrees when neither MACCS nor RDKit2D is used, and it is more pronounced without MACCS, indicating that MACCS contributes more to prediction than RDKit2D.



Supplementary Figure 1. Accuracy (blue histogram) and F1-score (orange histogram) for different input features in the last active learning cycle.



Supplementary Figure 2. Active learning results of mutagenicity classification of TOX-AL and five active learning strategy baselines. (A) Precision of the six algorithms in different active learning cycles. (B) F1-score of the six algorithms in different active learning cycles. (C) Sensitivity of the six algorithms in different active learning cycles. (D) AUC of the six algorithms in different active learning cycles. The red line represents the TOX-AL proposed in this paper, the dark blue line represents the TOD active learning strategy, the purple line represents the Core-set active learning strategy, the blue line represents the margin active learning strategy, the green line represents the entropy active learning strategy, and the orange line represents random strategy.



Supplementary Figure 3. The schematic diagram of the two uncertainty estimation modules. (A) Using two different hidden layer features of the backbone module. (B) The same hidden layer features as in Figure 2, but with an additional linear layer and ReLU activation function.

Supplementary Table 1. The accuracy of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	71.72±0.56	71.09±0.66	71.58±0.47	71.37±0.65	71.41±0.43	71.46±0.60
2	74.79±0.41	74.52±0.45	74.44±0.42	75.19±0.40	75.27±0.34	75.02±0.44
3	77.24±0.42	76.56±0.46	76.73±0.31	77.32±0.35	76.72±0.29	77.14±0.42
4	78.89±0.36	77.7±0.35	78.11±0.40	78.81±0.28	77.95±0.32	78.69±0.33
5	80.38±0.43	78.63±0.32	79.14±0.31	79.96±0.32	79.19±0.33	79.84±0.26
6	81.50±0.32	79.44±0.35	80.02±0.36	81.12±0.32	80.15±0.27	81.01±0.25
7	82.20±0.30	80.16±0.31	80.85±0.46	81.86±0.19	81.03±0.16	81.83±0.33
8	82.71±0.31	80.69±0.34	81.58±0.28	82.39±0.20	81.7±0.27	82.39±0.24
9	83.09±0.26	81.23±0.27	82.04±0.23	82.92±0.29	82.38±0.27	82.94±0.19

Supplementary Table 2. The F1-score of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	75.59±0.68	74.79±0.78	75.3±0.52	75.08±0.64	75.23±0.65	75.01±0.51
2	78.39±0.52	77.71±0.49	77.36±0.50	78.44±0.45	78.53±0.40	78.21±0.41
3	80.45±0.36	79.58±0.41	79.42±0.31	80.21±0.35	79.67±0.29	80.08±0.45
4	81.70±0.33	80.51±0.41	80.65±0.42	81.51±0.26	80.71±0.49	81.41±0.31
5	82.90±0.39	81.22±0.32	81.50±0.20	82.52±0.30	81.72±0.43	82.37±0.22
6	83.80±0.35	81.93±0.33	82.41±0.32	83.48±0.29	82.61±0.30	83.34±0.23
7	84.40±0.26	82.52±0.24	83.09±0.39	84.02±0.19	83.33±0.24	84.01±0.27
8	84.76±0.26	82.95±0.33	83.74±0.25	84.47±0.19	83.91±0.31	84.45±0.22
9	85.10±0.28	83.46±0.23	84.15±0.20	84.88±0.25	84.48±0.31	84.91±0.17

Supplementary Table 3. The Precision of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	71.72±0.56	71.09±0.66	71.58±0.47	71.37±0.65	71.41±0.43	71.46±0.6
2	74.79±0.41	74.52±0.45	74.44±0.42	75.19±0.40	75.27±0.34	75.02±0.44
3	77.24±0.42	76.56±0.46	76.73±0.31	77.32±0.35	76.72±0.29	77.14±0.42
4	78.89±0.36	77.7±0.35	78.11±0.40	78.81±0.28	77.95±0.32	78.69±0.33
5	80.38±0.43	78.63±0.32	79.14±0.31	79.96±0.32	79.19±0.33	79.84±0.26
6	81.5±0.32	79.44±0.35	80.02±0.36	81.12±0.32	80.15±0.27	81.01±0.25
7	82.2±0.30	80.16±0.31	80.85±0.46	81.86±0.19	81.03±0.16	81.83±0.33
8	82.71±0.31	80.69±0.34	81.58±0.28	82.39±0.20	81.7±0.27	82.39±0.24
9	83.09±0.26	81.23±0.27	82.04±0.23	82.92±0.29	82.38±0.27	82.94±0.19

Supplementary Table 4. The Recall of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	78.20±1.87	76.60±2.21	77.41±1.83	77.05±1.25	78.03±1.49	76.53±1.42
2	81.64±1.61	79.31±1.11	77.98±1.45	80.52±0.89	81.13±0.72	80.00±0.86
3	83.61±0.88	81.50±1.06	80.11±0.66	82.03±0.90	81.78±0.90	81.97±1.34
4	84.12±0.94	82.17±0.99	81.39±1.19	83.36±0.69	82.70±1.33	83.22±0.65
5	84.89±0.97	82.43±0.78	81.95±0.50	84.38±0.97	83.29±0.88	84.00±0.60
6	85.36±0.96	83.11±0.61	83.45±0.65	85.05±0.78	84.39±0.59	84.69±0.81
7	85.85±0.50	83.54±0.75	83.91±0.74	85.08±0.70	84.82±0.84	85.09±0.78
8	85.75±0.43	83.77±0.96	84.61±0.70	85.45±0.47	85.32±0.54	85.28±0.68
9	86.08±0.77	84.48±0.66	85.06±0.56	85.47±0.49	85.76±0.95	85.61±0.73

Supplementary Table 5. The Specificity of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	63.43±2.26	64.07±2.74	64.11±2.75	64.10±1.58	62.94±1.87	64.95±2.39
2	66.03±1.86	68.39±1.17	69.93±1.79	68.38±0.49	67.77±1.04	68.63±1.15
3	69.09±1.34	70.27±1.57	72.43±0.77	71.28±1.15	70.24±1.51	70.95±1.67
4	72.20±1.35	71.99±0.81	73.94±1.51	73.01±1.00	71.87±1.64	72.90±0.83
5	74.61±1.39	73.78±0.89	75.59±1.17	74.30±1.35	73.94±0.92	74.52±0.98
6	76.57±0.97	74.73±0.73	75.64±0.90	76.10±1.07	74.73±1.17	76.30±1.16
7	77.52±0.77	75.84±1.27	76.96±1.21	77.73±0.95	76.19±1.27	77.65±1.28
8	78.80±0.76	76.75±1.26	77.72±1.03	78.48±0.58	77.08±0.93	78.70±0.96
9	79.24±0.77	77.07±1.02	78.18±0.87	79.66±0.76	78.05±1.11	79.52±1.08

Supplementary Table 6. The AUC of TOX-AL with five active learning baselines

Cycle	TOX-AL	Random	Coreset	Entropy	TOD	Margin
1	77.08±0.50	76.40±0.69	76.96±0.62	76.61±0.65	76.67±0.46	76.66±0.61
2	80.19±0.51	80.53±0.32	80.43±0.52	80.44±0.35	80.77±0.33	80.30±0.56
3	82.45±0.48	82.61±0.38	82.91±0.39	82.41±0.32	82.17±0.16	82.22±0.46
4	84.02±0.38	84.07±0.29	84.65±0.31	83.85±0.28	83.57±0.28	83.63±0.33
5	85.27±0.36	85.10±0.20	85.69±0.24	84.99±0.29	84.80±0.16	84.82±0.22
6	86.22±0.27	85.97±0.30	86.50±0.23	85.92±0.26	85.81±0.13	85.87±0.22
7	86.99±0.26	86.60±0.22	87.32±0.29	86.74±0.21	86.67±0.23	86.81±0.38
8	87.58±0.27	87.18±0.25	87.94±0.22	87.43±0.21	87.35±0.22	87.54±0.41
9	88.16±0.24	87.75±0.21	88.48±0.19	88.14±0.21	88.02±0.21	88.23±0.26

Supplementary Table 7. The overall performance between different kinds of backbone module using all samples

Methods	Accuracy	F1	Precision	Recall	Specificity	AUC
Adaboost	80.20±0.05	82.52±0.05	81.75±0.06	83.32±0.04	76.22±0.08	86.52±0.10
RF	84.10±0.18	85.91±0.17	85.36±0.16	86.49±0.23	81.06±0.23	**91.70±0.05**
GBDT	80.01±0.03	82.34±0.03	81.59±0.03	83.14±0.03	76.04±0.04	87.33±0.02
GCN	82.47±0.24	84.37±0.23	84.41±0.44	84.38±0.56	80.18±0.79	89.64±0.11
TOX-AL	**84.76±0.17**	**86.54±0.18**	**85.80±0.32**	**87.33±0.55**	**81.47±0.56**	90.84±0.10

Supplementary Table 8. The sets of molecules with similar structures but opposite properties selected by each active learning cycle.

Cycle	TAID	Label	Canonical SMILES	InChIKey
	TOX-87607	0	Nc1ccc(Cl)c(Cl)c1Cl	RRJUYQOFOMFVQS-UHFFFAOYSA-N
	TOX-89106	0	Nc1cc(Cl)c(Cl)c(Cl)c1	XOGYQVITULCUGU-UHFFFAOYSA-N
	TOX-7784	1	Nc1ccc(Cl)cc1Cl	KQCMTOWTPBNWDB-UHFFFAOYSA-N
	TOX-5093	0	Nc1ccc(Cl)c(Cl)c1	SDYWXFYBZPNOFX-UHFFFAOYSA-N
	TOX-5097	0	Nc1cc(Cl)ccc1Cl	AVYGCQXNNJPXSS-UHFFFAOYSA-N
	TOX-5080	0	Nc1ccccc1Cl	AKCRQHGQIJBRMN-UHFFFAOYSA-N
	TOX-89106	0	Nc1cc(Cl)c(Cl)c(Cl)c1	XOGYQVITULCUGU-UHFFFAOYSA-N
cycle1	TOX-86842	0	Clc1cccc2oc3ccccc3c12	WRSMJZYBNIAAEE-UHFFFAOYSA-N
	TOX-87598	1	Clc1ccc2c(c1)oc1ccccc12	BBOZMMAURMEVAR-UHFFFAOYSA-N
	TOX-89767	0	Clc1cccc2c1oc1ccccc12	RHRYBWFAHXCUCR-UHFFFAOYSA-N
	TOX-89271	1	Clc1ccc2oc3ccccc3c2c1	PRKTYWJFCODJOA-UHFFFAOYSA-N
	TOX-4091	1	N=C(N)c1ccc(OCCCCCOc2ccc(C(=N)N)cc2)cc1	XDRYMKDFEDOLFX-UHFFFAOYSA-N
	TOX-4943	0	Clc1ccc2ccccc2c1	CGYGETOMCSJHJU-UHFFFAOYSA-N
	TOX-86309	1	CC(O)=Nc1ccc(C=CC(=O)c2ccccc2)cc1	MJYUMWOKEJNYEM-UHFFFAOYSA-N
	TOX-86764	0	CC(O)=Nc1ccc(C(=O)C=Cc2ccccc2)cc1	CUYJTAUCYGHYLW-UHFFFAOYSA-N
	TOX-90321	1	Cc1ccc2nc3c(ccc4ccccc43)cc2c1	AQGXIFMUJLIPBG-UHFFFAOYSA-N
cycle2	TOX-88554	1	Cc1c2ccccc2nc2c1ccc1ccccc12	LZUFUDJVBHLHFC-UHFFFAOYSA-N
	TOX-90346	1	Cc1cccc2nc3c(ccc4ccccc	YZUYUDXHUYPYGT-U

			43)cc12	HFFFAOYSA-N
	TOX-87874	0	Cc1cccc2cc3ccc4ccccc4c3nc12	WHCOTCZBAOPNIV-UHFFFAOYSA-N
	TOX-89619	0	Cn1c(N)nc2nc3cccnc3cc21	XPPWPWDVLBEMIF-UHFFFAOYSA-N
	TOX-88211	1	Cn1c(N)nc2nc3ccccc3cc21	NWZPKLJMJMZZEO-UHFFFAOYSA-N
	TOX-87763	1	Cn1c(N)nc2nc3cnccc3cc21	ISTKKTAIKRCSBN-UHFFFAOYSA-N
	TOX-87730	1	CC(=O)Nc1cccc2cccnc12	YBWWRMMAWZGAPQ-UHFFFAOYSA-N
	TOX-89403	0	CC(=O)Nc1cnc2ccccc2c1	VAECZJVLTOQMGP-UHFFFAOYSA-N
	TOX-89607	1	CC/C=C/C=O	DTCCTIQRPGSLPT-ONEGZZNKSA-N
	TOX-87239	0	CC/C=C/C=C/C=C/C=C/C=O	WUGCJUPOLHRZFG-BYFNFPHLSA-N
	TOX-87136	1	Clc1c2ccccc2cc2ccccc12	KULLJOPUZUWTMF-UHFFFAOYSA-N
	TOX-87624	0	Clc1cccc2cc3ccccc3cc12	SRIHSAFSOOUEGL-UHFFFAOYSA-N
	TOX-89058	1	Clc1cc2ccccc2c2ccccc12	CJWHZQNUDAJJSB-UHFFFAOYSA-N
cycle3	TOX-80973	1	O=[N+]([O-])c1ccc(Cl)c(Cl)c1Cl	BGKIECJVXXHLDP-UHFFFAOYSA-N
	TOX-4887	1	O=[N+]([O-])c1cc(Cl)c(Cl)cc1Cl	IBRBMZRLVYKVRF-UHFFFAOYSA-N
	TOX-83355	0	O=[N+]([O-])c1cc(Cl)c(Cl)c(Cl)c1Cl	MTBYTWZDRVOMBR-UHFFFAOYSA-N
	TOX-4723	0	O=[N+]([O-])c1c(Cl)c(Cl)c(Cl)c(Cl)c1Cl	LKPLKUMXSAEKID-UHFFFAOYSA-N
	TOX-5892	0	O=[N+]([O-])c1c(Cl)c(Cl)cc(Cl)c1Cl	XQTLDIFVVHJORV-UHFFFAOYSA-N
	TOX-86809	1	Fc1cnc2ccc3cccnc3c2c1	UEXDUNFEZBIYAH-UHFFFAOYSA-N
	TOX-87558	0	Fc1cnc2c(ccc3ncc(F)cc32)c1	OMYJICXWQGYTMH-UHFFFAOYSA-N
	TOX-87690	1	Fc1cc2cccnc2c2cccnc12	UUADHOPXMOODSO-U

				HFFFAOYSA-N
	TOX-88922	1	Fc1ccc2ccc3cccnc3c2c1	OOQHWHBPBYDOAU-UHFFFAOYSA-N
	TOX-88775	0	Fc1cnc2c(ccc3ccccc32)c1	DGXWXKVFHBWVQP-UHFFFAOYSA-N
	TOX-90170	1	Fc1cnc2ccc3ccccc3c2c1	QLVVFCDIRRCTND-UHFFFAOYSA-N
	TOX-87659	0	CC(=O)Nc1cccc2ncccc12	JXDASSXQWMYQGN-UHFFFAOYSA-N
	TOX-87941	1	CC(=O)Nc1ccc2cccnc2c1	MWWLXWJCCZTCCQ-UHFFFAOYSA-N
	TOX-87943	1	CC(=O)Nc1snc2ccc(Cl)cc12	NDQCLZBDMKRDBW-UHFFFAOYSA-N
	TOX-88290	0	CC(=O)Nc1snc2c(Cl)cc(Cl)cc12	WYUCDVXYCQOKMM-UHFFFAOYSA-N
cycle4	TOX-87811	0	Cc1cccc2cc3ccccc3cc12	KZNJSFHJUQDYHE-UHFFFAOYSA-N
	TOX-84672	1	Cc1cccc2c1ccc1ccccc12	DOWJXOHBNXRUOD-UHFFFAOYSA-N
	TOX-87807	1	Cc1ccc2cc3ccccc3cc2c1	GYMFBYTZOGMSQJ-UHFFFAOYSA-N
	TOX-113045	1	[O-][NH+](O)c1c(Cl)c(Cl)cc(Cl)c1Cl	na
	TOX-113044	0	[O-][NH+](O)c1c(Cl)c(Cl)c(Cl)c(Cl)c1Cl	na
	TOX-4884	1	O=[N+]([O-])c1cc(Cl)ccc1Cl	RZKKOBGFCAHLCZ-UHFFFAOYSA-N
	TOX-80942	0	O=[N+]([O-])c1cccc(Cl)c1Cl	CMVQZRLQEOAYSW-UHFFFAOYSA-N
	TOX-80943	1	O=[N+]([O-])c1ccc(Cl)cc1Cl	QUIMTLZDMCNYGY-UHFFFAOYSA-N
cycle5	TOX-87001	1	Clc1ccc2ncccc2c1	GKJSZXGYFJBYRQ-UHFFFAOYSA-N
	TOX-88825	0	Clc1ccc2ccccc2n1	OFUFXTHGZWIDDB-UHFFFAOYSA-N
	TOX-87803	1	Clc1cccc2ncccc12	HJSRGOVAIOPERP-UHFFFAOYSA-N
	TOX-88716	0	Clc1cnc2ccccc2c1	FLORQCMDMHHIHN-U

				HFFFAOYSA-N
	TOX-113053	1	[O-][NH+](O)c1cc(Cl)c(Cl)cc1Cl	na
	TOX-113052	0	[O-][NH+](O)c1cc(Cl)c(Cl)c(Cl)c1Cl	na
	TOX-87130	1	Fc1ccc(F)c2ncccc12	RUGDURZULGJEHD-UHFFFAOYSA-N
	TOX-88313	1	Fc1cc(F)c2ncccc2c1	TWMLBTXFCAIQRJ-UHFFFAOYSA-N
	TOX-88083	0	Fc1cnc2cccc(F)c2c1	FLPAHYRBIDJSAP-UHFFFAOYSA-N
	TOX-88536	0	Fc1ccc2ccccc2n1	na
	TOX-87336	0	CCN(CC)c1ccc(-c2ccc(-c3ccc(-c4ccc(N(CC)CC)cc4)cc3)cc2)cc1	OVVJZWNHLPSJOT-UHFFFAOYSA-N
	TOX-88409	1	CCN(CC)c1ccc(-c2ccc(-c3ccc(N(CC)CC)cc3)cc2)cc1	PWYBEZFEEJBETN-UHFFFAOYSA-N
cycle6	TOX-87627	1	OCc1ccc2nc3c(ccc4ccccc43)c(CO)c2c1	TYYRHOGUEROKSQ-UHFFFAOYSA-N
	TOX-89293	0	OCc1ccc2c(CO)c3ccc4ccccc4c3nc2c1	KNRURXFLNGOFAI-UHFFFAOYSA-N
	TOX-88946	1	CCc1cccc2cccnc12	WSNATRDCOFYLCB-UHFFFAOYSA-N
	TOX-87641	0	CCc1ccnc2ccccc12	UOYJBGYNFMPCLV-UHFFFAOYSA-N
cycle7	TOX-4883	1	Cc1ccc(C)c([N+](=O)[O-])c1	BSFHJMGROOFSRA-UHFFFAOYSA-N
	TOX-83843	1	Cc1ccc([N+](=O)[O-])c(C)c1	BBUPBICWUURTNP-UHFFFAOYSA-N
	TOX-81700	1	Cc1cccc([N+](=O)[O-])c1C	FVHAWXWFPBPFOS-UHFFFAOYSA-N
	TOX-4696	0	Cc1cccc(C)c1[N+](=O)[O-]	HDFQKJQEWGVKCQ-UHFFFAOYSA-N
	TOX-87347	1	CC(=O)Nc1ccc(/C=C/C(=O)c2ccccc2)cc1	MJYUMWOKEJNYEM-FMIVXFBMSA-N
	TOX-87029	0	CC(=O)Nc1ccc(C(=O)/C=C/c2ccccc2)cc1	CUYJTAUCYGHYLW-KPKJPENVSA-N

	TOX-86884	0	O=C(OCc1cccc([N+](=O)[O-])c1)c1ccccc1	XYHCAQKUZAEVJN-UHFFFAOYSA-N
	TOX-86444	1	O=C(OCc1ccc([N+](=O)[O-])cc1)c1ccccc1	CRTADXBPEXSZLH-UHFFFAOYSA-N
cycle8	TOX-31595	1	NCCNCCNCCNCCN	FAGUFWYHJQFNRV-UHFFFAOYSA-N
	TOX-5770	0	NCCNCCN	RPNUMPOLZDHAAY-UHFFFAOYSA-N
	TOX-13576	1	NCCNCCNCCNCCNCCN	LSHROXHEILXKHM-UHFFFAOYSA-N
	TOX-86731	0	O=C(O)Cc1cccc([N+](=O)[O-])c1	WUKHOVCMWXMOOA-UHFFFAOYSA-N
	TOX-5420	1	O=C(O)Cc1ccc([N+](=O)[O-])cc1	YBADLXQNJCMBKR-UHFFFAOYSA-N
	TOX-113078	0	[O-][NH+](O)c1ccc(O)c([NH+]([O-])O)c1	na
	TOX-113056	1	[O-][NH+](O)c1cc([NH+]([O-])O)c(O)c([NH+]([O-])O)c1	na
	TOX-89498	1	O=C(Cl)c1ccccc1Cl	ONIKNECPXCLUHT-UHFFFAOYSA-N
	TOX-81420	1	O=C(Cl)c1c(Cl)cccc1Cl	JBLIDPPHFGWTKU-UHFFFAOYSA-N
	TOX-4890	1	O=C(Cl)c1ccc(Cl)cc1Cl	CEOCVKWBUWKBKA-UHFFFAOYSA-N
	TOX-81537	0	O=Cc1ccc(Cl)cc1Cl	YSFBEAASFUWWHU-UHFFFAOYSA-N
	TOX-83375	0	O=Cc1ccc(Cl)c(Cl)c1	ZWUSBSHBFFPRNE-UHFFFAOYSA-N
cycle9	TOX-88839	0	OC1c2ccc3nc4c(ccc5ccccc54)cc3c2C2OC2C1O	XPMBTWVVTCOLCP-UHFFFAOYSA-N
	TOX-87403	0	OC1c2c(ccc3ncccc23)C2OC2C1O	WAQXNJCLUIKZTK-UHFFFAOYSA-N
	TOX-88005	1	OC1c2ccc3nc4ccc5ccccc5c4cc3c2C2OC2C1O	PIJVASQAQLIDRI-UHFFFAOYSA-N
	TOX-86709	0	O=[N+]([O-])c1cccc([N+](=O)[O-])c1CO	JOFAFRXWWUXIEE-UHFFFAOYSA-N
	TOX-86378	1	O=[N+]([O-])c1ccc(CO)	ZKKOBDGQUYKWFN-U

		c([N+](=O)[O-])c1	HFFFAOYSA-N
TOX-87174	1	O=C(c1ccc(F)cc1)[C@H]1O[C@@H]1c1ccccc1	XHUCGUZYMJDQMH-HUUCEWRRSA-N
TOX-87824	0	O=C(c1ccccc1)C1OC1c1ccc(F)cc1	JSNXRAUQMWIZHM-UHFFFAOYSA-N
TOX-89884	1	O=C(c1ccc(F)cc1)C1OC1c1ccccc1	XHUCGUZYMJDQMH-UHFFFAOYSA-N
TOX-86759	0	O=C(C=Cc1cccc([N+](=O)[O-])c1)c1ccccc1	SMFBODMWKWBFOK-UHFFFAOYSA-N
TOX-87813	0	O=C(/C=C/c1cccc([N+](=O)[O-])c1)c1ccccc1	SMFBODMWKWBFOK-MDZDMXLPSA-N
TOX-86550	1	O=C(C=Cc1ccccc1)c1ccc([N+](=O)[O-])cc1	QLXBMZAOTHBHSM-UHFFFAOYSA-N
TOX-81239	1	O=C(/C=C/c1ccc([N+](=O)[O-])cc1)c1ccccc1	WDZGGAFMGIOIQS-DHZHZOJOSA-N
TOX-72298	1	O=C(C=Cc1ccc([N+](=O)[O-])cc1)c1ccccc1	WDZGGAFMGIOIQS-UHFFFAOYSA-N
TOX-89575	0	COc1ccc(C(=O)C2OC2c2ccccc2)cc1	GAMOPGGZIQWBEM-UHFFFAOYSA-N
TOX-87129	1	COc1ccc([C@H]2O[C@@H]2C(=O)c2ccccc2)cc1	OJNVQVXKXBOSSW-HZPDHXFCSA-N
TOX-89800	1	COc1ccc(C2OC2C(=O)c2ccccc2)cc1	OJNVQVXKXBOSSW-UHFFFAOYSA-N