SUPPLEMENTARY MATERIAL

The index of ideality of correlation: Improvement of models for toxicity to algae

Alla P. Toropova*, Andrey A. Toropov

Department of Environmental Health Science, Laboratory of Environmental Chemistry and Toxicology, IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Via La Masa 19, 20156 Milano, Italy

Abstract

Toxicity to algae is important characteristic of substances from ecologic point of view. The CORAL software (http://www.insilico.eu/coral) gives possibility to build up model of toxicity to algae using data on the molecular architecture and experimental toxicity, without additional data on physicochemical and/or biochemical parameters. Considerable improvement of the model is observed in the case of using the index of ideality of correlation (*IIC*) in the role of additional criterion of predictive potential. The *IIC* is calculated with using of the correlation coefficient between experimental and calculated values of endpoint for the calibration set, with taking into account the positive and negative dispersions between experimental and calculated values. The best model calculated with use the *IIC* is characterized (the validation set) by n=50, r^2=0.947, RMSE=0.401 whereas, model calculated without use the *IIC* is characterized by n=50, r^2=0.805, and RMSE=0.539. The suggested models are built up in accordance to five OECD principles.

Containing

Table S14
Correlation weights of model based on Split 2

Table S15
Correlation weights of model based on Split 3

Table S1
Split 1: Distribution into training (+), invisible training (-), calibration (#), and validation (*) sets; ID; SMILES; Optimal descriptor (DCW); Experimental (Expr) and Calculated (Calc) toxicity to algae (pEC50); Defect-SMILES; domain of applicability

Set	ID	SMILES	DCW	Expr	Calc	DefectSMILES	Applicability
-	1	O(CCO)C	-1.59538	0.5000	1.9291	0.0183	YES
-	2	O=C(CCCCCCCCCCC)C	10.78879	6.2200	5.3266	1.0340	YES
#	3	O(CCOCC)CCO	-1.14837	0.9600	2.0517	0.0213	YES
-	4	O=C(C)C	-0.04105	0.9600	2.3555	1.0236	YES
#	5	O(CCO)CC	-0.51240	1.0000	2.2262	0.0194	YES
+	6	O(CC)CC	0.54648	1.5100	2.5167	0.0134	YES
#	7	OCC(C)C	-2.38607	1.6400	1.7122	0.0128	YES
#	8	OC(C)(C)C	-0.82381	1.6600	2.1408	0.0173	YES
-	9	O(CCCC)CCO	1.29064	1.8100	2.7208	0.0194	YES
+	10	O(CCOCCO)CCCC	1.38053	2.1700	2.7455	0.0254	YES
#	11	O1C(C)(C)CCC1(C)C	4.97448	2.3400	3.7315	0.0289	YES
-	12	ClCCCl	3.88089	2.6300	3.4315	0.0105	YES
*	13	O(c(c(O)ccc1)c1)C	2.96853	2.6600	3.1812	0.0232	YES
*	14	C(Cl)(Cl)Cl	4.48385	2.7100	3.5969	0.0214	YES
*	15	OCCCCCC	2.15086	2.9500	2.9568	0.0129	YES
#	16	C(C=C)(=C)C	5.91695	3.0100	3.9900	0.0379	YES
*	17	ClCC(Cl)Cl	3.99482	3.1100	3.4627	0.0185	YES
-	18	N(CC)CC	4.75705	3.1300	3.6718	0.0185	YES
*	19	ClCC(Cl)CCl	4.40670	3.2700	3.5757	0.0184	YES

-	20	ClCCCC	4.81131	3.4100	3.6867	0.0094	YES
#	21	c(cccc1)c1	5.04432	3.4300	3.7506	0.0079	YES
-	22	C(C(Cl)Cl)(Cl)Cl	5.00966	3.4900	3.7411	0.0294	YES
+	23	O=C(O)CCCCCCC	6.56186	3.5700	4.1669	1.0373	YES
*	24	c(cccc1)(c1)Cl	6.28130	3.5800	4.0900	0.0121	YES
-	25	O=C(O)C(=O)O	-0.76359	3.6100	2.1573	1.0600	YES
#	26	C(CCCC1)C1	3.82015	3.6400	3.4148	0.0114	YES
+	27	BrCCCBr	4.76999	3.6400	3.6754	1.0212	YES
-	28	O(CCCC)CCCC	4.87842	3.7700	3.7051	0.0176	YES
-	29	O(c(c(OC)cc(c1)CC=C)c1)C	7.74934	3.9100	4.4927	1.0333	YES
#	30	c(c(ccc1)Cl)(c1)Cl	7.66776	4.0100	4.4703	0.0184	YES
+	31	c(ccc(c1)C)(c1)C	7.23919	4.0400	4.3528	0.0135	YES
-	32	N(CC)(CC)CC	6.70154	4.1000	4.2053	0.0212	YES
-	33	C(=CCCC=CC1)C1	7.49560	4.1200	4.4231	0.0377	YES
-	34	BrC(Cl)Cl	6.83127	4.1400	4.2409	0.0270	YES
+	35	c(c(ccc1)Cl)(c1)C	6.84409	4.2100	4.2444	0.0151	YES
-	36	C(C(C=C1)CC=C2)(C2)C1	5.73267	4.2300	3.9395	0.0513	YES
#	37	COc1c(Cl)cc(OC)cc1	7.44570	4.2400	4.4094	0.0277	YES
*	38	O(c(ccc(c1)C(c(ccc(OCCO)c2)c2)(C)C)c1)CCO	8.61990	4.2700	4.7315	0.0491	YES
#	39	c(ccc(c1)Cl)(c1)C	6.84409	4.3200	4.2444	0.0151	YES
#	40	CCCCCCCCCc1ccc(cc1)OCCOCCOCCOCCOCCOCCO	9.29746	4.3300	4.9174	0.0502	YES
+	41	BrC(Br)Cl	5.79871	4.3400	3.9576	0.0364	YES
#	42	c(cccc1)(c1)CC	7.66652	4.3600	4.4700	0.0122	YES
#	43	c(ccc(c1)C)(c1)C(C)C	7.03420	4.3600	4.2965	0.0175	YES

#	44	OCCCCCCCCC(C)C	4.11183	4.3700	3.4948	0.0190	YES
+	45	C(C(Cl)Cl)(Cl)(Cl)Cl	8.38569	4.3800	4.6673	0.0389	YES
#	46	C(C(C=CC12)C1)(=CC)C2	8.39201	4.3900	4.6690	0.0491	YES
-	47	c(ccc(c1Cl)Cl)(c1)Cl	8.70401	4.5000	4.7546	0.0237	YES
*	48	COc1ccc(O)c(c1)C(C)(C)C	7.88975	4.5400	4.5312	0.0331	YES
+	49	c(cccc1Cl)(c1)C	6.49389	4.5500	4.1483	0.0142	YES
+	50	c(cccc1)(c1)C(C)C	6.37854	4.6600	4.1167	0.0152	YES
-	51	O(Cc(cccc1)c1)Cc(cccc2)c2	4.73787	4.6800	3.6666	0.0205	YES
+	52	c(c(c(cc1)C)ccc2)(c2)c1	10.47468	4.7100	5.2404	0.0178	YES
+	53	c(c(c(cc1)Cl)C)(c1)Cl	8.32341	4.7800	4.6502	0.0207	YES
+	54	c(ccc(c1Cl)C)(c1)Cl	7.97321	4.8000	4.5541	0.0197	YES
-	55	OCCCCCCCCC	5.39981	4.8200	3.8482	0.0160	YES
+	56	C=Cc1ccccc1C=C	10.13896	4.8600	5.1483	0.0352	YES
-	57	c(cccc1)(c1)CCCC	9.83248	4.9200	5.0642	0.0143	YES
-	58	c(ccc(c1C)Cl)(c1)Cl	8.70565	4.9800	4.7551	0.0194	YES
-	59	c(c(c(cc1)Cl)Cl)(c1)Cl	9.05421	5.0500	4.8507	0.0247	YES
#	60	c(ccc(c1Cl)Cl)(c1)C	7.88034	5.0600	4.5287	0.0204	YES
*	61	c(c(ccc1)C)(c1)C	7.23919	5.1200	4.3528	0.0135	YES
+	62	c(cc(ccc1C)c2c1)c(c2)C	9.92648	5.1900	5.0900	0.0165	YES
-	63	CCC1CCCCC1	8.49304	5.2500	4.6967	0.0119	YES
#	64	Cc2cc(C)c1ccccc1c2	11.25993	5.4000	5.4558	0.0148	YES
+	65	C(CCCC1)(C1)C	5.35936	5.4600	3.8371	0.0146	YES
-	66	O(c(cccc1)c1)c(cccc2)c2	5.89133	5.4700	3.9830	0.0180	YES
+	67	C(Cl)(Cl)(Cl)Cl	7.85989	5.5200	4.5230	0.0309	YES
-	68	O(C(C)C)CCO	-2.163	1.2100	1.7733	0.0203	YES

			30				
+	69	OCCCC	-0.01511	1.6800	2.3626	0.0108	YES
-	70	O=C(CCCC1)C1	2.71244	1.9200	3.1109	1.0259	YES
-	71	ClCCl	2.79791	2.0400	3.1343	0.0094	YES
+	72	OC(CC)CC	-1.30309	2.1300	2.0093	0.0138	YES
-	73	OCCCCC	1.06787	2.3800	2.6597	0.0119	YES
+	74	ClCC(Cl)C	3.17115	2.9100	3.2367	0.0153	YES
-	75	C(Cl)(Cl)C	3.66018	2.9700	3.3709	0.0181	YES
+	76	C(=CCl)(Cl)Cl	6.87550	3.2300	4.2530	0.0321	YES
-	77	O=C(O)CCCCCC	5.47888	3.3400	3.8698	1.0362	YES
+	78	N(CCCCCCN(C)C)(C)C	9.23221	3.3900	4.8995	0.0390	YES
+	79	c(cccc1)(c1)C	6.58353	3.5000	4.1729	0.0111	YES
+	80	OCCCCCCC	3.23384	3.5300	3.2539	0.0140	YES
#	81	BrCC(Br)CCl	7.29357	3.5800	4.3677	0.0374	YES
#	82	OCCCCCCCC	4.31682	3.6700	3.5510	0.0150	YES
#	83	C(=C(Cl)Cl)(Cl)Cl	6.98943	3.7900	4.2842	0.0401	YES
*	84	c(cccc1C)(c1)C	7.62143	4.0800	4.4576	0.0122	YES
*	85	c(cccc1)(c1)Br	4.81840	4.1200	3.6886	1.0175	YES
-	86	Fc1c(F)cc(Br)cc1	7.63216	4.1400	4.4606	1.0385	YES
*	87	CCCCCCCCCC(=O)O	8.74435	4.1600	4.7657	1.0396	YES
-	88	c(C(=C)C)(cccc1)c1	8.01759	4.3900	4.5663	0.0285	YES
+	89	c(ccc(c1)Cl)(c1)Cl	7.66776	4.4300	4.4703	0.0184	YES
-	90	O=C(CCCCCCCC)C	7.53984	4.5000	4.4352	1.0308	YES
+	91	c(cccc1Cl)(c1)Cl	7.31755	4.5800	4.3743	0.0174	YES
*	92	c(c(ccc1C)ccc2)(c2)c1	10.85	4.87	5.34	0.0166	YES

			692	00	52		
-	93	O=C(CCCCCCCCC)C	8.622 82	4.95 00	4.73 23	1.0319	YES
+	94	c(c(ccc1)ccc2)(c1CC3)c23	9.339 25	5.04 00	4.92 89	2.0311	No
+	95	O=C(c(cccc1)c1)c(c(O)cc(OC)c2)c2	8.693 94	5.53 00	4.75 19	1.0468	YES
-	96	Oc(cccc1)c1	3.947 21	2.68 00	3.44 96	0.0133	YES
+	97	N(c(c(c(ccc1)cc2)c1)c2)c(cccc3)c3	13.04 380	6.81 00	5.94 52	0.0500	YES
#	98	Nc(c(ccc1)C)c1	3.423 27	2.95 00	3.30 59	0.0148	YES
#	99	Nc(c(ccc1C)C)c1	4.461 16	3.23 00	3.59 06	0.0159	YES
+	100	Oc(ccc(c1)C)c1	4.602 87	3.27 00	3.62 95	0.0156	YES
-	101	Oc(c(ccc1)Cl)c1	6.177 61	3.39 00	4.06 15	0.0226	YES
*	102	Nc(ccc(c1)C)c1	3.423 27	3.40 00	3.30 59	0.0148	YES
*	104	Nc(c(cc(c1)C)C)c1	4.078 92	3.49 00	3.48 58	0.0171	YES
+	105	[O-][N+](=O)c(c(N)ccc1)c1	4.879 47	3.50 00	3.70 54	0.1239	YES
+	106	[O-][N+](=O)c(cccc1N)c1	4.454 85	3.51 00	3.58 89	1.1200	No
+	107	[O-][N+](=O)c(ccc(N)c1)c1	4.879 47	3.51 00	3.70 54	0.1239	YES
-	108	N(c(cccc1)c1)CC	7.501 42	3.56 00	4.42 47	0.0195	YES
-	109	Oc(ccc(c1C)C)c1	5.640 76	3.58 00	3.91 43	0.0167	YES
+	110	Oc(c(ccc1)CC)c1	5.685 85	3.59 00	3.92 66	0.0166	YES
-	111	Nc(cccc1C)c1	3.805 51	3.60 00	3.41 08	0.0135	YES
*	112	Oc(ccc(c1)Cl)c1	6.177 61	3.61 00	4.06 15	0.0226	YES
-	113	Nc(cc(cc1C)C)c1	4.461 16	3.62 00	3.59 06	0.0159	YES
+	114	Nc(cccc1Cl)c1	6.434 80	3.68 00	4.13 21	0.0184	YES
+	115	NCCCCCCN	6.394 19	3.79 00	4.12 10	0.0192	YES
+	116	Oc(c(cc(c1)C)C)c1	5.258 53	3.80 00	3.80 94	0.0180	YES
*	117	Nc(c(ccc1)Cl)c1Cl	7.195 52	3.84 00	4.34 08	0.0237	YES
#	11	CC(C)c1ccc(N)cc1	9.093	3.88	4.86	0.0187	YES

	8		66	00	15		
-	119	CCc1cccc(N)c1	8.72193	3.9300	4.7595	0.0159	YES
*	120	Nc(c(ccc1Cl)Cl)c1	7.82126	3.9900	4.5124	0.0247	YES
-	121	Oc(c(c(cc1C)C)C)c1	6.29642	4.0000	4.0941	0.0190	YES
-	122	Oc(c(ccc1C)C(C)C)c1	5.43577	4.0300	3.8580	0.0207	YES
-	123	Oc(cccc1Cl)c1	5.82741	4.0500	3.9655	0.0217	YES
+	124	Oc(ccc(c1)Cc(ccc(O)c2)c2)c1	7.95415	4.1000	4.5489	0.0317	YES
+	125	[O-][N+](=O)c(ccc(c1)C)c1	3.44862	4.1400	3.3129	0.1178	YES
-	126	Nc(ccc(c1)CC)c1	4.50625	4.1400	3.6030	0.0158	YES
-	127	Nc(ccc(c1C)C)c1	4.46116	4.1500	3.5906	0.0159	YES
*	128	Oc(c(cc(c1)C)C)c1C	5.57781	4.1500	3.8970	0.0187	YES
+	129	Oc(c(c(cc1)Cl)Cl)c1	7.56406	4.1700	4.4419	0.0289	YES
#	130	Nc(ccc(c1)Cc(ccc(N)c2)c2)c1	8.10039	4.2200	4.5890	0.0295	YES
+	131	Oc(ccc(c1)Br)c1	9.42977	4.2700	4.9537	1.0265	YES
-	132	Oc(c(cc(c1)Cl)Cl)c1	7.56406	4.3000	4.4419	0.0289	YES
+	133	Oc(ccc(c1)C(CC)C)c1	5.48086	4.3000	3.8704	0.0207	YES
-	134	[O-][N+](=O)c(c(N)ccc1Cl)c1	12.48681	4.3100	5.7924	1.1226	No
#	135	Oc(c(ccc1)C(CC)C)c1	5.48086	4.3400	3.8704	0.0207	YES
-	136	n(ccc(c1)C=C)c1	7.34795	4.3600	4.3826	0.0344	YES
-	137	Nc(c(c(cc1)Cl)Cl)c1	8.17146	4.3800	4.6085	0.0257	YES
+	138	Nc(ccc(c1Cl)Cl)c1	7.82126	4.4000	4.5124	0.0247	YES
+	139	Oc(c(ccc1)C(C)(C)C)c1C	7.36241	4.4200	4.3866	0.0260	YES
#	140	Oc(c(cc(c1Cl)Cl)Cl)c1	8.60031	4.4200	4.7262	0.0342	YES
-	141	[O-][N+](=O)c(c(O)ccc1Cl)c1	11.16097	4.4500	5.4287	1.1240	No
-	142	Oc(ccc(c(ccc(O)c1)c1)c2)c2	8.32839	4.5100	4.6516	0.0297	YES
*	14	Nc(c(cc(c(ccc(N)c1C)c1)c2)	10.16	4.53	5.15	0.0309	YES

	3	C)c2	818	00	63		
*	144	Nc(c(cc(c1)Cl)Cl)c1	8.17146	4.6100	4.6085	0.0257	YES
+	145	Oc(c(ccc1Cl)Cl)c1	7.21386	4.6500	4.3458	0.0279	YES
*	146	Oc1ccc(Cl)c(Cl)c1Cl	8.71338	4.6800	4.7572	0.0326	YES
+	147	Nc1ccc(Cl)c(Cl)c1Cl	9.32077	4.7400	4.9238	0.0294	YES
*	148	Nc(c(cc(c1Cl)Cl)Cl)c1	9.20771	4.8000	4.8928	0.0310	YES
#	149	Oc(ccc(c(cccc1)c1)c2)c2	7.56772	4.8500	4.4429	0.0199	YES
+	150	Oc(ccc(c1)CCCCC)c1	8.93480	4.8700	4.8179	0.0198	YES
*	151	[O-][N+](=O)c(ccc(c1Cl)Cl)c1	11.78676	4.8900	5.6003	1.1204	No
-	152	Oc1ccc(cc1)C(c2ccc(cc2)O)(c3ccc(cc3)O)C	14.17672	4.9000	6.2560	0.0762	YES
*	153	Oc(c(ccc1C)C(C)(C)C)c1	8.08101	4.9400	4.5837	0.0263	YES
+	154	Oc1cc(Cl)cc(Cl)c1Cl	8.71338	4.9400	4.7572	0.0326	YES
-	155	Oc(c(cc(c1)C)C(C)(C)C)c1	7.69877	4.9600	4.4788	0.0275	YES
*	156	Oc(c(cc(c1)C(C)(C)C)C(C)(C)C)c1	10.13902	5.3100	5.1483	0.0371	YES
-	157	Oc(cccc1C)c1	4.98511	2.8700	3.7344	0.0143	YES
*	158	Nc(cccc1)c1	2.76761	2.9300	3.1260	0.0124	YES
*	159	Oc(c(ccc1)C)c1	4.60287	2.9300	3.6295	0.0156	YES
#	160	Nc(c(ccc1)CC)c1	4.50625	3.3900	3.6030	0.0158	YES
+	161	Oc(c(ccc1)C)c1C	4.92216	3.4100	3.7171	0.0164	YES
-	162	Oc(c(c(cc1)C)C)c1	5.25853	3.4800	3.8094	0.0180	YES
#	163	Oc(c(ccc1C)C)c1	5.64076	3.5800	3.9143	0.0167	YES
-	164	Nc(c(c(cc1)C)C)c1	4.07892	3.5900	3.4858	0.0171	YES
-	165	Oc(cc(cc1C)C)c1	5.64076	3.6500	3.9143	0.0167	YES
-	166	Nc(c(cc(c1)C)C)c1C	4.39821	3.7000	3.5734	0.0179	YES
+	167	Oc(ccc(c1)CC)c1	5.68585	3.7500	3.9266	0.0166	YES
-	16	Oc(c(ccc1C)C)c1C	5.960	3.98	4.00	0.0175	YES

	8		05	00	18		
-	169	Oc(c(ccc1)Cl)c1Cl	6.58813	4.0100	4.1742	0.0270	YES
*	170	Oc(c(ccc1Cl)Cl)c1Cl	7.62437	4.3900	4.4584	0.0323	YES
+	171	Oc(ccc(C(=C)C)c1)c1	6.67073	4.4000	4.1968	1.0286	YES
#	172	Nc(cc(cc1Cl)Cl)c1	7.82126	4.5700	4.5124	0.0247	YES
-	173	Oc(ccc(c1ccc2)c2)c1	7.85855	4.8400	4.5227	0.0190	YES
-	174	Oc1ccc(Cl)c(Cl)c1	8.30286	4.8700	4.6446	0.0283	YES
*	175	Oc1cc(Cl)cc(Cl)c1	8.30286	4.8900	4.6446	0.0283	YES
#	176	Nc1cc(Cl)c(Cl)c(Cl)c1	10.29671	5.1400	5.1916	0.0314	YES
+	177	Oc(c(cc(c1)Br)Br)c1Br	10.61785	5.2400	5.2797	2.0505	No
#	178	CCC(C)c1cccc(C(C)CC)c1O	8.59395	5.2700	4.7244	0.0288	YES
+	179	Oc(c(cc(c1)Br)Br)c1	9.78367	5.3600	5.0508	1.0422	YES
+	180	N(c(cccc1)c1)c(cccc2)c2	10.10189	5.6000	5.1381	0.0230	YES
+	181	Oc(ccc(c1)CCCCCCCC)c1	12.18376	6.1700	5.7093	0.0229	YES
+	182	N#CC(=C)C	5.02437	3.4300	3.7452	1.0224	YES
*	183	CCc1cccc(CC)c1N(C(=O)CCl)CCOCCC	17.10168	7.9900	7.0584	2.0544	No
-	184	C(#N)C=C	4.24781	3.7200	3.5321	1.0223	YES
-	185	O=C1N(CCOC(=O)C=C)C(=O)N(C(=O)N1CCOC(=O)C=C)CCOC(=O)C=C	4.42277	4.2100	3.5801	2.2264	No
-	186	C1Oc2cc3N(CC)C=C(C(=O)O)C(=O)c3cc2O1	6.31544	4.2100	4.0993	2.1361	No
+	187	Fc1cc2C(=O)C(C(=O)O)=CN3c2c(CCC3C)c1	8.67355	4.7200	4.7463	5.1149	No
*	188	N(=Nc(cccc1)c1)c(ccc(N)c2)c2	10.03560	4.8300	5.1199	2.0494	No
+	189	N#CC(=C)Cl	13.89633	6.1000	6.1791	1.0256	YES
+	190	CCc1cccc(C)c1N(C(C)COC)C(=O)CCl	12.65763	6.6000	5.8393	2.0535	No
-	191	[O-][N+](=O)c(ccc(c1)CCl)c1	11.16239	6.6500	5.4291	1.1150	No
*	192	N(Nc(cccc1)c1)c(cccc2)c2	9.37771	5.2200	4.9394	0.0379	YES

-	193	N#CC=CC#N	6.70546	5.3100	4.2063	1.0247	YES
-	194	C(C=CC#N)(F)(F)F	7.69297	5.5300	4.4773	0.0632	YES
+	195	N(=C=S)C	10.28532	5.5900	5.1884	0.0436	YES
+	196	CCCCOCN(C(=O)CCl)c1c(CC)cccc1CC	17.72020	7.9800	7.2281	1.0574	YES
+	197	OCc(cccc1)c1	0.70517	2.1500	2.5602	0.0113	YES
-	198	Brc2c(c(cc(c2)OCC1OC1)C)Br	10.76274	5.7200	5.3194	1.0447	YES
-	199	OC(CCl)CCl	-0.24689	2.3100	2.2990	0.0219	YES
#	200	O(CCCl)CCCl	3.76865	2.6200	3.4007	0.0236	YES
+	201	S(SC)C	4.30398	3.5100	3.5475	0.0270	YES
+	202	O(C1c(cccc2)c2)C1	2.44684	3.6800	3.0380	0.0173	YES
+	203	O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F	7.08157	4.1600	4.3095	1.3076	No
+	204	C(=CCl)CCl	4.29876	4.7200	3.5461	0.0245	YES
-	205	C(Cl)(Cl)(Cl)C(Cl)COCC(Cl)C(Cl)(Cl)Cl	11.21517	5.5000	5.4435	0.0721	YES
-	206	C(C(Cl)CCl)=C	3.58901	3.3300	3.3514	0.0293	YES
+	207	c(cccc1)(c1)C(Cl)Cl	6.80711	3.7800	4.2342	0.0200	YES
+	208	S(SCC)CC	6.46995	4.4400	4.1417	0.0291	YES
*	209	c(c(ccc1)Cl)(c1)CCl	8.07964	5.1300	4.5833	0.0183	YES
+	210	O=C(Oc(c(ccc1)C(CC)C)c1)NC	9.82686	3.8000	5.0627	0.0608	YES
*	211	CC(C)Oc1cc(c(Cl)cc1Cl)N2N=C(OC2=O)C(C)(C)C	23.12259	6.7900	8.7102	3.0916	No
-	212	Cc1cccc(C)c1N(C(=O)COC)N2C(=O)OCC2	4.62068	4.5000	3.6344	1.0965	No
-	213	Oc(c(cc(c1)Cl)Cl)c1Cl	7.97458	4.5400	4.5545	0.0333	YES
#	214	Nc(c(cc(c1)Cl)Cl)c1Cl	8.58197	4.6000	4.7211	0.0300	YES
#	215	Oc(c(cc(c1Cl)Cl)Cl)c1Cl	9.01083	5.0400	4.8388	0.0386	YES
+	216	N(N)c(c(cc(c1)Cl)Cl)c1Cl	14.38226	6.4000	6.3124	0.0424	YES

+	217	c1ccc(C)cc1NC(=O)Oc2cccc(NC(=O)OC)c2	14.53885	6.4700	6.3553	1.1024	No
+	218	CNC(=O)Oc1ccccc1OC(C)C	9.79962	4.3200	5.0552	0.0688	YES
+	219	Oc(c(c(c(c1Cl)Cl)Cl)Cl)c1Cl	10.39728	5.4900	5.2192	0.0449	YES
+	220	COP(=S)(OC)Oc1ccc([N+](=O)[O-])c(C)c1	12.72583	5.7500	5.8580	5.1603	No
+	221	CCN(CC)C(=O)SCc1ccc(Cl)cc1	15.16410	6.4500	6.5269	1.0629	YES
+	222	N1C(=O)NC(=O)NC1=O	5.04608	2.1300	3.7511	2.1035	No
-	223	Clc1cc(C(F)(F)F)ccc1Oc2cc(OCC)c([N+](=O)[O-])cc2	24.33000	9.6600	9.0414	0.1898	YES
+	224	O=C(Nc(c(ccc1)C)c1)CC(=O)C	4.29886	2.4100	3.5461	0.0740	YES
-	225	N(CCNC1)C1	4.37954	2.8200	3.5683	0.0319	YES
-	226	Oc(cccc1N)c1	8.75654	2.8300	4.7690	1.0156	YES
-	227	O=S(=O)(N)c(c(ccc1)C)c1	3.81370	3.0000	3.4130	1.0617	YES
-	228	N#CCCCCCCC#N	6.23069	3.0500	4.0761	0.0176	YES
*	229	n1nc(N)nc1	8.27026	3.1100	4.6356	0.0426	YES
-	230	OC(CC(NC1(C)C)(C)C)C1	5.03564	3.1200	3.7482	0.0393	YES
-	231	O=S(=O)(O)c(c(ccc1[N+](=O)[O-])C)c1	6.12822	3.2600	4.0480	2.1742	No
-	232	n(ccc1)c(CCC2)c1C2	6.72914	3.3200	4.2128	0.0313	YES
+	233	OCC(N)(CC)CO	4.49863	3.3700	3.6009	0.0272	YES
-	234	C1(C#N)=CCCCC1	8.37621	3.4200	4.6647	1.0273	YES
-	235	[O-][N+](=O)c(c(OC)ccc1)c1	3.25648	3.4300	3.2602	0.1222	YES
*	236	NC(CCCC1)C1	3.46034	3.4600	3.3161	0.0151	YES
+	237	n(ccc(N)c1)c1	6.18360	3.5000	4.0632	0.0248	YES
*	238	Nc(cccc1N)c1	4.42950	3.5600	3.5820	1.0171	YES
+	239	Cc1ncc([N+](=O)[O-])n1CCO	7.79437	3.6300	4.5051	0.1523	YES
+	240	NCc(cccc1CN)c1	6.58475	3.6900	4.1732	0.0212	YES
-	241	Nc(c(ccc1N)C)c1	5.08516	3.8300	3.7618	1.0194	YES

#	242	O(c(ccc(N)c1)c1)c(ccc(N)c2)c2	13.21178	3.8500	5.9913	0.0326	YES
#	243	N#Cc(c(ccc1)Cl)c1	4.67885	3.8600	3.6504	0.0157	YES
-	244	n(c(N)ccc1)c1	6.18360	3.8900	4.0632	0.0248	YES
*	245	[O-][N+](=O)c(cccc1OC)c1	5.23417	3.9300	3.8027	0.1244	YES
+	246	N(O)=C(CCCC1)C1	4.89851	4.0600	3.7106	0.0462	YES
#	247	[O-][N+](=O)c(c(c([N+](=O)[O-])cc1)C)c1	8.11339	4.0800	4.5926	0.2350	YES
+	248	[O-][N+](=O)c(c(N)ccc1OC)c1	7.32068	4.1500	4.3751	0.1329	YES
-	249	Nc1c(Cl)ccc(C(=O)O)c1	10.70336	4.1900	5.3031	1.0493	YES
-	250	N(N)(C)C	4.88344	4.2500	3.7065	0.0242	YES
+	251	N=C(Nc(cccc1)c1)Nc(cccc2)c2	7.17222	4.4400	4.3344	2.0513	No
-	252	N=C(Nc(c(ccc1)C)c1)Nc(c(ccc2)C)c2	8.48353	4.4400	4.6941	2.0560	No
-	253	[O-][N+](=O)c(c(O)ccc1N)c1	5.21552	4.4900	3.7976	1.1299	No
-	254	O=S(=O)(Nc(nccc1)n1)c(ccc(N)c2)c2	5.74685	4.5100	3.9434	1.1032	No
*	255	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C)c1	9.25630	4.5500	4.9061	0.2436	YES
#	256	NC(C(CC(C1)CC(CCC(N)C2C)C2)C)C1	4.67511	4.5900	3.6493	0.0507	YES
+	257	COCC(=O)N(C(C)C(=O)OC)c1c(C)cccc1C	5.47280	4.6500	3.8682	0.0862	YES
+	258	N(c(cccc1)c1)(CC)CC	9.44591	4.7300	4.9582	0.0222	YES
+	259	NCc1cc(cc(c1)C(F)(F)F)C(F)(F)F	9.15656	4.8300	4.8788	1.0927	No
-	260	CC(C)(NC(=O)c1cc(Cl)cc(Cl)c1)C#C	10.64307	4.8800	5.2866	0.0665	YES
+	261	O=C(N)C(=C(O)C(N(C)C)C(C1(O)C(O)=C(C2C(O)(c(c3c(O)cc4)c4)C)C3(=O))C2O)C1(=O)	9.33263	5.0100	4.9271	8.2390	No
+	262	Nc(c(N)ccc1)c1	4.85412	5.1200	3.6984	0.0209	YES
+	263	N(N)C	4.02194	5.2200	3.4701	0.0225	YES
+	264	Nc(c(cc(c(ccc(N)c1Cl)c1)c2)Cl)c2	13.52827	5.2600	6.0781	0.0398	YES

+	265	N(c(c(Sc1cccc2)ccc3)c3)c12	13.06826	5.4300	5.9519	0.0607	YES
-	266	n(c(c(ccc1)cc2)c1O)c2	7.11884	5.4500	4.3197	0.0344	YES
-	267	O(c(ccc(NC(C=C1C)(C)C)c12)c2)CC	13.88934	5.4900	6.1772	0.0608	YES
+	268	c(c(c(N)cc1)c(N)cc2)(c1)c2	13.18125	5.5200	5.9829	0.0298	YES
#	269	N(c(c(S1)ccc2)c2)=C1S	13.05114	5.5200	5.9472	0.0565	YES
-	270	c1c(Cl)cc(Cl)cc1N2C(=O)C(C)(C3)C3(C)C2(=O)	12.59072	5.7000	5.8209	4.1101	No
+	271	N(CCNCCNCCNCCN)CCN	12.34203	5.7400	5.7527	0.0803	YES
+	272	Oc(c(N)ccc1)c1	9.18116	5.8600	4.8855	0.0195	YES
#	273	n(c(nc(n1)NCC)NCC)c1Cl	14.24032	5.9600	6.2734	0.0776	YES
+	274	Oc(ccc(N)c1)c1	9.18116	6.0400	4.8855	0.0195	YES
-	275	N#CCCCCCCCCCCC	11.23598	6.0800	5.4492	0.0172	YES
-	276	c12c(nc3c(o2)cccc3)cc(c(c1)=O)N	9.56407	6.1400	4.9906	0.0835	YES
+	277	CCCN(CCC)c1c(cc(cc1[N+](=O)[O-])C(F)(F)F)[N+](=O)[O-]	14.18119	6.1800	6.2572	2.2890	No
+	278	N(CCNCCNCCN)CCN	10.05811	6.2000	5.1261	0.0641	YES
+	279	n(c(nc(n1)NC(C)C)NCC)c1Cl	12.95234	6.2200	5.9201	0.0806	YES
+	280	N(c(c(S1)ccc2)c2)=C1SNC(CCCC3)C3	14.31727	6.2500	6.2946	3.0893	No
*	281	O=N(CCCCCCCCCCCC)(C)C	16.89296	6.3200	7.0012	1.0454	YES
-	282	c1c(C)cc2nc3SC(=O)Sc3nc2c1	14.45294	6.3800	6.3318	1.0852	YES
-	283	N(c(c(S1)ccc2)c2)=C1SNC(C)(C)C	14.73809	6.4000	6.4100	1.0779	YES
-	284	C1(C(C)(C(C)CC3)C(OC(CSCCN(CC)CC)=O)CC(C=C)(C)C(O)C2C)C23CCC1=O	15.88082	6.4800	6.7235	5.1614	No
-	285	O(c(nc(nc1NC(C)(C)C)NCC)n1)C	13.64888	6.5800	6.1112	1.0833	YES
#	286	CSc1nc(NC(C)(C)C)nc(N)n1	16.58861	6.6700	6.9177	0.0807	YES
*	287	n(c(nc(n1)NC(C)(C)C)NCC)c1Cl	15.59758	6.8000	6.6458	0.0861	YES
*	288	O=C(Nc(ccc(c1)Cl)c1)Nc(ccc(c2Cl)Cl)c2	16.96402	6.8200	7.0207	1.0760	YES

+	289	n(c(c(c1cccc2)ccc3)c3)(c12)C=C	14.09938	6.9600	6.2348	0.0695	YES
*	290	[O-][N+](=O)c(c(c(c(c(c1)ccc2)c2cc3)c3c4)c1)c4	17.78092	7.7500	7.2448	5.1521	No
+	291	CCNc1nc(NC(C)(C)C)nc(SC)n1	20.36997	7.8600	7.9550	1.0913	No
-	292	Clc1nc(N)nc(NCC)n1	15.23358	7.9400	6.5459	0.0666	YES
#	293	N(CCCCCCCC)(CCCCCCCC)CCCCCCCC	26.19525	8.2100	9.5531	0.0399	YES
#	294	N(CCCCCCCCCCCCCCCCCC)(C)C	21.86331	8.2200	8.3647	0.0357	YES
-	295	C(#N)c(c(C(#N))ccc1)c1	4.38416	2.9600	3.5695	0.0230	YES
-	296	O(CCNC1)C1	3.31641	3.1800	3.2766	0.0246	YES
+	297	n(c(c(ccc1)cc2)c1)c2	7.67544	3.2900	4.4724	0.0239	YES
*	298	S(=O)(=O)C(N)=N	5.30281	3.3400	3.8215	2.0720	No
-	299	COc2cc(Cc1cnc(N)nc1N)cc(OC)c2OC	5.41802	3.3500	3.8531	2.0598	No
+	300	n(c(nc(n1)c(cccc2)c2)N)c1N	10.04401	3.4200	5.1222	1.0564	YES
-	301	NC(=S)NN	6.73344	3.6800	4.2140	1.0384	YES
+	302	N(O)=C(CC)C	3.22801	3.7400	3.2523	0.0449	YES
+	303	c12n(=O)c(C(=O)NCCO)c(C)n(=O)c1cccc2	5.81360	3.8200	3.9617	1.1262	No
-	304	O=S(=O)(N)c(ccc(N)c1)c1	5.24455	3.8700	3.8056	1.0678	YES
+	305	O(c(ccc(N)c1)c1)C	7.44181	3.9800	4.4084	0.0196	YES
+	306	OCCN	1.03379	4.3900	2.6504	0.0118	YES
#	307	O(CCN(c(c(OCC)cc(N)c1OCC)c1)C2)C2	12.84302	4.5600	5.8901	0.0621	YES
-	308	Nc1cccc2c(N)cccc12	10.39051	4.8800	5.2173	0.0251	YES
+	309	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C(CC)C)c1	10.13429	5.2300	5.1470	0.2487	YES
+	310	N(c(ccc(NC(CC)C)c1)c1)C(CC)C	10.79608	5.3700	5.3286	0.0445	YES
-	311	Fc1ccc(cc1Cl)[N+](=O)[O-]	10.32613	5.4700	5.1996	1.1277	No
-	312	CN(C)c1ccc(cc1)C(=C2C=CC(C=C2)=N(Cl)(C)C)c3cccc	18.37424	5.5200	7.4075	1.1256	No

		c3					
-	313	n(cccc1N)c1	5.75898	5.5800	3.9467	1.0209	YES
+	314	c1ccc3c(c1)OC=2C(C=C(C(C=2)=O)NC(C)=O)=N3	12.84957	5.7700	5.8919	4.1514	No
-	315	O=C2C(N)=CC1=Nc3c(OC1=C2)cc(OC)cc3	11.07834	6.6600	5.4060	3.1248	No
+	316	N=C(NC1C(C(C(OC3C(C(C(O3)C)(C(=O))O)OC2C(C(O)C(C(O2)CO)O)NC)C(C1O)O)NC(=NOS(=O)(=O)O)N)O)N	16.04388	6.7100	6.7682	9.3150	No
+	317	O=C1N(N)C(SC)=NN=C1C(C)(C)C	15.09475	6.7400	6.5079	5.0930	No
-	318	CCNc1nc(NCC)nc(SC)n1	19.01270	6.8800	7.5827	1.0828	YES
-	319	CCC(CC)Nc1c(cc(C)c(C)c1[N+](=O)[O-])[N+](=O)[O-]	15.37698	7.3000	6.5853	1.2598	No
-	320	n(c(nc(n1)NC(C2)C2)NC(C)(C)C)c1SC	18.35265	8.0400	7.4016	1.1043	No
+	321	OCCO	-3.60289	0.2300	1.3784	0.0127	YES
-	322	c(c(c(c(ccc1)c2)c1cc3)c3cc4)(c2c(c5ccc6)c6)c45	25.41119	9.1400	9.3381	5.0479	No
*	323	O=S(=O)(c(ccc(O)c1)c1)c(ccc(O)c2)c2	6.75851	3.5900	4.2209	0.0771	YES
+	324	O=P(OCCOCCCC)(OCCOCCCC)OCCOCCCC	4.35264	3.8000	3.5609	5.0570	No
+	325	SCC	9.80838	4.3200	5.0576	0.0131	YES
#	326	FC(F)(F)c(cccc1)c1	4.76784	4.4300	3.6748	0.0515	YES
+	327	FC(F)(F)c(cccc1C(F)(F)F)c1	9.02818	4.4400	4.8436	0.0930	YES
-	328	CC(C)OP(=O)(OC(C)C)SCc1ccccc1	7.96151	4.4900	4.5509	4.0593	No
*	329	c(ccc1C(C)C)cc1C(C)C	5.99152	4.6800	4.0105	0.0178	YES
+	330	Oc(c(cc(c1)C(c(cc(c(O)c2Br)Br)c2)(C)C)Br)c1Br	9.03867	4.7600	4.8464	3.0880	No
+	331	O(O)C(C)(C)C	0.92577	4.9100	2.6208	0.0284	YES
-	332	s(c(c(c1cccc2)ccc3)c3)c12	11.62705	5.1200	5.5565	0.0429	YES
*	333	Oc(ccc(c1)C(c(cccc2)c2)(C)C)c1	10.21046	5.1800	5.1679	0.0302	YES
#	334	c(c(cccc1)c1)(cccc2)c2	10.26699	5.3000	5.1834	0.0157	YES
-	33	c(c(c(c(c1)ccc2)c2)ccc3)(c1)	12.55	5.44	5.81	0.0417	YES

	5	c3	844	00	20		
#	336	CC(c1cccc2c1cccc2)C	11.54287	5.8300	5.5334	0.0167	YES
+	337	Sc(cccc1)c1	11.28980	5.8400	5.4640	0.0134	YES
+	338	CCc2ccc(cc2)c1ccccc1	11.80140	6.0800	5.6044	0.0158	YES
+	339	SCCCCCCCC	16.30628	7.0200	6.8402	0.0193	YES
+	340	OCCS	12.38427	5.6600	5.7643	1.0135	YES
*	341	OC(CCC(C(C(CCC(O)C1)C1)(C)C)C2)C2	0.46194	3.4700	2.4935	0.0519	YES
-	342	O=P(OCC(CC)CCCC)(CC(CC)CCCC)O	5.97114	3.5200	4.0049	5.0509	No
#	343	O=S(=O)(c(ccc(c1)C)c1)Cl	3.03304	3.5200	3.1989	0.0585	YES
#	344	c(c(c(c1ccc2)c2)ccc3)(c3)C1	11.27538	5.3400	5.4600	0.0455	YES
*	345	c(cccc1)(c1)C(CCCC2)C2	8.51034	5.3700	4.7015	0.0261	YES
+	346	SCCCCC	13.05733	5.8700	5.9489	0.0162	YES
-	347	Oc(ccc(O)c1)c1	4.70788	6.3200	3.6583	0.0231	YES
#	348	SCCCCCCCCCC	18.47225	7.3500	7.4344	0.0214	YES

Table S2
Split 2: Distribution into training (+), invisible training (-), calibration (#), and validation (*) sets; ID; SMILES; Optimal descriptor (DCW); Experimental (Expr) and Calculated (Calc) toxicity to algae (pEC50); Defect-SMILES; domain of applicability

Set	**ID**	**SMILES**	**DCW**	**Expr**	**Calc**	**DefectSMILES**	**Applicability**
+	1	O(CCO)C	-1.42369	0.5000	1.3980	0.0090	YES
-	2	O=C(CCCCCCCCCCC)C	10.30897	6.2200	5.6362	0.0189	YES
+	3	O(CCOCC)CCO	-1.08351	0.9600	1.5209	0.0132	YES
+	4	O=C(C)C	0.30398	0.9600	2.0221	0.0090	YES
+	5	O(CCO)CC	-0.42319	1.0000	1.7594	0.0100	YES
-	6	O(CC)CC	0.46637	1.5100	2.0808	0.0083	YES

#	7	OCC(C)C	-0.34251	1.6400	1.7886	0.0060	YES
#	8	OC(C)(C)C	-0.08118	1.6600	1.8830	0.0101	YES
#	9	O(CCCC)CCO	1.23037	1.8100	2.3567	0.0114	YES
*	10	O(CCOCCO)CCCC	1.26493	2.1700	2.3692	0.0157	YES
-	11	O1C(C)(C)CCC1(C)C	2.16854	2.3400	2.6956	0.0262	YES
#	12	ClCCCl	3.09606	2.6300	3.0307	0.0202	YES
*	13	O(c(c(O)ccc1)c1)C	1.52569	2.6600	2.4634	0.0133	YES
+	14	C(Cl)(Cl)Cl	3.99464	2.7100	3.3553	0.0261	YES
+	15	OCCCCCC	3.03497	2.9500	3.0086	0.0089	YES
+	16	C(C=C)(=C)C	5.22181	3.0100	3.7986	0.0181	YES
*	17	ClCC(Cl)Cl	3.60497	3.1100	3.2145	0.0257	YES
+	18	N(CC)CC	4.31661	3.1300	3.4716	0.0150	YES
#	19	ClCC(Cl)CCl	4.32503	3.2700	3.4746	0.0282	YES
*	20	ClCCCC	4.26925	3.4100	3.4545	0.0146	YES
*	21	c(cccc1)c1	4.83454	3.4300	3.6587	0.0050	YES
+	22	C(C(Cl)Cl)(Cl)Cl	5.22360	3.4900	3.7992	0.0341	YES
#	23	O=C(O)CCCCCCC	5.48145	3.5700	3.8923	0.0169	YES
+	24	c(cccc1)(c1)Cl	5.19599	3.5800	3.7892	0.0156	YES
-	25	O=C(O)C(=O)O	-1.95470	3.6100	1.2062	0.0187	YES
*	26	C(CCCC1)C1	3.79241	3.6400	3.2822	0.0099	YES
+	27	BrCCCBr	4.89782	3.6400	3.6815	7.0022	No
#	28	O(CCCC)CCCC	4.46837	3.7700	3.5264	0.0123	YES
-	29	O(c(c(OC)cc(c1)CC=C)c1)C	5.31641	3.9100	3.8327	0.0214	YES
#	30	c(c(ccc1)Cl)(c1)Cl	6.276	4.010	4.17	0.0232	YES

			76	0	96		
#	31	c(ccc(c1)C)(c1)C	6.625 01	4.040 0	4.30 54	0.0089	YES
#	32	N(CC)(CC)CC	5.986 02	4.100 0	4.07 46	0.0145	YES
+	33	C(=CCCC=CC1)C1	6.276 28	4.120 0	4.17 95	0.0240	YES
-	34	BrC(Cl)Cl	4.643 16	4.140 0	3.58 95	4.0136	No
*	35	c(c(ccc1)Cl)(c1)C	6.169 01	4.210 0	4.14 07	0.0181	YES
-	36	C(C(C=C1)CC=C2)(C2)C 1	4.774 07	4.230 0	3.63 68	0.0439	YES
+	37	COc1c(Cl)cc(OC)cc1	5.395 69	4.240 0	3.86 14	0.0233	YES
*	38	O(c(ccc(c1)C(c(ccc(OCC O)c2)c2)(C)C)c1)CCO	6.751 83	4.270 0	4.35 12	0.0480	YES
*	39	c(ccc(c1)Cl)(c1)C	6.169 01	4.320 0	4.14 07	0.0181	YES
+	40	CCCCCCCCCc1ccc(cc1) OCCOCCOCCOCCOCC OCCOCCO	8.635 51	4.330 0	5.03 17	0.0487	YES
+	41	BrC(Br)Cl	4.708 78	4.340 0	3.61 32	8.0061	No
-	42	c(cccc1)(c1)CC	7.149 18	4.360 0	4.49 48	0.0092	YES
-	43	c(ccc(c1)C)(c1)C(C)C	7.249 53	4.360 0	4.53 10	0.0101	YES
-	44	OCCCCCCCC(C)C	5.660 49	4.370 0	3.95 70	0.0120	YES
-	45	C(C(Cl)Cl)(Cl)(Cl)Cl	7.089 87	4.380 0	4.47 34	0.0461	YES
-	46	C(C(C=CC12)C1)(=CC)C 2	6.582 28	4.390 0	4.29 00	0.0450	YES
*	47	c(ccc(c1Cl)Cl)(c1)Cl	7.502 66	4.500 0	4.62 25	0.0321	YES
-	48	COc1ccc(O)c(c1)C(C)(C) C	6.290 45	4.540 0	4.18 46	0.0204	YES
#	49	c(cccc1Cl)(c1)C	6.314 14	4.550 0	4.19 31	0.0194	YES
#	50	c(cccc1)(c1)C(C)C	6.773 20	4.660 0	4.35 90	0.0094	YES
-	51	O(Cc(cccc1)c1)Cc(cccc2) c2	4.617 08	4.680 0	3.58 01	0.0400	YES
+	52	c(c(c(cc1)C)ccc2)(c2)c1	9.155 68	4.710 0	5.21 96	0.0294	YES
*	53	c(c(c(cc1)Cl)C)(c1)Cl	6.753 08	4.780 0	4.35 17	0.0239	YES
-	54	c(ccc(c1Cl)C)(c1)Cl	6.898 21	4.800 0	4.40 41	0.0252	YES

+	55	OCCCCCCCCC	6.03646	4.8200	4.0928	0.0118	YES
#	56	C=Cc1ccccc1C=C	8.14119	4.8600	4.8531	0.0247	YES
+	57	c(cccc1)(c1)CCCC	9.15018	4.9200	5.2176	0.0112	YES
+	58	c(ccc(c1C)Cl)(c1)Cl	6.85755	4.9800	4.3894	0.0247	YES
-	59	c(c(c(cc1)Cl)Cl)(c1)Cl	7.35752	5.0500	4.5700	0.0308	YES
-	60	c(ccc(c1Cl)Cl)(c1)C	7.39490	5.0600	4.5835	0.0270	YES
-	61	c(c(ccc1)C)(c1)C	6.62501	5.1200	4.3054	0.0089	YES
#	62	c(cc(ccc1C)c2c1)c(c2)C	8.43402	5.1900	4.9589	0.0287	YES
+	63	CCC1CCCCC1	6.64606	5.2500	4.3130	0.0136	YES
-	64	Cc2cc(C)c1ccccc1c2	9.37323	5.4000	5.2982	0.0284	YES
+	65	C(CCCC1)(C1)C	5.10655	5.4600	3.7569	0.0131	YES
-	66	O(c(cccc1)c1)c(cccc2)c2	4.77800	5.4700	3.6382	0.0286	YES
-	67	C(Cl)(Cl)(Cl)Cl	5.86091	5.5200	4.0294	0.0381	YES
-	68	O(C(C)C)CCO	-1.14661	1.2100	1.4981	0.0096	YES
*	69	OCCCC	1.03397	1.6800	2.2858	0.0069	YES
+	70	O=C(CCCC1)C1	1.86737	1.9200	2.5868	0.0163	YES
+	71	ClCCl	2.09557	2.0400	2.6693	0.0192	YES
+	72	OC(CC)CC	0.65799	2.1300	2.1500	0.0070	YES
#	73	OCCCCC	2.03447	2.3800	2.6472	0.0079	YES
+	74	ClCC(Cl)C	3.49722	2.9100	3.1756	0.0207	YES
#	75	C(Cl)(Cl)C	3.88688	2.9700	3.3163	0.0210	YES
*	76	C(=CCl)(Cl)Cl	5.63156	3.2300	3.9466	1.0302	YES
#	77	O=C(O)CCCCCC	4.48095	3.3400	3.5309	0.0159	YES
+	78	N(CCCCCCN(C)C)(C)C	9.37512	3.3900	5.2989	0.0258	YES
+	79	c(cccc1)(c1)C	6.148	3.500	4.13	0.0082	YES

			68	0	34		
-	80	OCCCCCCC	4.035 47	3.530 0	3.37 00	0.0098	YES
-	81	BrCC(Br)CCl	6.429 34	3.580 0	4.23 48	8.0095	No
+	82	OCCCCCCCC	5.035 97	3.670 0	3.73 14	0.0108	YES
-	83	C(=C(Cl)Cl)(Cl)Cl	6.140 47	3.790 0	4.13 04	1.0358	YES
+	84	c(cccc1C)(c1)C	6.729 48	4.080 0	4.34 32	0.0097	YES
+	85	c(cccc1)(c1)Br	6.389 14	4.120 0	4.22 02	4.0067	No
-	86	Fc1c(F)cc(Br)cc1	3.295 78	4.140 0	3.10 28	4.0423	No
+	87	CCCCCCCCCC(=O)O	8.155 14	4.160 0	4.85 82	0.0212	YES
-	88	c(C(=C)C)(cccc1)c1	7.028 47	4.390 0	4.45 12	0.0182	YES
*	89	c(ccc(c1)Cl)(c1)Cl	6.276 76	4.430 0	4.17 96	0.0232	YES
+	90	O=C(CCCCCCCC)C	7.307 47	4.500 0	4.55 20	0.0160	YES
*	91	c(cccc1Cl)(c1)Cl	6.421 89	4.580 0	4.23 21	0.0245	YES
*	92	c(c(ccc1C)ccc2)(c2)c1	9.260 15	4.870 0	5.25 73	0.0302	YES
#	93	O=C(CCCCCCCCC)C	8.307 97	4.950 0	4.91 34	0.0170	YES
+	94	c(c(ccc1)ccc2)(c1CC3)c2 3	7.475 31	5.040 0	4.61 26	0.0437	YES
-	95	O=C(c(cccc1)c1)c(c(O)cc (OC)c2)c2	5.494 38	5.530 0	3.89 70	0.0376	YES
+	96	Oc(cccc1)c1	3.628 86	2.680 0	3.22 31	0.0089	YES
+	97	N(c(c(c(ccc1)cc2)c1)c2)c(cccc3)c3	11.50 564	6.810 0	6.06 85	0.0579	YES
*	98	Nc(c(ccc1)C)c1	3.401 92	2.950 0	3.14 12	0.0097	YES
+	99	Nc(c(ccc1C)C)c1	3.982 71	3.230 0	3.35 10	0.0112	YES
#	100	Oc(ccc(c1)C)c1	4.105 18	3.270 0	3.39 52	0.0096	YES
#	101	Oc(c(ccc1)Cl)c1	5.325 30	3.390 0	3.83 59	0.0184	YES
-	102	Nc(ccc(c1)C)c1	3.401 92	3.400 0	3.14 12	0.0097	YES
#	104	Nc(c(cc(c1)C)C)c1	3.878 24	3.490 0	3.31 32	0.0104	YES
*	105	[O-	5.036	3.500	3.73	0.0303	YES

		][N+](=O)c(c(N)ccc1)c1	16	0	15		
#	106	[O-][N+](=O)c(cccc1N)c1	4.90483	3.5100	3.6841	0.0302	YES
*	107	[O-][N+](=O)c(ccc(N)c1)c1	5.03616	3.5100	3.7315	0.0303	YES
+	108	N(c(cccc1)c1)CC	6.72563	3.5600	4.3418	0.0184	YES
-	109	Oc(ccc(c1C)C)c1	4.68597	3.5800	3.6050	0.0111	YES
-	110	Oc(c(ccc1)CC)c1	5.10568	3.5900	3.7566	0.0106	YES
-	111	Nc(cccc1C)c1	3.50639	3.6000	3.1789	0.0105	YES
#	112	Oc(ccc(c1)Cl)c1	5.32530	3.6100	3.8359	0.0184	YES
+	113	Nc(cc(cc1C)C)c1	3.98271	3.6200	3.3510	0.0112	YES
*	114	Nc(cccc1Cl)c1	5.90152	3.6800	4.0441	0.0203	YES
-	115	NCCCCCCN	4.74612	3.7900	3.6267	0.0158	YES
*	116	Oc(c(cc(c1)C)C)c1	4.58150	3.8000	3.5673	0.0103	YES
-	117	Nc(c(ccc1)Cl)c1Cl	6.29031	3.8400	4.1845	0.0269	YES
+	118	CC(C)c1ccc(N)cc1	8.28357	3.8800	4.9046	0.0175	YES
-	119	CCc1cccc(N)c1	7.57859	3.9300	4.6499	0.0221	YES
+	120	Nc(c(ccc1Cl)Cl)c1	6.98228	3.9900	4.4345	0.0279	YES
-	121	Oc(c(c(cc1C)C)C)c1	5.16229	4.0000	3.7771	0.0119	YES
-	122	Oc(c(ccc1C)C(C)C)c1	5.31049	4.0300	3.8306	0.0123	YES
#	123	Oc(cccc1Cl)c1	5.47044	4.0500	3.8884	0.0197	YES
#	124	Oc(ccc(c1)Cc(ccc(O)c2)c2)c1	7.08836	4.1000	4.4728	0.0398	YES
-	125	[O-][N+](=O)c(ccc(c1)C)c1	4.32818	4.1400	3.4758	0.0208	YES
+	126	Nc(ccc(c1)CC)c1	4.40242	4.1400	3.5026	0.0106	YES
-	127	Nc(ccc(c1C)C)c1	3.98271	4.1500	3.3510	0.0112	YES
+	128	Oc(c(cc(c1)C)C)c1C	4.96701	4.1500	3.7065	0.0126	YES
-	129	Oc(c(c(cc1)Cl)Cl)c1	6.40607	4.1700	4.2263	0.0260	YES
-	130	Nc(ccc(c1)Cc(ccc(N)c2)c	7.918	4.220	4.77	0.0475	YES

		2)c1	62	0	27		
+	131	Oc(ccc(c1)Br)c1	7.602 29	4.270 0	4.65 85	5.0070	No
+	132	Oc(c(cc(c1)Cl)Cl)c1	6.406 07	4.300 0	4.22 63	0.0260	YES
*	133	Oc(ccc(c1)C(CC)C)c1	5.730 20	4.300 0	3.98 22	0.0118	YES
+	134	[O-][N+](=O)c(c(N)ccc1Cl)c1	9.721 42	4.310 0	5.42 39	0.0382	YES
-	135	Oc(c(ccc1)C(CC)C)c1	5.730 20	4.340 0	3.98 22	0.0118	YES
*	136	n(ccc(c1)C=C)c1	6.015 71	4.360 0	4.08 53	0.0199	YES
-	137	Nc(c(c(cc1)Cl)Cl)c1	6.837 15	4.380 0	4.38 21	0.0266	YES
#	138	Nc(ccc(c1Cl)Cl)c1	6.982 28	4.400 0	4.43 45	0.0279	YES
+	139	Oc(c(ccc1)C(C)(C)C)c1C	6.377 04	4.420 0	4.21 59	0.0182	YES
*	140	Oc(c(cc(c1Cl)Cl)Cl)c1	7.631 97	4.420 0	4.66 92	0.0348	YES
-	141	[O-][N+](=O)c(c(O)ccc1Cl)c1	8.187 91	4.450 0	4.87 00	0.0306	YES
*	142	Oc(ccc(c(ccc(O)c1)c1)c2)c2	6.463 80	4.510 0	4.24 72	0.0309	YES
+	143	Nc(c(cc(c(ccc(N)c1C)c1)c2)C)c2	8.351 17	4.530 0	4.92 90	0.0408	YES
*	144	Nc(c(cc(c1)Cl)Cl)c1	6.837 15	4.610 0	4.38 21	0.0266	YES
-	145	Oc(c(ccc1Cl)Cl)c1	6.551 20	4.650 0	4.27 88	0.0273	YES
+	146	Oc1ccc(Cl)c(Cl)c1Cl	7.859 78	4.680 0	4.75 15	0.0334	YES
*	147	Nc1ccc(Cl)c(Cl)c1Cl	8.290 86	4.740 0	4.90 72	0.0340	YES
#	148	Nc(c(cc(c1Cl)Cl)Cl)c1	8.063 05	4.800 0	4.82 49	0.0354	YES
+	149	Oc(ccc(c(cccc1)c1)c2)c2	6.813 01	4.850 0	4.37 33	0.0283	YES
-	150	Oc(ccc(c1)CCCCC)c1	8.107 17	4.870 0	4.84 08	0.0136	YES
+	151	[O-][N+](=O)c(ccc(c1Cl)Cl)c1	9.617 88	4.890 0	5.38 65	0.0355	YES
-	152	Oc1ccc(cc1)C(c2ccc(cc2)O)(c3ccc(cc3)O)C	11.66 387	4.900 0	6.12 56	0.0641	YES
-	153	Oc(c(ccc1C)C(C)(C)C)c1	6.572 32	4.940 0	4.28 64	0.0174	YES

+	154	Oc1cc(Cl)cc(Cl)c1Cl	7.85978	4.9400	4.7515	0.0334	YES
-	155	Oc(c(cc(c1)C)C(C)(C)C)c1	6.46785	4.9600	4.2487	0.0166	YES
+	156	Oc(c(cc(c1)C(C)(C)C)C(C)(C)C)c1	8.35420	5.3100	4.9301	0.0229	YES
#	157	Oc(cccc1C)c1	4.20965	2.8700	3.4329	0.0104	YES
+	158	Nc(cccc1)c1	2.92560	2.9300	2.9691	0.0089	YES
*	159	Oc(c(ccc1)C)c1	4.10518	2.9300	3.3952	0.0096	YES
#	160	Nc(c(ccc1)CC)c1	4.40242	3.3900	3.5026	0.0106	YES
-	161	Oc(c(ccc1)C)c1C	4.49069	3.4100	3.5345	0.0119	YES
#	162	Oc(c(c(cc1)C)C)c1	4.58150	3.4800	3.5673	0.0103	YES
*	163	Oc(c(ccc1C)C)c1	4.68597	3.5800	3.6050	0.0111	YES
*	164	Nc(c(c(cc1)C)C)c1	3.87824	3.5900	3.3132	0.0104	YES
#	165	Oc(cc(cc1C)C)c1	4.68597	3.6500	3.6050	0.0111	YES
+	166	Nc(c(cc(c1)C)C)c1C	4.26376	3.7000	3.4525	0.0127	YES
-	167	Oc(ccc(c1)CC)c1	5.10568	3.7500	3.7566	0.0106	YES
-	168	Oc(c(ccc1C)C)c1C	5.07148	3.9800	3.7443	0.0134	YES
#	169	Oc(c(ccc1)Cl)c1Cl	5.85923	4.0100	4.0288	0.0263	YES
+	170	Oc(c(ccc1Cl)Cl)c1Cl	7.08513	4.3900	4.4716	0.0351	YES
-	171	Oc(ccc(C(=C)C)c1)c1	7.13961	4.4000	4.4913	0.0174	YES
+	172	Nc(cc(cc1Cl)Cl)c1	6.98228	4.5700	4.4345	0.0279	YES
+	173	Oc(ccc(c1ccc2)c2)c1	7.36036	4.8400	4.5711	0.0304	YES
+	174	Oc1ccc(Cl)c(Cl)c1	7.32585	4.8700	4.5586	0.0255	YES
-	175	Oc1cc(Cl)cc(Cl)c1	7.32585	4.8900	4.5586	0.0255	YES
-	176	Nc1cc(Cl)c(Cl)c(Cl)c1	8.83770	5.1400	5.1047	0.0337	YES
+	177	Oc(c(cc(c1)Br)Br)c1Br	9.07984	5.2400	5.1922	12.0070	No
+	178	CCC(C)c1cccc(C(C)CC)c1O	8.95358	5.2700	5.1466	0.0176	YES

-	179	Oc(c(cc(c1)Br)Br)c1	8.74867	5.3600	5.0726	9.0071	No
-	180	N(c(cccc1)c1)c(cccc2)c2	8.62824	5.6000	5.0291	0.0353	YES
#	181	Oc(ccc(c1)CCCCCCCC)c1	11.10867	6.1700	5.9251	0.0166	YES
*	182	N#CC(=C)C	3.81680	3.4300	3.2910	4.0108	No
*	183	CCc1cccc(CC)c1N(C(=O)CCl)CCOCCC	15.45158	7.9900	7.4938	0.0466	YES
-	184	C(#N)C=C	1.89347	3.7200	2.5963	4.0113	No
-	185	O=C1N(CCOC(=O)C=C)C(=O)N(C(=O)N1CCOC(=O)C=C)CCOC(=O)C=C	4.84737	4.2100	3.6633	0.0977	YES
+	186	C1Oc2cc3N(CC)C=C(C(=O)O)C(=O)c3cc2O1	6.11032	4.2100	4.1195	1.0822	YES
#	187	Fc1cc2C(=O)C(C(=O)O)=CN3c2c(CCC3C)c1	8.30519	4.7200	4.9124	0.0794	YES
+	188	N(=Nc(cccc1)c1)c(ccc(N)c2)c2	7.88912	4.8300	4.7621	2.0559	No
-	189	N#CC(=C)Cl	7.75641	6.1000	4.7141	3.0159	No
-	190	CCc1cccc(C)c1N(C(C)COC)C(=O)CCl	11.57741	6.6000	6.0944	0.0420	YES
+	191	[O-][N+](=O)c(ccc(c1)CCl)c1	9.11204	6.6500	5.2038	0.0291	YES
-	192	N(Nc(cccc1)c1)c(cccc2)c2	9.20650	5.2200	5.2379	0.0463	YES
-	193	N#CC=CC#N	8.13517	5.3100	4.8509	7.0121	No
+	194	C(C=CC#N)(F)(F)F	9.92636	5.5300	5.4980	4.0787	No
+	195	N(=C=S)C	10.66540	5.5900	5.7649	2.0304	No
#	196	CCCCOCN(C(=O)CCl)c1c(CC)cccc1CC	15.59574	7.9800	7.5459	0.0402	YES
-	197	OCc(cccc1)c1	1.55681	2.1500	2.4747	0.0132	YES
+	198	Brc2c(c(cc(c2)OCC1OC1)C)Br	10.32696	5.7200	5.6427	7.0383	No
#	199	OC(CCl)CCl	1.42499	2.3100	2.4270	0.0238	YES
+	200	O(CCCl)CCCl	3.23437	2.6200	3.0806	0.0270	YES
#	201	S(SC)C	4.38167	3.5100	3.4951	0.0197	YES
+	202	O(C1c(cccc2)c2)C1	1.83622	3.6800	2.5756	0.0328	YES

-	203	O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F	5.71909	4.1600	3.9782	0.4357	YES
-	204	C(=CCl)CCl	4.23776	4.7200	3.4431	1.0234	YES
+	205	C(Cl)(Cl)(Cl)C(Cl)COCC(Cl)C(Cl)(Cl)Cl	10.54641	5.5000	5.7220	0.0783	YES
+	206	C(C(Cl)CCl)=C	4.63891	3.3300	3.5880	1.0239	YES
+	207	c(cccc1)(c1)C(Cl)Cl	6.42496	3.7800	4.2332	0.0236	YES
-	208	S(SCC)CC	6.38266	4.4400	4.2179	0.0216	YES
-	209	c(c(ccc1)Cl)(c1)CCl	6.99682	5.1300	4.4397	0.0256	YES
-	210	O=C(Oc(c(ccc1)C(CC)C)c1)NC	8.18582	3.8000	4.8692	0.0305	YES
-	211	CC(C)Oc1cc(c(Cl)cc1Cl)N2N=C(OC2=O)C(C)(C)C	17.56557	6.7900	8.2575	4.0682	No
+	212	Cc1cccc(C)c1N(C(=O)COC)N2C(=O)OCC2	5.73474	4.5000	3.9838	1.0634	YES
#	213	Oc(c(cc(c1)Cl)Cl)c1Cl	6.93999	4.5400	4.4192	0.0338	YES
*	214	Nc(c(cc(c1)Cl)Cl)c1Cl	7.37107	4.6000	4.5749	0.0344	YES
+	215	Oc(c(cc(c1Cl)Cl)Cl)c1Cl	8.16589	5.0400	4.8620	0.0427	YES
-	216	N(N)c(c(cc(c1)Cl)Cl)c1Cl	11.07387	6.4000	5.9125	0.0517	YES
+	217	c1ccc(C)cc1NC(=O)Oc2ccccc(NC(=O)OC)c2	10.97448	6.4700	5.8766	0.0690	YES
-	218	CNC(=O)Oc1ccccc1OC(C)C	8.02787	4.3200	4.8122	0.0351	YES
#	219	Oc(c(c(c(c1Cl)Cl)Cl)Cl)c1Cl	9.24666	5.4900	5.2524	0.0503	YES
+	220	COP(=S)(OC)Oc1ccc([N+](=O)[O-])c(C)c1	10.51066	5.7500	5.7090	6.0441	No
-	221	CCN(CC)C(=O)SCc1ccc(Cl)cc1	12.65742	6.4500	6.4845	0.0446	YES
+	222	N1C(=O)NC(=O)NC1=O	4.02013	2.1300	3.3645	0.0622	YES
-	223	Clc1cc(C(F)(F)F)ccc1Oc2cc(OCC)c([N+](=O)[O-])cc2	26.84800	9.6600	11.6106	2.1198	No
+	224	O=C(Nc(c(ccc1)C)c1)CC(=O)C	5.53302	2.4100	3.9110	0.0303	YES
+	225	N(CCNC1)C1	3.30999	2.8200	3.1080	0.0288	YES
-	226	Oc(cccc1N)c1	7.826	2.830	4.73	1.0172	YES

			97	0	96		
+	227	O=S(=O)(N)c(c(ccc1)C)c1	3.44130	3.0000	3.1554	0.0435	YES
+	228	N#CCCCCCCC#N	6.88762	3.0500	4.4003	7.0135	No
+	229	n1nc(N)nc1	4.46750	3.1100	3.5261	0.0481	YES
+	230	OC(CC(NC1(C)C)(C)C)C1	5.12780	3.1200	3.7646	1.0310	YES
+	231	O=S(=O)(O)c(c(ccc1[N+](=O)[O-])C)c1	4.33484	3.2600	3.4782	1.0461	YES
-	232	n(ccc1)c(CCC2)c1C2	6.54224	3.3200	4.2755	0.0326	YES
*	233	OCC(N)(CC)CO	4.53596	3.3700	3.5508	1.0209	YES
+	234	C1(C#N)=CCCCC1	6.51908	3.4200	4.2672	4.0215	No
-	235	[O-][N+](=O)c(c(OC)ccc1)c1	3.07884	3.4300	3.0245	0.0238	YES
-	236	NC(CCCC1)C1	2.53456	3.4600	2.8278	0.0148	YES
#	237	n(ccc(N)c1)c1	5.78518	3.5000	4.0021	0.0213	YES
*	238	Nc(cccc1N)c1	3.97857	3.5600	3.3495	0.0191	YES
-	239	Cc1ncc([N+](=O)[O-])n1CCO	4.65564	3.6300	3.5940	0.0458	YES
*	240	NCc(cccc1CN)c1	4.98765	3.6900	3.7140	0.0251	YES
*	241	Nc(c(ccc1N)C)c1	4.45489	3.8300	3.5215	0.0198	YES
-	242	O(c(ccc(N)c1)c1)c(ccc(N)c2)c2	10.29176	3.8500	5.6300	1.0472	YES
*	243	N#Cc(c(ccc1)Cl)c1	9.53315	3.8600	5.3559	3.0206	No
#	244	n(c(N)ccc1)c1	5.78518	3.8900	4.0021	0.0213	YES
-	245	[O-][N+](=O)c(cccc1OC)c1	4.40343	3.9300	3.5029	0.0266	YES
-	246	N(O)=C(CCCC1)C1	4.43236	4.0600	3.5134	0.0286	YES
-	247	[O-][N+](=O)c(c(c([N+](=O)[O-])cc1)C)c1	6.62086	4.0800	4.3039	0.0318	YES
#	248	[O-][N+](=O)c(c(N)ccc1OC)c1	5.58773	4.1500	3.9307	0.0369	YES
#	249	Nc1c(Cl)ccc(C(=O)O)c1	7.43489	4.1900	4.5980	0.0244	YES
+	250	N(N)(C)C	3.692	4.250	3.24	0.0210	YES

			51	0	61		
*	251	N=C(Nc(cccc1)c1)Nc(ccc c2)c2	6.438 53	4.440 0	4.23 81	2.0515	No
+	252	N=C(Nc(c(ccc1)C)c1)Nc(c(ccc2)C)c2	7.391 18	4.440 0	4.58 22	2.0530	No
-	253	[O-][N+](=O)c(c(O)ccc1N)c1	4.555 62	4.490 0	3.55 79	0.0328	YES
-	254	O=S(=O)(Nc(nccc1)n1)c(ccc(N)c2)c2	5.486 71	4.510 0	3.89 43	0.0937	YES
#	255	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C)c1	6.376 13	4.550 0	4.21 55	0.0353	YES
-	256	NC(C(CC(C1)CC(CCC(N)C2C)C2)C)C1	5.699 40	4.590 0	3.97 11	0.0504	YES
-	257	COCC(=O)N(C(C)C(=O) OC)c1c(C)cccc1C	6.096 99	4.650 0	4.11 47	0.0401	YES
*	258	N(c(cccc1)c1)(CC)CC	8.395 03	4.730 0	4.94 48	0.0178	YES
#	259	NCc1cc(cc(c1)C(F)(F)F) C(F)(F)F	6.530 84	4.830 0	4.27 14	0.1516	YES
-	260	CC(C)(NC(=O)c1cc(Cl)cc (Cl)c1)C#C	8.517 19	4.880 0	4.98 89	3.0459	No
+	261	O=C(N)C(=C(O)C(N(C) C)C(C1(O)C(O)=C(C2C(O)(c(c3c(O)cc4)c4)C)C3(=O))C2O)C1(=O)	8.318 75	5.010 0	4.91 73	7.1203	No
-	262	Nc(c(N)ccc1)c1	4.109 91	5.120 0	3.39 69	0.0192	YES
+	263	N(N)C	3.023 60	5.220 0	3.00 45	0.0226	YES
-	264	Nc(c(cc(c(ccc(N)c1Cl)c1) c2)Cl)c2	11.35 074	5.260 0	6.01 25	0.0575	YES
-	265	N(c(c(Sc1cccc2)ccc3)c3)c 12	12.58 779	5.430 0	6.45 94	1.0663	YES
+	266	n(c(c(ccc1)cc2)c1O)c2	6.002 77	5.450 0	4.08 07	0.0361	YES
-	267	O(c(ccc(NC(C=C1C)(C)C)c12)c2)CC	10.72 958	5.490 0	5.78 81	0.0588	YES
+	268	c(c(c(N)cc1)c(N)cc2)(c1) c2	9.792 23	5.520 0	5.44 95	0.0461	YES
*	269	N(c(c(S1)ccc2)c2)=C1S	13.97 656	5.520 0	6.96 10	4.0535	No
+	270	c1c(Cl)cc(Cl)cc1N2C(=O)C(C)(C3)C3(C)C2(=O)	9.594 60	5.700 0	5.37 81	1.0810	YES
#	271	N(CCNCCNCCNCCN)CCN	11.54 693	5.740 0	6.08 34	0.0616	YES
+	272	Oc(c(N)ccc1)c1	7.958 31	5.860 0	4.78 71	1.0173	YES
+	273	n(c(nc(n1)NCC)NCC)c1Cl	11.80 598	5.960 0	6.17 69	0.0696	YES

-	274	Oc(ccc(N)c1)c1	7.95831	6.0400	4.7871	1.0173	YES
-	275	N#CCCCCCCCCCCC	7.27502	6.0800	4.5402	4.0147	No
+	276	c12c(nc3c(o2)cccc3)cc(c(c1)=O)N	11.52205	6.1400	6.0744	4.0715	No
-	277	CCCN(CCC)c1c(cc(cc1[N+](=O)[O-])C(F)(F)F)[N+](=O)[O-]	16.13341	6.1800	7.7401	1.1104	YES
-	278	N(CCNCCNCCN)CCN	9.11444	6.2000	5.2047	0.0502	YES
#	279	n(c(nc(n1)NC(C)C)NCC)c1Cl	11.43001	6.2200	6.0411	0.0698	YES
+	280	N(c(c(S1)ccc2)c2)=C1SNC(CCCC3)C3	11.43594	6.2500	6.0433	4.0784	No
-	281	O=N(CCCCCCCCCCCC)(C)C	13.86049	6.3200	6.9191	2.0244	No
+	282	c1c(C)cc2nc3SC(=O)Sc3nc2c1	12.14574	6.3800	6.2997	4.0803	No
+	283	N(c(c(S1)ccc2)c2)=C1SNC(C)(C)C	12.57881	6.4000	6.4561	4.0703	No
#	284	C1(C(C)(C(C)CC3)C(OC(CSCCN(CC)CC)=O)CC(C=C)(C)C(O)C2C)C23CCC1=O	13.36255	6.4800	6.7392	0.1077	YES
+	285	O(c(nc(nc1NC(C)(C)C)NCC)n1)C	10.06503	6.5800	5.5481	1.0729	YES
-	286	CSc1nc(NC(C)(C)C)nc(N)n1	12.99003	6.6700	6.6047	1.0742	YES
-	287	n(c(nc(n1)NC(C)(C)C)NCC)c1Cl	12.69184	6.8000	6.4969	0.0749	YES
+	288	O=C(Nc(ccc(c1)Cl)c1)Nc(ccc(c2Cl)Cl)c2	13.35930	6.8200	6.7381	1.0670	YES
#	289	n(c(c(c1cccc2)ccc3)c3)(c12)C=C	12.56187	6.9600	6.4500	0.0637	YES
+	290	[O-][N+](=O)c(c(c(c(c(c1)ccc2)c2cc3)c3c4)c1)c4	16.05768	7.7500	7.7128	5.0641	No
+	291	CCNc1nc(NC(C)(C)C)nc(SC)n1	16.28290	7.8600	7.7941	1.0781	YES
+	292	Clc1nc(N)nc(NCC)n1	16.59712	7.9400	7.9076	1.0695	YES
-	293	N(CCCCCCCC)(CCCCCCCC)CCCCCCCC	23.99500	8.2100	10.5800	0.0324	YES
-	294	N(CCCCCCCCCCCCCCCCCC)(C)C	19.99300	8.2200	9.1343	0.0284	YES
+	295	C(#N)c(c(C(#N))ccc1)c1	3.59792	2.9600	3.2120	7.0200	No
-	296	O(CCNC1)C1	2.60490	3.1800	2.8533	1.0202	YES

-	297	n(c(c(ccc1)cc2)c1)c2	7.25830	3.2900	4.5342	0.0296	YES
*	298	S(=O)(=O)C(N)=N	6.35254	3.3400	4.2070	1.0424	YES
-	299	COc2cc(Cc1cnc(N)nc1N)cc(OC)c2OC	2.88633	3.3500	2.9549	2.0849	No
-	300	n(c(nc(n1)c(cccc2)c2)N)c1N	8.15248	3.4200	4.8572	0.0718	YES
+	301	NC(=S)NN	4.59732	3.6800	3.5730	3.0344	No
*	302	N(O)=C(CC)C	3.86947	3.7400	3.3101	0.0223	YES
-	303	c12n(=O)c(C(=O)NCCO)c(C)n(=O)c1cccc2	3.72089	3.8200	3.2564	1.0609	YES
+	304	O=S(=O)(N)c(ccc(N)c1)c1	4.14929	3.8700	3.4111	0.0530	YES
*	305	O(c(ccc(N)c1)c1)C	6.20435	3.9800	4.1535	1.0191	YES
+	306	OCCN	1.57511	4.3900	2.4813	1.0074	YES
+	307	O(CCN(c(c(OCC)cc(N)c1OCC)c1)C2)C2	8.90031	4.5600	5.1273	1.0583	YES
*	308	Nc1cccc2c(N)cccc12	8.09321	4.8800	4.8358	0.0381	YES
-	309	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C(CC)C)c1	8.00115	5.2300	4.8025	0.0374	YES
+	310	N(c(ccc(NC(CC)C)c1)c1)C(CC)C	11.05559	5.3700	5.9059	0.0345	YES
+	311	Fc1ccc(cc1Cl)[N+](=O)[O-]	10.20733	5.4700	5.5995	1.0410	YES
+	312	CN(C)c1ccc(cc1)C(=C2C=CC(C=C2)=N(Cl)(C)C)c3ccccc3	9.92344	5.5200	5.4969	3.0889	No
-	313	n(cccc1N)c1	5.65385	5.5800	3.9546	0.0212	YES
+	314	c1ccc3c(c1)OC=2C(C=C(C(C=2)=O)NC(C)=O)=N3	9.96431	5.7700	5.5117	3.0840	No
-	315	O=C2C(N)=CC1=Nc3c(OC1=C2)cc(OC)cc3	12.19528	6.6600	6.3176	1.0955	YES
+	316	N=C(NC1C(C(C(OC3C(C(C(O3)C)(C(=O))O)OC2C(C(O)C(C(O2)CO)O)NC)C(C1O)O)NC(=NOS(=O)(=O)O)N)O)N	13.05616	6.7100	6.6285	6.1907	No
+	317	O=C1N(N)C(SC)=NN=C1C(C)(C)C	13.55180	6.7400	6.8076	3.0727	No
-	318	CCNc1nc(NCC)nc(SC)n1	15.39705	6.8800	7.4741	1.0728	YES

-	319	CCC(CC)Nc1c(cc(C)c(C)c1[N+](=O)[O-])[N+](=O)[O-]	14.30692	7.3000	7.0804	1.0441	YES
+	320	n(c(nc(n1)NC(C2)C2)NC(C)(C)C)c1SC	16.71722	8.0400	7.9510	2.0901	No
*	321	OCCO	-2.20403	0.2300	1.1161	0.0060	YES
*	322	c(c(c(c(ccc1)c2)c1cc3)c3cc4)(c2c(c5ccc6)c6)c45	20.89806	9.1400	9.4613	5.0518	No
#	323	O=S(=O)(c(ccc(O)c1)c1)c(ccc(O)c2)c2	4.64967	3.5900	3.5919	0.0591	YES
-	324	O=P(OCCOCCCC)(OCCOCCCC)OCCOCCCC	0.90038	3.8000	2.2375	3.0357	No
-	325	SCC	9.76688	4.3200	5.4404	1.0055	YES
+	326	FC(F)(F)c(cccc1)c1	6.37591	4.4300	4.2155	0.0752	YES
#	327	FC(F)(F)c(cccc1C(F)(F)F)c1	7.36866	4.4400	4.5741	0.1470	YES
-	328	CC(C)OP(=O)(OC(C)C)SCc1ccccc1	6.44713	4.4900	4.2412	4.0371	No
-	329	c(ccc1C(C)C)cc1C(C)C	7.04989	4.6800	4.4589	0.0112	YES
+	330	Oc(c(cc(c1)C(c(cc(c(O)c2Br)Br)c2)(C)C)Br)c1Br	7.96445	4.7600	4.7893	16.0344	No
-	331	O(O)C(C)(C)C	-0.47380	4.9100	1.7411	0.0145	YES
*	332	s(c(c(c1cccc2)ccc3)c3)c12	9.82218	5.1200	5.4603	0.0470	YES
+	333	Oc(ccc(c1)C(c(cccc2)c2)(C)C)c1	9.40438	5.1800	5.3094	0.0377	YES
#	334	c(c(cccc1)c1)(cccc2)c2	9.05179	5.3000	5.1821	0.0261	YES
+	335	c(c(c(c(c1)ccc2)c2)ccc3)(c1)c3	10.85174	5.4400	5.8323	0.0481	YES
*	336	CC(c1cccc2c1cccc2)C	11.11213	5.8300	5.9263	0.0281	YES
-	337	Sc(cccc1)c1	11.70651	5.8400	6.1410	2.0081	No
+	338	CCc2ccc(cc2)c1ccccc1	10.76009	6.0800	5.7991	0.0330	YES
-	339	SCCCCCCCC	15.76987	7.0200	7.6088	1.0115	YES
-	340	OCCS	8.52988	5.6600	4.9935	1.0067	YES
-	341	OC(CCC(C(C(CCC(O)C1)C1)(C)C)C2)C2	2.48780	3.4700	2.8110	0.0424	YES
-	342	O=P(OCC(CC)CCCC)(C	5.414	3.520	3.86	3.0231	No

		C(CC)CCCC)O	28	0	81		
-	343	O=S(=O)(c(ccc(c1)C)c1)Cl	4.00732	3.5200	3.3599	0.0362	YES
+	344	c(c(c(c1ccc2)c2)ccc3)(c3)C1	10.17977	5.3400	5.5895	0.0515	YES
+	345	c(cccc1)(c1)C(CCCC2)C2	8.30251	5.3700	4.9114	0.0290	YES
*	346	SCCCCC	12.76837	5.8700	6.5246	1.0085	YES
+	347	Oc(ccc(O)c1)c1	3.27965	6.3200	3.0970	0.0115	YES
-	348	SCCCCCCCCCC	17.77087	7.3500	8.3316	1.0135	YES

Table S3
Split 3: Distribution into training (+), invisible training (-), calibration (#), and validation (*) sets; ID; SMILES; Optimal descriptor (DCW); Experimental (Expr) and Calculated (Calc) toxicity to algae (pEC50); Defect-SMILES; domain of applicability

Set	**ID**	**SMILES**	**DCW**	**Expr**	**Calc**	**DefectSMILES**	**Applicability**
-	1	O(CCO)C	-0.54766	0.5000	1.9953	0.0140	YES
+	2	O=C(CCCCCCCCCCC)C	8.63096	6.2200	5.3825	0.0231	YES
-	3	O(CCOCC)CCO	-0.03856	0.9600	2.1832	0.0169	YES
-	4	O=C(C)C	0.59778	0.9600	2.4180	0.0130	YES
-	5	O(CCO)CC	0.25566	1.0000	2.2918	0.0151	YES
*	6	O(CC)CC	0.84676	1.5100	2.5099	0.0136	YES
+	7	OCC(C)C	-0.32083	1.6400	2.0790	0.0127	YES
*	8	OC(C)(C)C	0.06392	1.6600	2.2210	0.0176	YES
#	9	O(CCCC)CCO	1.69542	1.8100	2.8231	0.0159	YES
*	10	O(CCOCCO)CCCC	1.73495	2.1700	2.8377	0.0201	YES
-	11	O1C(C)(C)CCC1(C)C	3.04022	2.3400	3.3194	0.0345	YES
-	12	ClCCCl	2.70411	2.6300	3.1953	0.0173	YES
*	13	O(c(c(O)ccc1)c1)C	1.947	2.66	2.91	0.0239	YES

			10	00	60		
+	14	C(Cl)(Cl)Cl	4.098 85	2.71 00	3.71 00	0.0199	YES
*	15	OCCCCCC	2.347 75	2.95 00	3.06 38	0.0120	YES
+	16	C(C=C)(=C)C	4.739 91	3.01 00	3.94 66	0.0193	YES
#	17	ClCC(Cl)Cl	3.579 53	3.11 00	3.51 84	0.0199	YES
#	18	N(CC)CC	3.947 25	3.13 00	3.65 41	0.0114	YES
*	19	ClCC(Cl)CCl	3.820 65	3.27 00	3.60 74	0.0239	YES
-	20	ClCCCC	3.364 39	3.41 00	3.43 90	0.0137	YES
*	21	c(cccc1)c1	4.667 69	3.43 00	3.91 99	0.0115	YES
*	22	C(C(Cl)Cl)(Cl)Cl	5.215 39	3.49 00	4.12 21	0.0265	YES
+	23	O=C(O)CCCCCCC	5.031 86	3.57 00	4.05 43	0.0206	YES
-	24	c(cccc1)(c1)Cl	5.146 40	3.58 00	4.09 66	0.0194	YES
-	25	O=C(O)C(=O)O	- 0.477 01	3.61 00	2.02 14	0.0226	YES
#	26	C(CCCC1)C1	3.758 75	3.64 00	3.58 45	0.0176	YES
+	27	BrCCCBr	3.980 31	3.64 00	3.66 63	1.0234	YES
+	28	O(CCCC)CCCC	4.060 04	3.77 00	3.69 57	0.0177	YES
+	29	O(c(c(OC)cc(c1)CC=C)c1)C	5.010 33	3.91 00	4.04 64	0.0298	YES
+	30	c(c(ccc1)Cl)(c1)Cl	5.993 23	4.01 00	4.40 91	0.0243	YES
*	31	c(ccc(c1)C)(c1)C	5.965 06	4.04 00	4.39 87	0.0205	YES
+	32	N(CC)(CC)CC	4.898 78	4.10 00	4.00 52	0.0165	YES
*	33	C(=CCCC=CC1)C1	5.803 14	4.12 00	4.33 90	0.0236	YES
#	34	BrC(Cl)Cl	5.050 82	4.14 00	4.06 13	0.0212	YES
+	35	c(c(ccc1)Cl)(c1)C	5.287 99	4.21 00	4.14 88	0.0227	YES
+	36	C(C(C=C1)CC=C2)(C2)C1	5.022 99	4.23 00	4.05 11	0.0463	YES
#	37	COc1c(Cl)cc(OC)cc1	5.075 47	4.24 00	4.07 04	0.0271	YES

+	38	O(c(ccc(c1)C(c(ccc(OCCO)c2)c2)(C)C)c1)CCO	6.39818	4.2700	4.5585	0.0661	YES
*	39	c(ccc(c1)Cl)(c1)C	5.28799	4.3200	4.1488	0.0227	YES
#	40	CCCCCCCCCc1ccc(cc1)OCCOCCOCCOCCOCCOCCO	8.09688	4.3300	5.1854	0.0515	YES
+	41	BrC(Br)Cl	4.46659	4.3400	3.8457	2.0258	No
-	42	c(cccc1)(c1)CC	6.49339	4.3600	4.5937	0.0185	YES
-	43	c(ccc(c1)C)(c1)C(C)C	6.50976	4.3600	4.5997	0.0253	YES
-	44	OCCCCCCCC(C)C	4.49908	4.3700	3.8577	0.0187	YES
*	45	C(C(Cl)Cl)(Cl)(Cl)Cl	6.97531	4.3800	4.7715	0.0343	YES
#	46	C(C(C=CC12)C1)(=CC)C2	6.13992	4.3900	4.4632	0.0447	YES
+	47	c(ccc(c1Cl)Cl)(c1)Cl	6.61437	4.5000	4.6383	0.0284	YES
*	48	COc1ccc(O)c(c1)C(C)(C)C	5.94661	4.5400	4.3919	0.0378	YES
-	49	c(cccc1Cl)(c1)C	5.06229	4.5500	4.0656	0.0220	YES
*	50	c(cccc1)(c1)C(C)C	6.23477	4.6600	4.4982	0.0223	YES
-	51	O(Cc(cccc1)c1)Cc(cccc2)c2	4.92695	4.6800	4.0156	0.0509	YES
+	52	c(c(c(cc1)C)ccc2)(c2)c1	7.84851	4.7100	5.0938	0.0420	YES
+	53	c(c(c(cc1)Cl)C)(c1)Cl	6.26822	4.7800	4.5106	0.0273	YES
-	54	c(ccc(c1Cl)C)(c1)Cl	6.04253	4.8000	4.4273	0.0266	YES
-	55	OCCCCCCCCC	4.75771	4.8200	3.9532	0.0150	YES
#	56	C=Cc1ccccc1C=C	7.34938	4.8600	4.9096	0.0271	YES
#	57	c(cccc1)(c1)CCCC	8.10003	4.9200	5.1866	0.0205	YES
+	58	c(ccc(c1C)Cl)(c1)Cl	6.54196	4.9800	4.6116	0.0291	YES
*	59	c(c(c(cc1)Cl)Cl)(c1)Cl	6.84007	5.0500	4.7216	0.0291	YES
-	60	c(ccc(c1Cl)Cl)(c1)C	5.90913	5.0600	4.3781	0.0269	YES
+	61	c(c(ccc1)C)(c1)C	5.96506	5.1200	4.3987	0.0205	YES
-	62	c(cc(ccc1C)c2c1)c(c2)C	7.45191	5.1900	4.9474	0.0422	YES

+	63	CCC1CCCCC1	5.87946	5.2500	4.3671	0.0217	YES
#	64	Cc2cc(C)c1ccccc1c2	8.63395	5.4000	5.3836	0.0379	YES
+	65	C(CCCC1)(C1)C	4.78113	5.4600	3.9618	0.0236	YES
-	66	O(c(cccc1)c1)c(cccc2)c2	4.74927	5.4700	3.9500	0.0399	YES
-	67	C(Cl)(Cl)(Cl)Cl	5.85877	5.5200	4.3595	0.0277	YES
*	68	O(C(C)C)CCO	-0.16984	1.2100	2.1347	0.0177	YES
*	69	OCCCC	0.74111	1.6800	2.4709	0.0099	YES
-	70	O=C(CCCC1)C1	2.26156	1.9200	3.0320	0.0221	YES
-	71	ClCCl	1.90079	2.0400	2.8989	0.0163	YES
-	72	OC(CC)CC	0.48249	2.1300	2.3755	0.0137	YES
#	73	OCCCCC	1.54443	2.3800	2.7674	0.0109	YES
*	74	ClCC(Cl)C	2.87429	2.9100	3.2581	0.0183	YES
-	75	C(Cl)(Cl)C	3.39361	2.9700	3.4498	0.0184	YES
-	76	C(=CCl)(Cl)Cl	4.99606	3.2300	4.0411	1.0229	YES
#	77	O=C(O)CCCCCC	4.22854	3.3400	3.7579	0.0196	YES
-	78	N(CCCCCCN(C)C)(C)C	7.17793	3.3900	4.8463	0.0277	YES
+	79	c(cccc1)(c1)C	5.69007	3.5000	4.2972	0.0175	YES
-	80	OCCCCCCC	3.15107	3.5300	3.3603	0.0130	YES
+	81	BrCC(Br)CCl	5.51103	3.5800	4.2312	2.0308	No
+	82	OCCCCCCCC	3.95439	3.6700	3.6567	0.0140	YES
+	83	C(=C(Cl)Cl)(Cl)Cl	5.87149	3.7900	4.3642	1.0255	YES
*	84	c(cccc1C)(c1)C	6.23880	4.0800	4.4997	0.0223	YES
-	85	c(cccc1)(c1)Br	5.54353	4.1200	4.2431	2.0223	No
+	86	Fc1c(F)cc(Br)cc1	5.31427	4.1400	4.1585	4.0516	No
-	87	CCCCCCCCCC(=O)O	6.889	4.16	4.73	0.0231	YES

			54	00	99		
+	88	c(C(=C)C)(cccc1)c1	6.792 35	4.39 00	4.70 40	0.0251	YES
#	89	c(ccc(c1)Cl)(c1)Cl	5.993 23	4.43 00	4.40 91	0.0243	YES
*	90	O=C(CCCCCCCC)C	6.221 01	4.50 00	4.49 32	0.0200	YES
+	91	c(cccc1Cl)(c1)Cl	5.767 54	4.58 00	4.32 58	0.0236	YES
+	92	c(c(ccc1C)ccc2)(c2)c1	8.122 25	4.87 00	5.19 48	0.0438	YES
+	93	O=C(CCCCCCCCC)C	7.024 33	4.95 00	4.78 96	0.0210	YES
+	94	c(c(ccc1)ccc2)(c1CC3)c23	6.752 34	5.04 00	4.68 92	1.0504	YES
+	95	O=C(c(cccc1)c1)c(c(O)cc(OC)c 2)c2	5.545 68	5.53 00	4.24 39	0.0510	YES
-	96	Oc(cccc1)c1	3.760 21	2.68 00	3.58 50	0.0186	YES
-	97	N(c(c(c(ccc1)cc2)c1)c2)c(cccc3)c3	9.592 63	6.81 00	5.73 74	2.0519	No
-	98	Nc(c(ccc1)C)c1	2.806 72	2.95 00	3.23 32	0.0134	YES
-	99	Nc(c(ccc1C)C)c1	3.355 45	3.23 00	3.43 57	0.0183	YES
#	100	Oc(ccc(c1)C)c1	4.035 20	3.27 00	3.68 65	0.0216	YES
-	101	Oc(c(ccc1)Cl)c1	4.854 90	3.39 00	3.98 90	0.0212	YES
#	102	Nc(ccc(c1)C)c1	2.806 72	3.40 00	3.23 32	0.0134	YES
#	104	Nc(c(cc(c1)C)C)c1	3.081 71	3.49 00	3.33 47	0.0165	YES
#	105	[O-][N+](=O)c(c(N)ccc1)c1	4.488 13	3.50 00	3.85 37	0.0861	YES
+	106	[O-][N+](=O)c(cccc1N)c1	4.910 97	3.51 00	4.00 97	0.0892	YES
*	107	[O-][N+](=O)c(ccc(N)c1)c1	4.488 13	3.51 00	3.85 37	0.0861	YES
+	108	N(c(cccc1)c1)CC	5.961 75	3.56 00	4.39 75	0.0181	YES
-	109	Oc(ccc(c1C)C)c1	4.583 93	3.58 00	3.88 90	0.0264	YES
*	110	Oc(c(ccc1)CC)c1	4.838 52	3.59 00	3.98 30	0.0226	YES
+	111	Nc(cccc1C)c1	3.080 46	3.60 00	3.33 42	0.0152	YES
+	112	Oc(ccc(c1)Cl)c1	4.854 90	3.61 00	3.98 90	0.0212	YES
#	113	Nc(cc(cc1C)C)c1	3.355	3.62	3.43	0.0183	YES

			45	00	57		
+	114	Nc(cccc1Cl)c1	5.261 27	3.68 00	4.13 90	0.0163	YES
+	115	NCCCCCCN	4.339 46	3.79 00	3.79 88	0.0126	YES
#	116	Oc(c(cc(c1)C)C)c1	4.310 19	3.80 00	3.78 80	0.0246	YES
*	117	Nc(c(ccc1)Cl)c1Cl	5.860 01	3.84 00	4.35 99	0.0210	YES
-	118	CC(C)c1ccc(N)cc1	6.901 48	3.88 00	4.74 43	0.0187	YES
-	119	CCc1cccc(N)c1	6.445 62	3.93 00	4.57 61	0.0195	YES
-	120	Nc(c(ccc1Cl)Cl)c1	6.108 10	3.99 00	4.45 15	0.0211	YES
#	121	Oc(c(c(cc1C)C)C)c1	4.858 92	4.00 00	3.99 05	0.0295	YES
#	122	Oc(c(ccc1C)C(C)C)c1	5.128 62	4.03 00	4.09 00	0.0312	YES
*	123	Oc(cccc1Cl)c1	4.629 20	4.05 00	3.90 57	0.0206	YES
+	124	Oc(ccc(c1)Cc(ccc(O)c2)c2)c1	6.538 52	4.10 00	4.61 03	0.0559	YES
+	125	[O-][N+](=O)c(ccc(c1)C)c1	3.650 95	4.14 00	3.54 47	0.0854	YES
-	126	Nc(ccc(c1)CC)c1	3.610 04	4.14 00	3.52 96	0.0144	YES
#	127	Nc(ccc(c1C)C)c1	3.355 45	4.15 00	3.43 57	0.0183	YES
+	128	Oc(c(cc(c1)C)C)c1C	4.477 43	4.15 00	3.84 97	0.0296	YES
+	129	Oc(c(c(cc1)Cl)Cl)c1	5.701 74	4.17 00	4.30 15	0.0261	YES
+	130	Nc(ccc(c1)Cc(ccc(N)c2)c2)c1	6.533 05	4.22 00	4.60 83	0.0468	YES
+	131	Oc(ccc(c1)Br)c1	7.044 62	4.27 00	4.79 71	3.0239	No
*	132	Oc(c(cc(c1)Cl)Cl)c1	5.701 74	4.30 00	4.30 15	0.0261	YES
#	133	Oc(ccc(c1)C(CC)C)c1	5.383 22	4.30 00	4.18 40	0.0274	YES
+	134	[O-][N+](=O)c(c(N)ccc1Cl)c1	8.594 81	4.31 00	5.36 92	0.0904	YES
#	135	Oc(c(ccc1)C(CC)C)c1	5.383 22	4.34 00	4.18 40	0.0274	YES
#	136	n(ccc(c1)C=C)c1	5.420 06	4.36 00	4.19 76	0.0242	YES
*	137	Nc(c(c(cc1)Cl)Cl)c1	6.333 80	4.38 00	4.53 48	0.0218	YES
#	138	Nc(ccc(c1Cl)Cl)c1	6.108	4.40	4.45	0.0211	YES

			10	00	15		
*	139	Oc(c(ccc1)C(C)(C)C)c1C	5.935 21	4.42 00	4.38 77	0.0373	YES
#	140	Oc(c(cc(c1Cl)Cl)Cl)c1	6.322 88	4.42 00	4.53 08	0.0302	YES
-	141	[O-][N+](=O)c(c(O)ccc1Cl)c1	7.371 79	4.45 00	4.91 78	0.0913	YES
+	142	Oc(ccc(c(ccc(O)c1)c1)c2)c2	6.234 20	4.51 00	4.49 80	0.0469	YES
+	143	Nc(c(cc(c(ccc(N)c1C)c1)c2)C)c 2	7.052 45	4.53 00	4.80 00	0.0457	YES
*	144	Nc(c(cc(c1)Cl)Cl)c1	6.333 80	4.61 00	4.53 48	0.0218	YES
+	145	Oc(c(ccc1Cl)Cl)c1	5.476 04	4.65 00	4.21 82	0.0254	YES
*	146	Oc1ccc(Cl)c(Cl)c1Cl	6.918 44	4.68 00	4.75 05	0.0291	YES
+	147	Nc1ccc(Cl)c(Cl)c1Cl	7.550 50	4.74 00	4.98 38	0.0248	YES
+	148	Nc(c(cc(c1Cl)Cl)Cl)c1	6.954 94	4.80 00	4.76 40	0.0259	YES
-	149	Oc(ccc(c(cccc1)c1)c2)c2	6.345 04	4.85 00	4.53 89	0.0423	YES
-	150	Oc(ccc(c1)CCCCC)c1	7.248 48	4.87 00	4.87 23	0.0257	YES
+	151	[O-][N+](=O)c(ccc(c1Cl)Cl)c1	8.329 47	4.89 00	5.27 12	0.0916	YES
-	152	Oc1ccc(cc1)C(c2ccc(cc2)O)(c3 ccc(cc3)O)C	11.06 148	4.90 00	6.27 94	3.0717	No
-	153	Oc(c(ccc1C)C(C)(C)C)c1	6.316 70	4.94 00	4.52 85	0.0372	YES
-	154	Oc1cc(Cl)cc(Cl)c1Cl	6.918 44	4.94 00	4.75 05	0.0291	YES
+	155	Oc(c(cc(c1)C)C(C)(C)C)c1	6.042 96	4.96 00	4.42 75	0.0354	YES
*	156	Oc(c(cc(c1)C(C)(C)C)C(C)(C)C)c1	7.775 73	5.31 00	5.06 69	0.0461	YES
-	157	Oc(cccc1C)c1	4.308 94	2.87 00	3.78 75	0.0234	YES
-	158	Nc(cccc1)c1	2.531 73	2.93 00	3.13 17	0.0104	YES
*	159	Oc(c(ccc1)C)c1	4.035 20	2.93 00	3.68 65	0.0216	YES
+	160	Nc(c(ccc1)CC)c1	3.610 04	3.39 00	3.52 96	0.0144	YES
-	161	Oc(c(ccc1)C)c1C	4.202 44	3.41 00	3.74 82	0.0266	YES
+	162	Oc(c(c(cc1)C)C)c1	4.310 19	3.48 00	3.78 80	0.0246	YES
#	163	Oc(c(ccc1C)C)c1	4.583	3.58	3.88	0.0264	YES

			93	00	90		
*	164	Nc(c(c(cc1)C)C)c1	3.081 71	3.59 00	3.33 47	0.0165	YES
*	165	Oc(cc(cc1C)C)c1	4.583 93	3.65 00	3.88 90	0.0264	YES
+	166	Nc(c(cc(c1)C)C)c1C	3.248 95	3.70 00	3.39 64	0.0214	YES
*	167	Oc(ccc(c1)CC)c1	4.838 52	3.75 00	3.98 30	0.0226	YES
#	168	Oc(c(ccc1C)C)c1C	4.751 17	3.98 00	3.95 07	0.0314	YES
#	169	Oc(c(ccc1)Cl)c1Cl	5.227 95	4.01 00	4.12 67	0.0253	YES
+	170	Oc(c(ccc1Cl)Cl)c1Cl	5.849 09	4.39 00	4.35 59	0.0295	YES
-	171	Oc(ccc(C(=C)C)c1)c1	6.352 10	4.40 00	4.54 15	0.0272	YES
*	172	Nc(cc(cc1Cl)Cl)c1	6.108 10	4.57 00	4.45 15	0.0211	YES
-	173	Oc(ccc(c1ccc2)c2)c1	6.380 82	4.84 00	4.55 21	0.0422	YES
#	174	Oc1ccc(Cl)c(Cl)c1	6.545 39	4.87 00	4.61 29	0.0250	YES
*	175	Oc1cc(Cl)cc(Cl)c1	6.545 39	4.89 00	4.61 29	0.0250	YES
#	176	Nc1cc(Cl)c(Cl)c(Cl)c1	8.024 29	5.14 00	5.15 86	0.0255	YES
-	177	Oc(c(cc(c1)Br)Br)c1Br	8.440 78	5.24 00	5.31 23	6.0425	No
+	178	CCC(C)c1cccc(C(C)CC)c1O	8.069 67	5.27 00	5.17 54	0.0321	YES
-	179	Oc(c(cc(c1)Br)Br)c1	7.307 23	5.36 00	4.89 40	5.0333	No
-	180	N(c(cccc1)c1)c(cccc2)c2	7.849 76	5.60 00	5.09 42	0.0377	YES
*	181	Oc(ccc(c1)CCCCCCCC)c1	9.658 43	6.17 00	5.76 17	0.0287	YES
-	182	N#CC(=C)C	4.097 75	3.43 00	3.70 96	1.0109	YES
#	183	CCc1cccc(CC)c1N(C(=O)CCl) CCOCCC	12.87 994	7.99 00	6.95 05	0.0547	YES
-	184	C(#N)C=C	2.905 52	3.72 00	3.26 96	1.0089	YES
#	185	O=C1N(CCOC(=O)C=C)C(=O) N(C(=O)N1CCOC(=O)C=C)C COC(=O)C=C	4.408 79	4.21 00	3.82 44	0.0951	YES
-	186	C1Oc2cc3N(CC)C=C(C(=O)O) C(=O)c3cc2O1	5.714 18	4.21 00	4.30 61	0.0720	No
-	187	Fc1cc2C(=O)C(C(=O)O)=CN3c 2c(CCC3C)c1	7.349 13	4.72 00	4.90 95	4.0821	No

-	188	N(=Nc(cccc1)c1)c(ccc(N)c2)c2	7.43833	4.8300	4.9424	1.0467	YES
+	189	N#CC(=C)Cl	10.32366	6.1000	6.0072	1.0125	YES
#	190	CCc1cccc(C)c1N(C(C)COC)C(=O)CCl	10.34476	6.6000	6.0150	0.0552	YES
+	191	[O-][N+](=O)c(ccc(c1)CCl)c1	7.94945	6.6500	5.1310	0.0914	YES
*	192	N(Nc(cccc1)c1)c(cccc2)c2	8.20507	5.2200	5.2253	0.0416	YES
+	193	N#CC=CC#N	6.14294	5.3100	4.4644	1.0088	YES
#	194	C(C=CC#N)(F)(F)F	7.18419	5.5300	4.8486	0.0770	YES
+	195	N(=C=S)C	9.19053	5.5900	5.5890	0.0217	YES
*	196	CCCCOCN(C(=O)CCl)c1c(CC)cccc1CC	12.34413	7.9800	6.7528	0.0504	YES
+	197	OCc(cccc1)c1	1.61930	2.1500	2.7950	0.0200	YES
+	198	Brc2c(c(cc(c2)OCC1OC1)C)Br	9.38542	5.7200	5.6609	3.0636	No
-	199	OC(CCl)CCl	0.88302	2.3100	2.5233	0.0209	YES
+	200	O(CCCl)CCCl	2.85393	2.6200	3.2506	0.0228	YES
#	201	S(SC)C	3.84256	3.5100	3.6154	0.0211	YES
-	202	O(C1c(cccc2)c2)C1	2.13791	3.6800	2.9864	0.0439	YES
+	203	O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F	5.38509	4.1600	4.1847	1.4421	No
-	204	C(=CCl)CCl	3.26146	4.7200	3.4010	1.0192	YES
*	205	C(Cl)(Cl)(Cl)C(Cl)COCC(Cl)C(Cl)(Cl)Cl	10.16988	5.5000	5.9504	0.0576	YES
+	206	C(C(Cl)CCl)=C	3.43164	3.3300	3.4638	1.0202	YES
+	207	c(cccc1)(c1)C(Cl)Cl	6.26294	3.7800	4.5086	0.0260	YES
-	208	S(SCC)CC	5.44919	4.4400	4.2083	0.0232	YES
-	209	c(c(ccc1)Cl)(c1)CCl	6.23435	5.1300	4.4981	0.0283	YES
+	210	O=C(Oc(c(ccc1)C(CC)C)c1)NC	7.51608	3.8000	4.9711	0.0420	YES
-	211	CC(C)Oc1cc(c(Cl)cc1Cl)N2N=C(OC2=O)C(C)(C)C	14.15907	6.7900	7.4226	1.0829	YES
+	212	Cc1cccc(C)c1N(C(=O)COC)N2	3.800	4.50	3.59	0.0699	YES

		C(=O)OCC2	28	00	98		
*	213	Oc(c(cc(c1)Cl)Cl)c1Cl	6.074 78	4.54 00	4.43 92	0.0301	YES
#	214	Nc(c(cc(c1)Cl)Cl)c1Cl	6.706 84	4.60 00	4.67 24	0.0258	YES
+	215	Oc(c(cc(c1Cl)Cl)Cl)c1Cl	6.695 92	5.04 00	4.66 84	0.0343	YES
+	216	N(N)c(c(cc(c1)Cl)Cl)c1Cl	10.55 222	6.40 00	6.09 15	0.0331	YES
#	217	c1ccc(C)cc1NC(=O)Oc2cccc(NC(=O)OC)c2	11.58 874	6.47 00	6.47 40	0.0789	YES
+	218	CNC(=O)Oc1ccccc1OC(C)C	7.970 19	4.32 00	5.13 87	0.0451	YES
+	219	Oc(c(c(c(c1Cl)Cl)Cl)Cl)c1Cl	7.542 76	5.49 00	4.98 09	0.0391	YES
+	220	COP(=S)(OC)Oc1ccc([N+](=O)[O-])c(C)c1	9.431 29	5.75 00	5.67 79	5.1185	No
#	221	CCN(CC)C(=O)SCc1ccc(Cl)cc1	10.79 537	6.45 00	6.18 12	0.0440	YES
+	222	N1C(=O)NC(=O)NC1=O	3.904 51	2.13 00	3.63 83	1.0508	YES
#	223	Clc1cc(C(F)(F)F)ccc1Oc2cc(OCC)c([N+](=O)[O-])cc2	16.92 792	9.66 00	8.44 44	0.1954	YES
+	224	O=C(Nc(c(ccc1)C)c1)CC(=O)C	4.302 59	2.41 00	3.78 52	0.0362	YES
-	225	N(CCNC1)C1	3.483 38	2.82 00	3.48 29	0.0253	YES
-	226	Oc(cccc1N)c1	7.186 61	2.83 00	4.84 95	1.0214	YES
#	227	O=S(=O)(N)c(c(ccc1)C)c1	2.307 73	3.00 00	3.04 90	0.0379	YES
-	228	N#CCCCCCCC#N	5.100 67	3.05 00	4.07 97	1.0118	YES
-	229	n1nc(N)nc1	4.980 18	3.11 00	4.03 53	0.0441	YES
-	230	OC(CC(NC1(C)C)(C)C)C1	3.976 27	3.12 00	3.66 48	1.0375	YES
+	231	O=S(=O)(O)c(c(ccc1[N+](=O)[O-])C)c1	3.711 98	3.26 00	3.56 72	1.1110	YES
-	232	n(ccc1)c(CCC2)c1C2	5.760 54	3.32 00	4.32 32	0.0387	YES
+	233	OCC(N)(CC)CO	3.641 16	3.37 00	3.54 11	1.0165	YES
+	234	C1(C#N)=CCCCC1	6.338 58	3.42 00	4.53 65	1.0208	YES
-	235	[O-][N+](=O)c(c(OC)ccc1)c1	2.925 56	3.43 00	3.27 70	0.0876	YES
*	236	NC(CCCC1)C1	2.569 07	3.46 00	3.14 55	0.0191	YES
+	237	n(ccc(N)c1)c1	4.670	3.50	3.92	0.0195	YES

			22	00	09		
-	238	Nc(cccc1N)c1	4.06674	3.5600	3.6982	0.0173	YES
+	239	Cc1ncc([N+](=O)[O-])n1CCO	5.30061	3.6300	4.1535	0.1230	YES
-	240	NCc(cccc1CN)c1	4.73678	3.6900	3.9454	0.0263	YES
*	241	Nc(c(ccc1N)C)c1	4.34173	3.8300	3.7996	0.0203	YES
-	242	O(c(ccc(N)c1)c1)c(ccc(N)c2)c2	8.86501	3.8500	5.4689	1.0431	YES
#	243	N#Cc(c(ccc1)Cl)c1	5.57680	3.8600	4.2554	0.0215	YES
#	244	n(c(N)ccc1)c1	4.67022	3.8900	3.9209	0.0195	YES
*	245	[O-][N+](=O)c(cccc1OC)c1	4.74800	3.9300	3.9496	0.0873	YES
-	246	N(O)=C(CCCC1)C1	4.21590	4.0600	3.7532	0.0296	YES
-	247	[O-][N+](=O)c(c(c([N+](=O)[O-])cc1)C)c1	6.59931	4.0800	4.6328	0.1629	YES
*	248	[O-][N+](=O)c(c(N)ccc1OC)c1	5.86017	4.1500	4.3600	0.0910	YES
+	249	Nc1c(Cl)ccc(C(=O)O)c1	7.00058	4.1900	4.7808	0.0299	YES
-	250	N(N)(C)C	3.32600	4.2500	3.4248	0.0141	YES
+	251	N=C(Nc(cccc1)c1)Nc(cccc2)c2	6.09629	4.4400	4.4471	1.0454	YES
*	252	N=C(Nc(c(ccc1)C)c1)Nc(c(ccc2)C)c2	6.64627	4.4400	4.6501	1.0514	YES
+	253	[O-][N+](=O)c(c(O)ccc1N)c1	4.80013	4.4900	3.9688	0.0938	YES
+	254	O=S(=O)(Nc(nccc1)n1)c(ccc(N)c2)c2	4.92076	4.5100	4.0133	0.0844	YES
-	255	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C)c1	6.76221	4.5500	4.6929	0.1693	YES
-	256	NC(C(CC(C1)CC(CCC(N)C2C)C2)C)C1	6.06199	4.5900	4.4345	0.0543	YES
-	257	COCC(=O)N(C(C)C(=O)OC)c1c(C)cccc1C	4.81328	4.6500	3.9737	0.0523	YES
+	258	N(c(cccc1)c1)(CC)CC	6.91328	4.7300	4.7486	0.0232	YES
-	259	NCc1cc(cc(c1)C(F)(F)F)C(F)(F)F	7.37515	4.8300	4.9191	0.1575	YES
+	260	CC(C)(NC(=O)c1cc(Cl)cc(Cl)c1)C#C	7.19591	4.8800	4.8529	1.0457	YES
-	261	O=C(N)C(=C(O)C(N(C)C)C(C1(O)C(O)=C(C2C(O)(c(c3c(O)c	7.66193	5.0100	5.0249	9.1261	No

		c4)c4)C)C3(=O))C2O)C1(=O)					
-	262	Nc(c(N)ccc1)c1	3.643 90	5.12 00	3.54 21	0.0141	YES
-	263	N(N)C	3.177 80	5.22 00	3.37 01	0.0101	YES
+	264	Nc(c(cc(c(ccc(N)c1Cl)c1)c2)Cl) c2	9.805 10	5.26 00	5.81 58	0.0485	YES
+	265	N(c(c(Sc1cccc2)ccc3)c3)c12	9.837 66	5.43 00	5.82 78	3.0457	No
-	266	n(c(c(ccc1)cc2)c1O)c2	5.521 92	5.45 00	4.23 52	0.0452	YES
*	267	O(c(ccc(NC(C=C1C)(C)C)c12) c2)CC	8.651 11	5.49 00	5.38 99	0.0665	YES
-	268	c(c(c(N)cc1)c(N)cc2)(c1)c2	8.723 73	5.52 00	5.41 67	0.0409	YES
-	269	N(c(c(S1)ccc2)c2)=C1S	10.86 456	5.52 00	6.20 68	0.0609	YES
#	270	c1c(Cl)cc(Cl)cc1N2C(=O)C(C)(C3)C3(C)C2(=O)	9.398 99	5.70 00	5.66 59	0.0920	YES
-	271	N(CCNCCNCCNCCN)CCN	10.11 564	5.74 00	5.93 04	0.0465	YES
-	272	Oc(c(N)ccc1)c1	6.763 78	5.86 00	4.69 35	1.0182	YES
+	273	n(c(nc(n1)NCC)NCC)c1Cl	10.91 096	5.96 00	6.22 39	0.0639	YES
-	274	Oc(ccc(N)c1)c1	6.763 78	6.04 00	4.69 35	1.0182	YES
+	275	N#CCCCCCCCCCCC	7.079 65	6.08 00	4.81 00	1.0148	YES
+	276	c12c(nc3c(o2)cccc3)cc(c(c1)=O)N	10.56 385	6.14 00	6.09 58	7.0597	No
+	277	CCCN(CCC)c1c(cc(cc1[N+](= O)[O-])C(F)(F)F)[N+](=O)[O-]	10.46 978	6.18 00	6.06 11	2.2410	No
+	278	N(CCNCCNCCN)CCN	8.226 70	6.20 00	5.23 33	0.0367	YES
#	279	n(c(nc(n1)NC(C)C)NCC)c1Cl	10.65 234	6.22 00	6.12 85	0.0676	YES
#	280	N(c(c(S1)ccc2)c2)=C1SNC(CC CC3)C3	11.02 805	6.25 00	6.26 71	0.0881	YES
+	281	O=N(CCCCCCCCCCCC)(C)C	8.619 86	6.32 00	5.37 84	1.0317	YES
-	282	c1c(C)cc2nc3SC(=O)Sc3nc2c1	10.76 028	6.38 00	6.16 83	3.0660	No
-	283	N(c(c(S1)ccc2)c2)=C1SNC(C)(C)C	10.24 027	6.40 00	5.97 64	0.0773	YES
+	284	C1(C(C)(C(C)CC3)C(OC(CSC CN(CC)CC)=O)CC(C=C)(C)C(O)C2C)C23CCC1=O	12.18 227	6.48 00	6.69 31	2.1178	No
+	285	O(c(nc(nc1NC(C)(C)C)NCC)n1)C	10.27 302	6.58 00	5.98 85	1.0814	YES

+	286	CSc1nc(NC(C)(C)C)nc(N)n1	12.31131	6.6700	6.7407	1.0744	YES
+	287	n(c(nc(n1)NC(C)(C)C)NCC)c1Cl	11.84042	6.8000	6.5669	0.0736	YES
-	288	O=C(Nc(ccc(c1)Cl)c1)Nc(ccc(c2Cl)Cl)c2	12.72690	6.8200	6.8940	0.0584	YES
+	289	n(c(c(c1cccc2)ccc3)c3)(c12)C=C	9.94858	6.9600	5.8688	2.0578	No
+	290	[O-][N+](=O)c(c(c(c(c(c1)ccc2)c2cc3)c3c4)c1)c4	13.95011	7.7500	7.3454	8.1174	No
-	291	CCNc1nc(NC(C)(C)C)nc(SC)n1	15.13306	7.8600	7.7820	1.0806	YES
-	292	Clc1nc(N)nc(NCC)n1	11.37982	7.9400	6.3969	0.0623	YES
+	293	N(CCCCCCCC)(CCCCCCCC)CCCCCCCC	19.35852	8.2100	9.3413	0.0346	YES
-	294	N(CCCCCCCCCCCCCCCCCC)(C)C	16.14524	8.2200	8.1555	0.0306	YES
-	295	C(#N)c(c(C(#N))ccc1)c1	3.01793	2.9600	3.3111	1.0228	YES
-	296	O(CCNC1)C1	2.27428	3.1800	3.0367	1.0234	YES
-	297	n(c(c(ccc1)cc2)c1)c2	5.97818	3.2900	4.4035	0.0389	YES
-	298	S(=O)(=O)C(N)=N	4.07807	3.3400	3.7024	0.0263	YES
+	299	COc2cc(Cc1cnc(N)nc1N)cc(OC)c2OC	4.24865	3.3500	3.7653	2.0853	No
-	300	n(c(nc(n1)c(cccc2)c2)N)c1N	7.99488	3.4200	5.1478	0.0747	YES
+	301	NC(=S)NN	3.65918	3.6800	3.5478	0.0283	YES
-	302	N(O)=C(CC)C	3.35543	3.7400	3.4357	0.0215	YES
-	303	c12n(=O)c(C(=O)NCCO)c(C)n(=O)c1cccc2	5.00728	3.8200	4.0453	0.0720	YES
+	304	O=S(=O)(N)c(ccc(N)c1)c1	3.14491	3.8700	3.3580	0.0386	YES
-	305	O(c(ccc(N)c1)c1)C	5.06151	3.9800	4.0653	1.0189	YES
-	306	OCCN	0.98813	4.3900	2.5621	1.0062	YES
+	307	O(CCN(c(c(OCC)cc(N)c1OCC)c1)C2)C2	8.36060	4.5600	5.2827	1.0522	YES
+	308	Nc1cccc2c(N)cccc12	7.39480	4.8800	4.9263	0.0312	YES
+	309	[O-][N+](=O)c(cc([N+](=O)[O-])c(O)c1C(CC)C)c1	8.11022	5.2300	5.1903	0.1751	YES
-	310	N(c(ccc(NC(CC)C)c1)c1)C(CC)	9.512	5.37	5.70	0.0379	YES

		C	94	00	80		
-	311	Fc1ccc(cc1Cl)[N+](=O)[O-]	8.952 46	5.47 00	5.50 12	1.0976	YES
-	312	CN(C)c1ccc(cc1)C(=C2C=CC(C=C2)=N(Cl)(C)C)c3ccccc3	11.97 514	5.52 00	6.61 66	4.0758	No
-	313	n(cccc1N)c1	5.093 05	5.58 00	4.07 69	0.0227	YES
+	314	c1ccc3c(c1)OC=2C(C=C(C(C= 2)=O)NC(C)=O)=N3	9.152 20	5.77 00	5.57 49	5.0696	No
-	315	O=C2C(N)=CC1=Nc3c(OC1=C 2)cc(OC)cc3	9.913 19	6.66 00	5.85 57	5.0647	No
-	316	N=C(NC1C(C(C(OC3C(C(C(O 3)C)(C(=O))O)OC2C(C(O)C(C(O2)CO)O)NC)C(C1O)O)NC(= NOS(=O)(=O)O)N)O)N	11.92 915	6.71 00	6.59 96	1.1861	YES
#	317	O=C1N(N)C(SC)=NN=C1C(C) (C)C	12.89 132	6.74 00	6.95 47	0.0673	YES
+	318	CCNc1nc(NCC)nc(SC)n1	14.20 361	6.88 00	7.43 90	1.0709	YES
+	319	CCC(CC)Nc1c(cc(C)c(C)c1[N+](=O)[O-])[N+](=O)[O-]	13.00 711	7.30 00	6.99 74	1.1763	YES
+	320	n(c(nc(n1)NC(C2)C2)NC(C)(C) C)c1SC	14.35 299	8.04 00	7.49 41	1.0985	YES
+	321	OCCO	-1.623 50	0.23 00	1.59 83	0.0082	YES
-	322	c(c(c(c(ccc1)c2)c1cc3)c3cc4)(c 2c(c5ccc6)c6)c45	20.07 494	9.14 00	9.60 57	8.0549	No
-	323	O=S(=O)(c(ccc(O)c1)c1)c(ccc(O)c2)c2	6.179 53	3.59 00	4.47 79	0.0631	YES
-	324	O=P(OCCOCCCC)(OCCOCCC C)OCCOCCCC	2.384 81	3.80 00	3.07 75	5.0386	No
-	325	SCC	7.234 27	4.32 00	4.86 71	0.0148	YES
*	326	FC(F)(F)c(cccc1)c1	5.863 93	4.43 00	4.36 14	0.0704	YES
#	327	FC(F)(F)c(cccc1C(F)(F)F)c1	7.717 84	4.44 00	5.04 55	0.1449	YES
+	328	CC(C)OP(=O)(OC(C)C)SCc1cc ccc1	6.232 26	4.49 00	4.49 73	5.0471	No
+	329	c(ccc1C(C)C)cc1C(C)C	6.473 04	4.68 00	4.58 62	0.0308	YES
+	330	Oc(c(cc(c1)C(c(cc(c(O)c2Br)Br)c2)(C)C)Br)c1Br	6.874 21	4.76 00	4.73 42	8.0899	No
+	331	O(O)C(C)(C)C	0.587 06	4.91 00	2.41 41	0.0239	YES
-	332	s(c(c(c1cccc2)ccc3)c3)c12	8.611 97	5.12 00	5.37 55	2.0395	No
#	333	Oc(ccc(c1)C(c(cccc2)c2)(C)C)c 1	8.293 29	5.18 00	5.25 79	0.0566	YES

-	334	c(c(cccc1)c1)(cccc2)c2	8.38140	5.3000	5.2904	0.0381	YES
-	335	c(c(c(c(c1)ccc2)c2)ccc3)(c1)c3	9.10091	5.4400	5.5559	2.0497	No
+	336	CC(c1cccc2c1cccc2)C	9.45212	5.8300	5.6855	0.0388	YES
+	337	Sc(cccc1)c1	9.14465	5.8400	5.5721	1.0162	YES
+	338	CCc2ccc(cc2)c1ccccc1	9.55746	6.0800	5.7244	0.0430	YES
#	339	SCCCCCCCC	12.05419	7.0200	6.6458	0.0208	YES
-	340	OCCS	8.14444	5.6600	5.2030	0.0094	YES
+	341	OC(CCC(C(C(CCC(O)C1)C1)(C)C)C2)C2	3.23746	3.4700	3.3921	0.0595	YES
+	342	O=P(OCC(CC)CCCC)(CC(CC)CCCC)O	5.55763	3.5200	4.2484	5.0362	No
*	343	O=S(=O)(c(ccc(c1)C)c1)Cl	4.69012	3.5200	3.9282	0.0379	YES
-	344	c(c(c(c1ccc2)c2)ccc3)(c3)C1	9.24074	5.3400	5.6075	2.0511	No
*	345	c(cccc1)(c1)C(CCCC2)C2	7.93969	5.3700	5.1274	0.0362	YES
+	346	SCCCCC	9.64423	5.8700	5.7564	0.0178	YES
+	347	Oc(ccc(O)c1)c1	3.64937	6.3200	3.5441	0.0232	YES
+	348	SCCCCCCCCCC	13.66082	7.3500	7.2387	0.0228	YES

Figure S4
The general scheme of building up the CORAL model

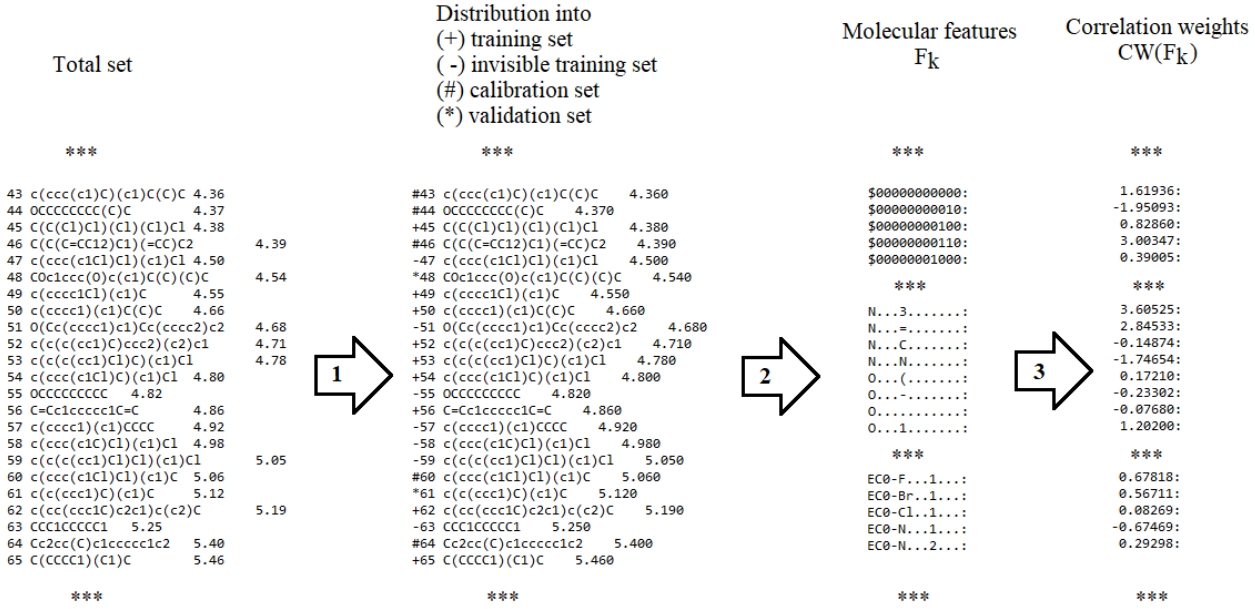

Total set
Distribution into
(+) training set
(-) invisible training set
(#) calibration set
(*) validation set
Molecular features
F_k
Correlation weights
$CW(F_k)$
***
43 c(ccc(c1)C)(c1)C(C)C 4.36
44 OCCCCCCCC(C)C 4.37
45 C(C(Cl)Cl)(Cl)(Cl)Cl 4.38
46 C(C(C=CC12)C1)(=CC)C2 4.39
47 c(ccc(c1Cl)Cl)(c1)Cl 4.50
48 COc1ccc(O)c(c1)C(C)(C)C 4.54
49 c(cccc1Cl)(c1)C 4.55
50 c(cccc1)(c1)C(C)C 4.66
51 O(Cc(cccc1)c1)Cc(cccc2)c2 4.68
52 c(c(c(cc1)C)ccc2)(c2)c1 4.71
53 c(c(c(cc1)Cl)C)(c1)Cl 4.78
54 c(ccc(c1Cl)C)(c1)Cl 4.80
55 OCCCCCCCCC 4.82
56 C=Cc1ccccc1C=C 4.86
57 c(cccc1)(c1)CCCC 4.92
58 c(ccc(c1C)Cl)(c1)Cl 4.98
59 c(c(c(cc1)Cl)Cl)(c1)Cl 5.05
60 c(ccc(c1Cl)Cl)(c1)C 5.06
61 c(c(ccc1)C)(c1)C 5.12
62 c(cc(ccc1C)c2c1)c(c2)C 5.19
63 CCC1CCCCC1 5.25
64 Cc2cc(C)c1ccccc1c2 5.40
65 C(CCCC1)(C1)C 5.46
***
1
#43 c(ccc(c1)C)(c1)C(C)C 4.360
#44 OCCCCCCCC(C)C 4.370
+45 C(C(Cl)Cl)(Cl)(Cl)Cl 4.380
#46 C(C(C=CC12)C1)(=CC)C2 4.390
-47 c(ccc(c1Cl)Cl)(c1)Cl 4.500
*48 COc1ccc(O)c(c1)C(C)(C)C 4.540
+49 c(cccc1Cl)(c1)C 4.550
+50 c(cccc1)(c1)C(C)C 4.660
-51 O(Cc(cccc1)c1)Cc(cccc2)c2 4.680
+52 c(c(c(cc1)C)ccc2)(c2)c1 4.710
+53 c(c(c(cc1)Cl)C)(c1)Cl 4.780
+54 c(ccc(c1Cl)C)(c1)Cl 4.800
-55 OCCCCCCCCC 4.820
+56 C=Cc1ccccc1C=C 4.860
-57 c(cccc1)(c1)CCCC 4.920
-58 c(ccc(c1C)Cl)(c1)Cl 4.980
-59 c(c(c(cc1)Cl)Cl)(c1)Cl 5.050
#60 c(ccc(c1Cl)Cl)(c1)C 5.060
*61 c(c(ccc1)C)(c1)C 5.120
+62 c(cc(ccc1C)c2c1)c(c2)C 5.190
-63 CCC1CCCCC1 5.250
#64 Cc2cc(C)c1ccccc1c2 5.400
+65 C(CCCC1)(C1)C 5.460
***
2
$00000000000:
$00000000010:
$00000000100:
$00000000110:
$00000001000:
***
N...3.....:
N...=.....:
N...C.....:
N...N.....:
O...(.....:
O...-.....:
O.....:
O...1.....:
***
EC0-F...1...:
EC0-Br..1...:
EC0-Cl..1...:
EC0-N...1...:
EC0-N...2...:
***
3
1.61936:
-1.95093:
0.82860:
3.00347:
0.39005:
***
3.60525:
2.84533:
-0.14874:
-1.74654:
0.17210:
-0.23302:
-0.07680:
1.20200:
***
0.67818:
0.56711:
0.08269:
-0.67469:
0.29298:
***

Figure S5
The graphical interpretation for T* and N*

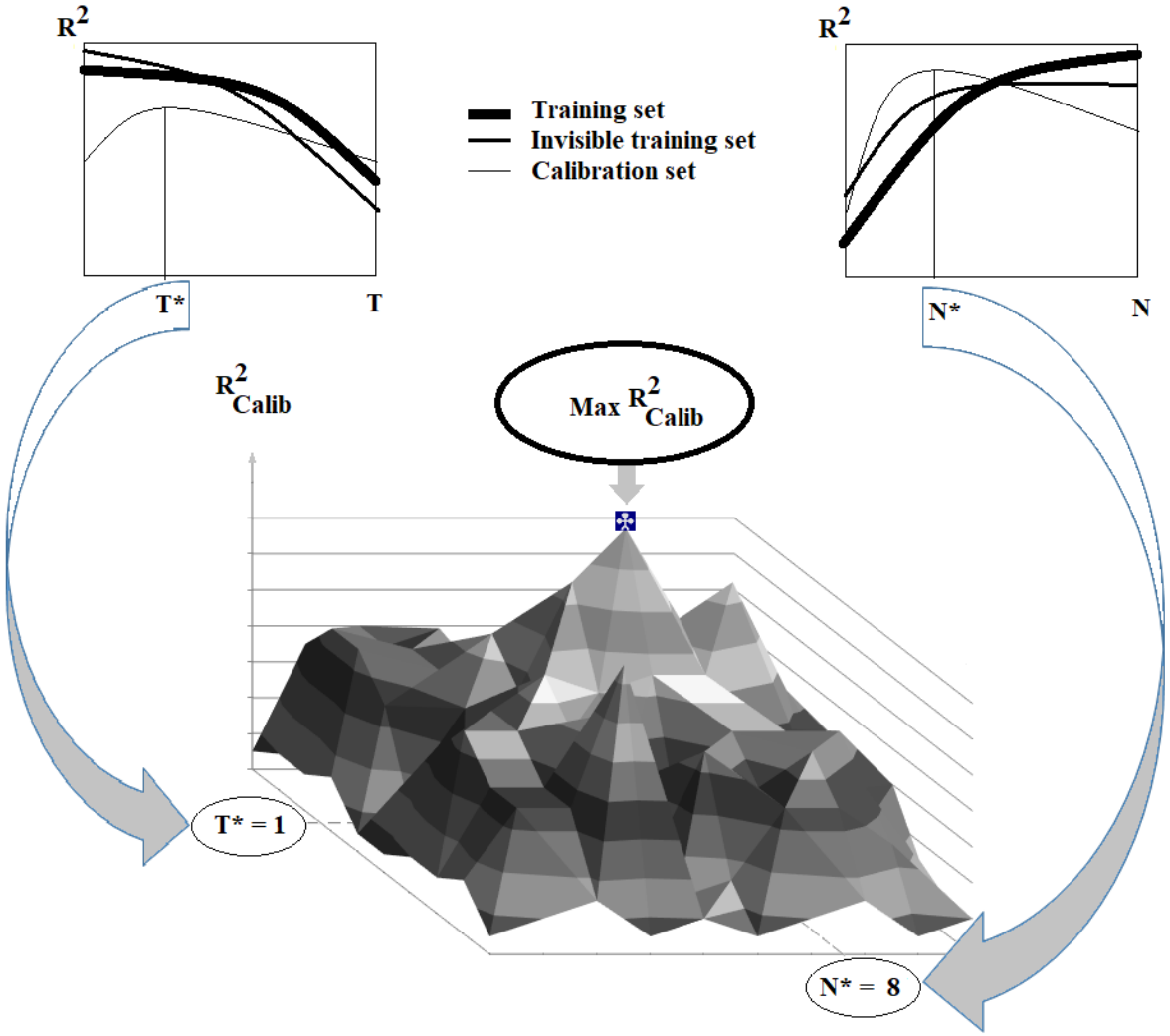


Figure S6
Co-evolution of correlation between observed and calculated endpoint values in the cases (i) Eq. 1 and (ii) Eq. 2

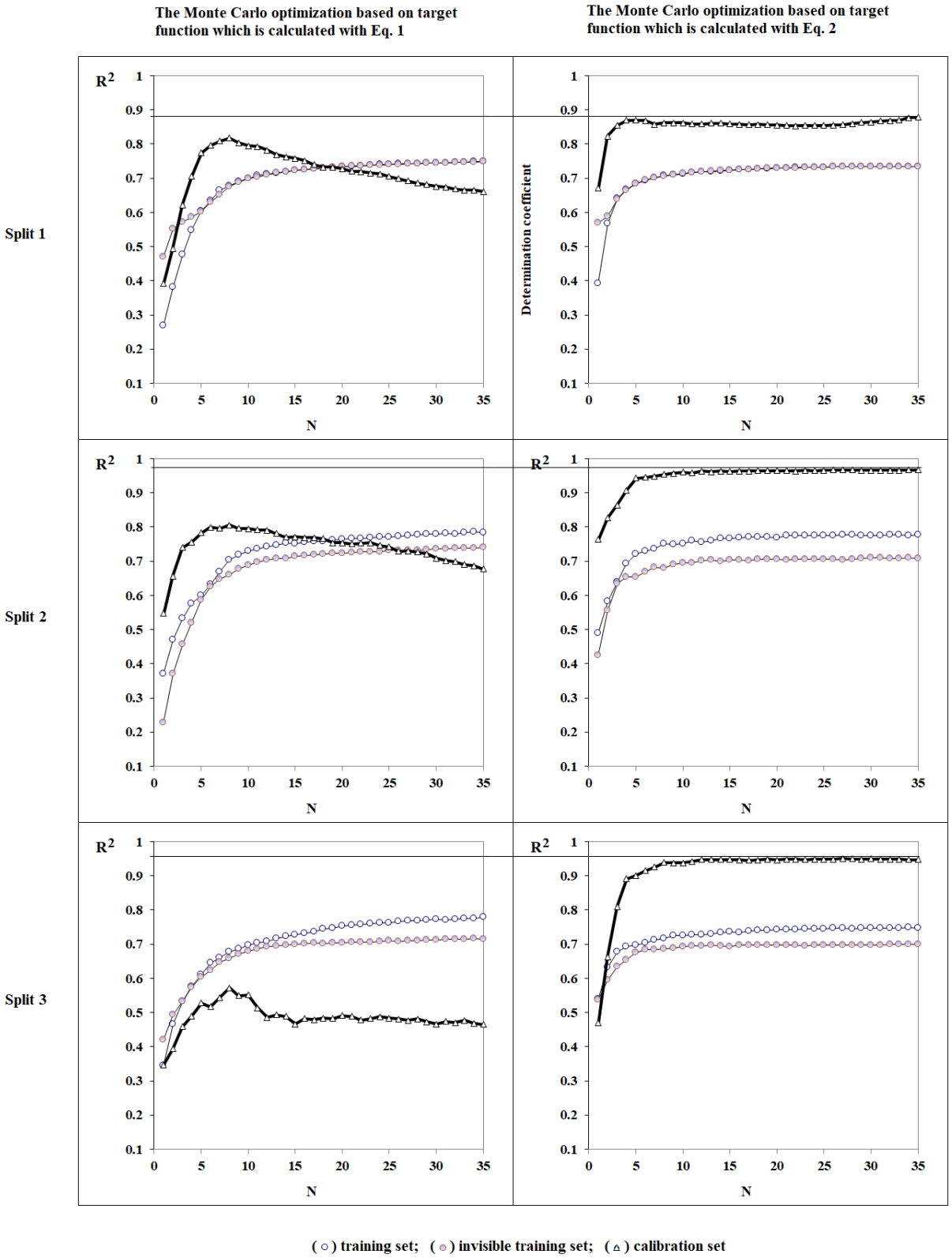
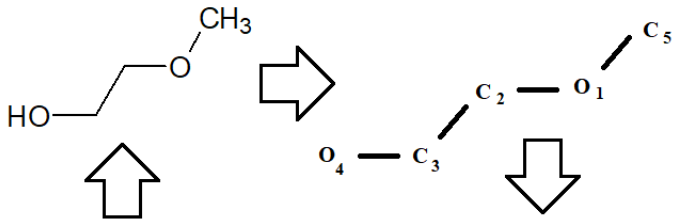


Figure S7
The scheme of translation of SMILES into the adjacency matrix of HSG



O(CCO)C

The adjacency matrix

	O_1	C_2	C_3	O_4	C_5	EC0
O_1	0	1	0	0	1	2
C_2	1	0	1	0	0	2
C_3	0	1	0	1	0	2
O_4	0	0	1	0	0	1
C_5	1	0	0	0	0	1

The CORAL representation by twelve symbols

```
EC0-O...2...
EC0-C...2...
EC0-C...2...
EC0-O...1...
EC0-C...1...
```

Table S8
An example of calculation hybrid optimal descriptor *DCW(T*,N*)* for substance represented by SMILES of "O(CCO)C"

Molecular features	**Correlation weights**
Vertex degree, $EC0_k$	$CW(ECO_k)$
EC0-O...2...	0.6464
EC0-C...2...	-0.3608
EC0-C...2...	-0.3608
EC0-O...1...	-0.2271
EC0-C...1...	-0.3676
S_k	$CW(S_k)$
O.....	-0.3905
(.....	-0.3158
C.....	0.3709
C.....	0.3709
O.....	-0.3905
(.....	-0.3158
C.....	0.3709
SS_k	$CW(SS_k)$
O...(.....	0.4193
C...(.....	0.0563
C...C.....	1.0729
O...C.....	-0.8110
O...(.....	0.4193
C...(.....	0.0563
HARD	***CW(HARD)***
$00001000000	-1.8388
	***DCW(T*,N*)* = -1.5954**

Table S9

The list of promoters of increase or decrease for toxicity to algae (pEC50)

F_k	CWs Run 1	CWs Run 2	CWs Run 3	N1*	N2	N3	Defect of F_k
Split 1 Promoters of increase for pEC50							
C...C.....	1.51416	0.66663	0.57884	49	49	23	0.0007
N...(.....	0.78985	0.38452	0.59511	45	39	10	0.0031
EC0-N...2...	0.57911	0.43432	0.89080	28	25	3	0.0054
c...O.....	2.78847	2.16543	2.18305	25	23	6	0.0028
c...2.....	0.73793	0.30894	0.33841	21	20	8	0.0005
Promoters of decrease for pEC50							
1...(.....	-0.52854	-0.19341	-0.32822	59	51	26	0.0003
O...=.....	-0.02813	-0.36305	-1.02058	28	31	2	0.0062
O...C.....	-0.34167	-0.36676	-1.10351	22	19	9	0.0001
$00001000000	-1.96031	-1.18719	-1.47674	19	22	12	0.0025
S...(.....	-1.04510	-0.41853	-0.59682	11	9	2	0.0039
c...N.....	-1.93641	-1.76189	-2.29432	8	15	7	0.0047
Split 2 Promoters of increase for pEC50							
C...C.....	1.11234	0.01411	1.13848	46	55	17	0.0008
N...(.....	0.92634	0.91987	0.76939	44	38	9	0.0035
EC0-N...2...	0.03905	0.80159	0.55920	29	25	4	0.0049
c...2.....	0.50710	0.35535	0.25323	25	18	6	0.0029
c...O.....	1.67101	1.76977	1.65442	22	25	12	0.0015
Promoters of decrease for pEC50							
1...(.....	-0.31550	-0.66247	-0.71163	60	55	22	0.0008
O...=.....	-0.82137	-0.70379	-0.62499	26	31	11	0.0000
$00001000000	-1.71342	-1.94460	-2.24723	19	23	11	0.0019
O...C.....	-0.65143	-0.61695	-0.24023	19	24	8	0.0001
c...N.....	-1.04827	-0.88506	-1.22974	13	14	5	0.0006
S...(.....	-0.64744	-0.26284	-0.96828	12	8	2	0.0043
Split 3 Promoters of increase for pEC50							
C...C.....	0.50733	0.95926	0.42534	52	44	20	0.0004
N...(.....	0.65440	0.97757	1.04223	41	44	13	0.0014
EC0-N...2...	1.31088	1.62623	0.97053	25	27	7	0.0020
c...2.....	0.07747	0.66693	0.41778	22	28	5	0.0030
c...O.....	2.10593	2.08116	1.94332	18	19	15	0.0044
Promoters of decrease for pEC50							
1...(.....	-0.84637	-0.64421	-0.84516	58	57	21	0.0008
O...=.....	-0.95004	-0.95869	-0.69568	33	22	11	0.0012
O...C.....	-0.83220	-0.69829	-0.97538	20	18	9	0.0005
c...N.....	-1.86497	-1.50865	-1.65916	16	11	7	0.0003
$00001000000	-1.42778	-1.61671	-1.95150	15	21	12	0.0041
S...(.....	-1.32307	-1.20639	-0.60257	10	8	5	0.0011

*) The N1, N2, and N3 are the numbers of a molecular feature in the training, invisible training, and calibration sets, respectively. Defect of F_k is the measure of reliability of a feature: maximal defect is equal to 1 (feature is absolutely unreliable), minimal defect is equal to zero (feature is absolutely reliable) (Gobbi et al. 2016).

Table S10
Criteria of predictive potential

Criterion	**Reference**
$Q^2 = 1 - \frac{\sum(y_k - \acute{y}_k)^2}{\sum(y_k - \bar{y}_k)^2}$	(Golbraikh and Tropsha, 2002)
$\langle R_m^2 \rangle = \frac{R_m^2(x,y) + R_m^2(y,x)}{2}$ $R_m^2(x,y) = r^2(1 - \sqrt{\lvert r^2 - r_0^2 \rvert}$	(Ojha et al., 2011)
$CCC = \frac{2\sum (x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2 + \sum(y - \bar{y})^2 + n(\bar{x} - \bar{y})^2}$	(I-Kuei Lin, 1989)

References

Golbraikh A, Tropsha A. 2002. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J Comput Aided Mol Des. 16 (5-6): 357-369.

I-Kuei Lin L. 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 45(1): 255-268.

Ojha PK, Mitra I, Das RN, Roy K. 2011. Further exploring rm2 metrics for validation of QSPR models. Chemometr Intell Lab Syst. 107(1): 194-205.

Table S11
Comments on molecular features, which are promoters of increase or decrease for toxicity to algae

F_k	**Comments**
C...C.....	A pair connected carbon atoms (sp^3)
N...(.....	Ring with nitrogen (sp^3)
EC0-N...2...	Vertex degree equal to 2 in HSG, which is image of nitrogen
c...O.....	Carbon (sp^2) connected with oxygen
c...2.....	Two rings with carbon (sp^2)
1...(.....	Ring
O...=.....	Double covalent bond started from oxygen
O...C.....	Carbon (sp^3) connected with oxygen
$00001000000	Presense of oxygen accompanied absence of halogens, double bonds, nitrogen, sulfur, and phosphorus
S...(.....	Ring contains sulfur
c...N.....	Carbon (sp^2) connected with nitrogen (sp^3)

Figure S12
Graphical representation of models for toxicity to algae which are calculated using the target function calculated with Eq. 2 (i.e. with using of the *IIC*)

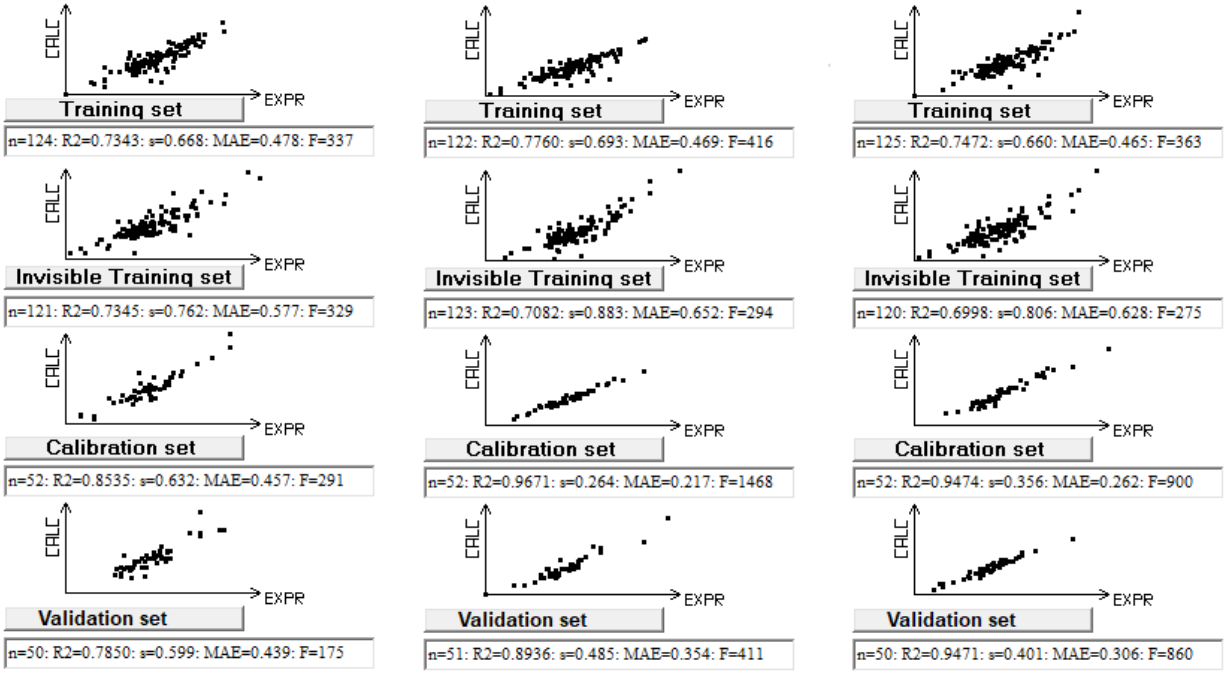


Table S13
Correlation weights of model based on Split 1

SAk	CW(SAk)	ID	N1	N2	N3	DEFECT[SAk]
#.....	-3.84406	1	2	8	1	0.0010
$00000000000	1.45619	2	10	6	8	0.0041
$00000000010	-0.29257	3	1	0	0	1.0000
$00000000100	0.33029	4	10	8	3	0.0018
$00000000110	4.03523	5	1	1	1	0.0056
$00000001000	-2.06219	6	1	0	1	0.0056
$00000001010	0.0	7	0	1	0	0.0000
$00000100000	-0.16581	8	5	0	1	0.0035
$00001000000	-1.83877	9	19	22	12	0.0025
$00001000010	3.28988	10	4	1	0	1.0000
$00001000100	-0.99483	11	4	8	4	0.0056
$00001100000	3.83189	12	1	0	0	1.0000
$00010000000	0.84789	13	15	16	9	0.0022
$00010000100	3.47883	14	6	2	4	0.0029
$00010001000	-2.37026	15	1	0	0	1.0000
$00010100000	5.79441	16	1	2	0	1.0000
$00011000000	1.30866	17	5	5	2	0.0003
$01010000000	0.0	18	0	3	0	0.0000
$01010000100	0.0	19	0	0	1	0.0000
$10000000000	1.77107	20	2	4	2	0.0056
$10000000100	1.70450	21	2	1	1	0.0010
$10001000000	-0.17148	22	3	8	0	1.0000
$10001010000	-0.76105	23	1	1	0	1.0000
$10001100000	0.0	24	0	1	0	0.0000
$10001100100	0.0	25	0	0	1	0.0000
$10001101000	-14.42152	26	1	0	0	1.0000
$10001110000	0.0	27	0	1	0	0.0000
$10010000000	-2.59591	28	1	1	0	1.0000
$10010000100	0.0	29	0	1	0	0.0000
$10010100000	0.57410	30	2	2	1	0.0010
$10011000000	-0.10375	31	18	9	1	0.0066
$10011000100	6.46734	32	2	5	0	1.0000
$10011001000	-6.24319	33	2	0	0	1.0000
$10011001100	0.0	34	0	2	0	0.0000
$10011100000	-0.54514	35	2	5	0	1.0000
$10011100100	3.65811	36	1	0	0	1.0000
$10011110000	-4.97581	37	1	0	0	1.0000
$11010000000	4.21766	38	1	3	0	1.0000
$11010000100	12.26595	39	1	0	0	1.0000
$11010001000	0.0	40	0	1	0	0.0000
$11011000100	0.0	41	0	1	0	0.0000
(...#.....	0.0	42	0	2	0	0.0000
(...(.....	2.58555	43	34	33	17	0.0010

(.....	-0.31576	44	113	112	50	0.0003
+.....	1.16966	45	9	9	1	0.0053
-.....	-0.06011	46	9	9	1	0.0053
1...(.....	-0.66226	47	59	51	26	0.0003
1.....	-0.02337	48	90	90	39	0.0002
2...(.....	-0.45978	49	20	21	10	0.0010
2.....	-0.75424	50	24	28	11	0.0005
2...1.....	2.39261	51	3	4	1	0.0012
3...(.....	-0.81646	52	7	7	1	0.0047
3.....	1.29010	53	9	11	1	0.0053
3...2.....	-2.92742	54	1	1	0	1.0000
4...(.....	5.29901	55	1	1	0	1.0000
4.....	-0.01850	56	1	1	0	1.0000
5.....	0.0	57	0	1	0	0.0000
5...4.....	0.0	58	0	1	0	0.0000
6...(.....	0.0	59	0	1	0	0.0000
6.....	0.0	60	0	1	0	0.0000
=...(.....	-0.26434	61	33	31	6	0.0039
=.....	0.19050	62	40	46	6	0.0045
=...1.....	0.88910	63	1	2	0	1.0000
=...2.....	-1.62110	64	1	0	0	1.0000
C...#.....	1.11095	65	2	6	1	0.0010
C...(.....	0.05634	66	80	81	37	0.0006
C.....	0.37088	67	96	91	42	0.0002
C...1.....	0.12819	68	19	27	8	0.0000
C...2.....	0.44696	69	4	8	3	0.0036
C...3.....	-1.37272	70	5	2	0	1.0000
C...=.....	0.73574	71	21	25	4	0.0037
C...C.....	1.07294	72	49	49	23	0.0007
F...(.....	0.33912	73	4	3	1	0.0026
F.....	0.77548	74	5	4	1	0.0035
F...C.....	-0.37940	75	1	0	1	0.0056
EC0-C...1...	-0.36760	76	71	69	33	0.0006
EC0-C...2...	-0.36084	77	114	114	49	0.0001
EC0-C...3...	-0.17302	78	99	102	41	0.0001
EC0-C...4...	0.40222	79	14	13	4	0.0020
EC0-F...1...	-1.07701	80	5	4	1	0.0035
EC0-Br..1...	0.59131	81	6	3	1	0.0042
EC0-Cl..1...	-0.19647	82	27	29	15	0.0017
EC0-N...1...	-0.23756	83	30	33	13	0.0002
EC0-N...2...	1.06942	84	28	25	3	0.0054
EC0-N...3...	-0.13908	85	20	17	4	0.0035
EC0-N...4...	0.0	86	0	1	0	0.0000
EC0-O...1...	-0.22706	87	61	59	15	0.0027
EC0-O...2...	0.64644	88	17	22	8	0.0007
EC0-P...4...	-1.42908	89	2	2	0	1.0000
EC0-S...1...	5.13312	90	7	1	2	0.0020
EC0-S...2...	0.58217	91	7	6	2	0.0020
EC0-S...4...	-4.47216	92	2	4	1	0.0010

EC0-o...2...	0.0	93	0	1	0	0.0000
EC0-s...2...	0.0	94	0	1	0	0.0000
Br..(.....	-0.22903	95	5	2	1	0.0035
Br.....	-0.31903	96	6	3	1	0.0042
Br..1.....	0.37408	97	2	0	0	1.0000
Br..2.....	-9.29393	98	1	0	0	1.0000
Br..C.....	1.17099	99	2	1	1	0.0010
Cl..(.....	-0.03653	100	25	24	13	0.0013
Cl.....	1.11629	101	27	29	15	0.0017
Cl..1.....	-0.69713	102	12	7	6	0.0010
Cl..C.....	0.30896	103	5	6	2	0.0003
N...#.....	1.81555	104	2	7	1	0.0010
N...(.....	0.65469	105	45	39	10	0.0031
N...+.....	0.57155	106	9	9	1	0.0053
N.....	0.47500	107	59	58	18	0.0017
N...1.....	-0.08032	108	6	9	0	1.0000
N...2.....	0.0	109	0	2	0	0.0000
N...3.....	4.35846	110	2	1	0	1.0000
N...=.....	2.74371	111	4	3	0	1.0000
N...C.....	-0.17676	112	22	13	4	0.0039
N...N.....	-2.37254	113	1	1	0	1.0000
O...(.....	0.41927	114	36	42	9	0.0026
O...-.....	0.61266	115	9	9	1	0.0053
O.....	-0.39051	116	64	69	20	0.0016
O...1.....	2.08657	117	3	4	3	0.0056
O...2.....	-12.02911	118	2	2	0	1.0000
O...3.....	6.11527	119	1	0	0	1.0000
O...=.....	0.02347	120	28	31	2	0.0062
O...C.....	-0.81100	121	22	19	9	0.0001
O...N.....	4.23775	122	1	0	0	1.0000
P...(.....	-1.81768	123	2	2	0	1.0000
P.....	-0.73100	124	2	2	0	1.0000
P...=.....	-0.47033	125	1	1	0	1.0000
P...O.....	1.43488	126	1	1	0	1.0000
S...(.....	-0.82668	127	11	9	2	0.0039
S.....	1.71627	128	16	11	4	0.0026
S...1.....	-0.75682	129	1	2	1	0.0056
S...3.....	0.0	130	0	1	0	0.0000
S...=.....	0.46309	131	3	5	1	0.0012
S...C.....	2.03853	132	9	5	2	0.0031
S...N.....	2.42671	133	1	1	0	1.0000
S...O.....	1.58052	134	1	0	0	1.0000
[...(.....	1.07451	135	9	9	1	0.0053
[...+.....	-0.28772	136	9	9	1	0.0053
[...-.....	-0.97318	137	9	9	1	0.0053
[.....	0.05899	138	9	9	1	0.0053
[...1.....	-0.26030	139	1	2	0	1.0000
[...N.....	0.18235	140	9	9	1	0.0053

[...O.....	-0.46228	141	9	9	1	0.0053
[...[.....	-2.61044	142	6	5	1	0.0042
c...(.....	0.82964	143	83	79	35	0.0000
c.....	0.01017	144	85	80	35	0.0001
c...1.....	1.25804	145	83	76	34	0.0001
c...2.....	1.45641	146	21	20	8	0.0005
c...3.....	-0.52491	147	6	9	1	0.0042
c...4.....	1.37422	148	1	1	0	1.0000
c...5.....	0.0	149	0	1	0	0.0000
c...6.....	0.0	150	0	1	0	0.0000
c...C.....	0.18652	151	9	5	4	0.0003
c...F.....	2.14866	152	1	2	0	1.0000
c...Br.....	0.0	153	0	1	0	0.0000
c...Cl.....	0.0	154	0	2	0	0.0000
c...N.....	-2.09366	155	8	15	7	0.0047
c...O.....	2.62761	156	25	23	6	0.0028
c...S.....	-0.79172	157	2	1	1	0.0010
c...c.....	0.67590	158	82	76	33	0.0002
n...(.....	0.02451	159	8	12	2	0.0026
n.....	-0.95389	160	8	13	2	0.0026
n...1.....	1.03030	161	4	5	2	0.0010
n...2.....	-0.78897	162	1	1	0	1.0000
n...3.....	0.0	163	0	1	0	0.0000
n...c.....	1.27397	164	4	8	2	0.0010
o...(.....	0.0	165	0	1	0	0.0000
o.....	0.0	166	0	1	0	0.0000
o...2.....	0.0	167	0	1	0	0.0000
s...(.....	0.0	168	0	1	0	0.0000
s.....	0.0	169	0	1	0	0.0000

Table S14
Correlation weights of model based on Split 2

SAk	CW(SAk)	ID	N1	N2	N3	DEFECT[SAk]
#.....	-0.08404	1	4	5	0	1.0000
$00000000000	0.18692	2	11	8	4	0.0009
$00000000010	0.02088	3	2	0	0	1.0000
$00000000100	-0.87352	4	7	6	5	0.0032
$00000000110	1.21174	5	1	2	0	1.0000
$00000001000	1.56522	6	1	0	1	0.0055
$00000001010	0.0	7	0	1	0	0.0000
$00000100000	0.0	8	0	4	1	0.0000
$00001000000	-2.04399	9	19	23	11	0.0019
$00001000010	0.78307	10	4	1	0	1.0000
$00001000100	-1.42831	11	9	3	7	0.0038
$00001100000	0.0	12	0	1	0	0.0000
$00010000000	-0.36380	13	17	18	6	0.0010
$00010000100	1.38622	14	4	6	3	0.0036
$00010001000	0.0	15	0	0	1	0.0000
$00010100000	3.99501	16	2	1	0	1.0000

$00011000000	1.10116	17	5	5	0	1.0000
$01010000000	-8.65852	18	2	1	0	1.0000
$10000000000	-0.84868	19	2	3	2	0.0055
$10000000100	-1.05224	20	1	2	0	1.0000
$10001000000	-0.72966	21	4	5	3	0.0036
$10001010000	0.0	22	0	2	0	0.0000
$10001100000	-0.97829	23	1	0	1	0.0055
$10001100100	0.0	24	0	1	0	0.0000
$10001101000	0.0	25	0	1	0	0.0000
$10001110000	0.0	26	0	1	0	0.0000
$10010000000	-7.11945	27	2	0	0	1.0000
$10010000100	-9.93482	28	1	0	0	1.0000
$10010100000	2.02771	29	4	0	0	1.0000
$10011000000	-0.70503	30	9	17	3	0.0013
$10011000100	2.75433	31	5	3	2	0.0004
$10011001000	0.0	32	0	1	1	0.0000
$10011001100	5.22693	33	1	1	0	1.0000
$10011100000	-1.36154	34	5	1	1	0.0036
$10011100100	0.0	35	0	1	0	0.0000
$10011110000	0.05597	36	1	0	0	1.0000
$11010000000	-3.83186	37	1	2	0	1.0000
$11010000100	0.0	38	0	1	0	0.0000
$11010001000	-0.38497	39	1	0	0	1.0000
$11011000100	0.0	40	0	1	0	0.0000
(...#.....	-3.78806	41	1	1	0	1.0000
(...(.....	1.72987	42	41	36	13	0.0016
(..........	-0.48065	43	114	115	49	0.0000
+..........	0.05478	44	7	11	3	0.0000
-..........	0.85320	45	7	11	3	0.0000
1...(.....	-0.19074	46	60	55	22	0.0008
1..........	-0.03636	47	93	93	37	0.0004
2...(.....	0.51428	48	28	20	5	0.0040
2..........	-0.45087	49	34	25	7	0.0035
2...1.....	1.44806	50	1	4	1	0.0055
3...(.....	-0.03793	51	10	2	2	0.0036
3..........	0.95505	52	14	3	3	0.0034
3...2.....	-1.22997	53	1	0	1	0.0055
4...(.....	4.60199	54	2	0	0	1.0000
4..........	0.52067	55	2	0	0	1.0000
=...(.....	0.82289	56	32	31	8	0.0027
=..........	-0.03304	57	38	43	13	0.0012
=...1.....	2.39430	58	1	1	1	0.0055
=...2.....	-4.40284	59	1	1	0	1.0000
C...#.....	-2.13485	60	3	4	0	1.0000
C...(.....	0.00454	61	81	86	35	0.0001
C..........	0.44812	62	92	97	38	0.0002
C...1.....	-0.34626	63	26	24	8	0.0017
C...2.....	-0.30130	64	9	7	2	0.0032
C...3.....	-2.13629	65	5	0	2	0.0004

C...=.....	0.31027	66	20	19	7	0.0011
C...C.....	0.16461	67	46	55	17	0.0008
F...(.....	0.20953	68	2	4	2	0.0055
F.....	0.06996	69	3	4	3	0.0055
F...C.....	-0.94542	70	1	0	1	0.0055
EC0-C...1...	0.62655	71	66	76	30	0.0004
EC0-C...2...	0.38776	72	113	115	48	0.0000
EC0-C...3...	0.04486	73	103	101	44	0.0000
EC0-C...4...	-0.54564	74	16	17	4	0.0027
EC0-F...1...	0.10327	75	3	4	3	0.0055
EC0-Br..1...	0.69202	76	7	4	0	1.0000
EC0-Cl..1...	0.44603	77	29	27	17	0.0019
EC0-N...1...	0.47468	78	31	29	11	0.0010
EC0-N...2...	1.30412	79	29	25	4	0.0049
EC0-N...3...	0.12070	80	16	19	8	0.0009
EC0-N...4...	-3.29931	81	1	1	0	1.0000
EC0-O...1...	0.00040	82	56	59	28	0.0009
EC0-O...2...	-0.59816	83	17	26	5	0.0020
EC0-P...4...	-1.69157	84	1	3	0	1.0000
EC0-S...1...	4.64201	85	3	5	0	1.0000
EC0-S...2...	-0.78410	86	6	6	2	0.0013
EC0-S...4...	-1.36650	87	4	3	1	0.0027
EC0-o...2...	1.98068	88	1	0	0	1.0000
Br..(.....	0.26810	89	6	3	0	1.0000
Br.....	0.52560	90	7	4	0	1.0000
Br..1.....	-0.54355	91	2	0	0	1.0000
Br..2.....	-7.17024	92	1	0	0	1.0000
Br..C.....	-0.19759	93	2	2	0	1.0000
Cl..(.....	0.50123	94	26	23	15	0.0018
Cl.....	0.23971	95	29	27	17	0.0019
Cl..1.....	0.19109	96	10	9	8	0.0040
Cl..2.....	-0.38912	97	1	0	0	1.0000
Cl..C.....	0.38087	98	5	4	4	0.0040
N...#.....	6.56739	99	4	4	0	1.0000
N...(.....	1.14343	100	44	38	9	0.0035
N...+.....	0.32643	101	7	11	3	0.0000
N.....	-0.96980	102	60	59	17	0.0021
N...1.....	0.55682	103	6	7	1	0.0043
N...2.....	2.03189	104	2	1	0	1.0000
N...3.....	5.93057	105	2	0	1	0.0009
N...=.....	4.60754	106	6	3	0	1.0000
N...C.....	0.13089	107	19	15	6	0.0016
N...N.....	-3.20366	108	2	0	0	1.0000
O...(.....	0.35198	109	34	43	13	0.0006
O...-.....	1.85187	110	7	11	3	0.0000
O.....	-0.44612	111	63	68	29	0.0004
O...1.....	1.22129	112	6	4	1	0.0043
O...2.....	-6.85921	113	3	1	0	1.0000
O...3.....	3.42268	114	1	0	0	1.0000

O...=.....	-0.49706	115	26	31	11	0.0000
O...C.....	-0.55249	116	19	24	8	0.0001
O...N.....	4.78173	117	1	0	0	1.0000
P...(.....	-1.84457	118	1	3	0	1.0000
P.....	-1.95948	119	1	3	0	1.0000
P...=.....	0.0	120	0	2	0	0.0000
P...O.....	2.09705	121	1	1	0	1.0000
S...(.....	0.42912	122	12	8	2	0.0043
S.....	0.84482	123	13	14	3	0.0031
S...1.....	0.29159	124	3	0	0	1.0000
S...3.....	2.48785	125	1	0	0	1.0000
S...=.....	-1.08716	126	6	3	1	0.0043
S...C.....	2.20487	127	4	9	2	0.0009
S...N.....	0.60213	128	2	0	0	1.0000
S...O.....	1.98440	129	1	0	0	1.0000
[...(.....	0.51200	130	7	11	3	0.0000
[...+.....	0.64893	131	7	11	3	0.0000
[...-.....	-0.08874	132	7	11	3	0.0000
[.....	0.04809	133	7	11	3	0.0000
[...1.....	0.19109	134	1	2	0	1.0000
[...N.....	-0.79155	135	7	11	3	0.0000
[...O.....	-0.13807	136	7	11	3	0.0000
[...[.....	-1.49125	137	4	7	3	0.0036
c...(.....	0.37929	138	83	82	35	0.0001
c.....	0.08915	139	83	84	36	0.0001
c...1.....	0.84380	140	78	83	36	0.0005
c...2.....	0.55496	141	25	18	6	0.0029
c...3.....	-0.54201	142	10	3	2	0.0036
c...4.....	0.90586	143	2	0	0	1.0000
c...C.....	-0.54159	144	3	10	3	0.0055
c...F.....	-3.23318	145	1	1	1	0.0055
c...Br.....	3.78277	146	1	0	0	1.0000
c...Cl.....	7.12548	147	1	1	0	1.0000
c...N.....	-0.52020	148	13	14	5	0.0006
c...O.....	1.81385	149	22	25	12	0.0015
c...S.....	1.72804	150	1	3	0	1.0000
c...c.....	0.06182	151	77	81	34	0.0002
n...(.....	-0.32907	152	8	11	4	0.0009
n.....	0.19821	153	9	11	4	0.0002
n...1.....	-0.08325	154	6	6	1	0.0043
n...2.....	-3.48547	155	1	1	0	1.0000
n...3.....	1.57860	156	1	0	0	1.0000
n...c.....	-0.88592	157	8	7	1	0.0051
o...(.....	-0.33795	158	1	0	0	1.0000
o.....	2.32760	159	1	0	0	1.0000
o...2.....	0.73433	160	1	0	0	1.0000

Table S15
Correlation weights of model based on Split 3

SAk	CW(SAk)	ID	N1	N2	N3	DEFECT[SAk]
#.....	0.40194	1	5	4	2	0.0002
$00000000000	1.04946	2	11	10	3	0.0022
$00000000010	0.91755	3	1	1	0	1.0000
$00000000100	-0.19946	4	8	10	2	0.0026
$00000000110	2.39721	5	2	0	1	0.0011
$00000001000	0.0	6	0	0	1	0.0000
$00000001010	-0.87344	7	1	0	0	1.0000
$00000100000	-1.66814	8	2	2	2	0.0056
$00001000000	-1.11840	9	15	21	12	0.0041
$00001000010	1.90339	10	3	2	0	1.0000
$00001000100	-0.87055	11	7	3	4	0.0019
$00001100000	0.0	12	0	1	0	0.0000
$00010000000	0.19080	13	15	23	6	0.0002
$00010000100	2.29920	14	7	2	4	0.0019
$00010001000	0.0	15	0	1	0	0.0000
$00010100000	2.95239	16	2	1	0	1.0000
$00011000000	0.77299	17	4	8	0	1.0000
$01010000000	-3.19867	18	1	2	0	1.0000
$01010000100	0.0	19	0	0	1	0.0000
$10000000000	0.71285	20	4	0	3	0.0037
$10000000100	-0.05607	21	2	2	0	1.0000
$10001000000	0.14110	22	5	5	1	0.0035
$10001010000	-0.28052	23	1	1	0	1.0000
$10001100000	0.0	24	0	2	0	0.0000
$10001101000	-6.85503	25	1	0	0	1.0000
$10001110000	0.01740	26	1	0	0	1.0000
$10010000000	-6.31935	27	1	1	0	1.0000
$10010000100	0.0	28	0	1	0	0.0000
$10010100000	1.49896	29	2	2	1	0.0011
$10011000000	-0.26927	30	15	10	3	0.0035
$10011000100	3.21626	31	4	3	3	0.0037
$10011001000	-4.41263	32	1	1	0	1.0000
$10011001100	0.0	33	0	1	1	0.0000
$10011100000	-2.89582	34	4	2	2	0.0011
$10011100100	0.0	35	0	0	1	0.0000
$10011110000	-1.01588	36	1	0	0	1.0000
$11010000000	0.79521	37	2	2	0	1.0000
$11010000100	6.31587	38	1	0	0	1.0000
$11010001000	0.0	39	0	0	1	0.0000
$11011000100	-5.30370	40	1	0	0	1.0000
(...#.....	0.0	41	0	2	0	0.0000
(...(.....	1.73053	42	43	34	10	0.0029
(.....	-0.45236	43	116	111	49	0.0001
+.....	1.15035	44	13	5	2	0.0044
-.....	-0.07609	45	13	5	2	0.0044
1...(.....	-0.22709	46	58	57	21	0.0008
1.....	-0.25782	47	94	87	43	0.0005
2...(.....	-0.01161	48	25	28	4	0.0042

2.....	-0.55642	49	29	35	7	0.0027
2...1.....	1.10256	50	4	2	1	0.0026
3...(.....	0.39266	51	6	8	2	0.0012
3.....	0.31801	52	7	13	2	0.0019
3...2.....	-2.06206	53	2	0	0	1.0000
4...(.....	2.15207	54	1	2	0	1.0000
4.....	1.39050	55	1	2	0	1.0000
5.....	0.0	56	0	1	0	0.0000
5...4.....	0.0	57	0	1	0	0.0000
6...(.....	0.0	58	0	1	0	0.0000
6.....	0.0	59	0	1	0	0.0000
=...(.....	0.32691	60	34	27	12	0.0009
=.....	0.17796	61	46	33	16	0.0010
=...1.....	0.87361	62	2	1	0	1.0000
=...2.....	-3.67646	63	1	1	0	1.0000
C...#.....	1.09457	64	5	2	2	0.0002
C...(.....	0.15440	65	82	74	38	0.0006
C.....	0.17181	66	95	86	43	0.0005
C...1.....	-0.10202	67	18	25	13	0.0034
C...2.....	0.10940	68	7	8	2	0.0019
C...3.....	-0.83169	69	2	3	2	0.0056
C...=.....	0.16224	70	20	20	8	0.0002
C...C.....	0.24078	71	52	44	20	0.0004
F...(.....	-0.14027	72	3	1	3	0.0056
F.....	0.37266	73	3	3	3	0.0056
F...C.....	0.0	74	0	0	1	0.0000
EC0-C...1...	0.21323	75	72	60	36	0.0011
EC0-C...2...	0.39073	76	116	113	49	0.0001
EC0-C...3...	0.27496	77	107	95	45	0.0001
EC0-C...4...	-0.05661	78	14	11	6	0.0002
EC0-F...1...	0.45921	79	3	3	3	0.0056
EC0-Br..1...	-0.52696	80	7	3	1	0.0046
EC0-Cl..1...	0.31913	81	32	22	17	0.0014
EC0-N...1...	-0.26742	82	31	35	13	0.0000
EC0-N...2...	0.75426	83	25	27	7	0.0020
EC0-N...3...	-0.61718	84	20	15	8	0.0002
EC0-N...4...	-7.08510	85	1	1	0	1.0000
EC0-O...1...	-0.10532	86	56	53	26	0.0006
EC0-O...2...	-0.06209	87	18	17	8	0.0004
EC0-P...4...	0.13996	88	3	1	0	1.0000
EC0-S...1...	5.05003	89	6	3	1	0.0041
EC0-S...2...	1.50149	90	6	5	4	0.0029
EC0-S...3...	0.0	91	0	1	0	0.0000
EC0-S...4...	0.74439	92	4	2	1	0.0026
EC0-o...2...	2.51704	93	1	0	0	1.0000
EC0-s...2...	0.0	94	0	1	0	0.0000
Br..(.....	0.15666	95	6	3	0	1.0000
Br.....	0.89510	96	7	3	1	0.0046
Br..1.....	0.88119	97	1	1	0	1.0000

Br..2.....	-6.40944	98	1	0	0	1.0000
Br..C.....	0.07865	99	3	0	1	0.0012
Cl..(.....	0.02101	100	29	18	14	0.0009
Cl.....	0.90455	101	32	22	17	0.0014
Cl..1.....	-0.73485	102	15	8	5	0.0012
Cl..2.....	0.0	103	0	1	0	0.0000
Cl..C.....	-0.45482	104	4	7	3	0.0037
N...#.....	-0.50784	105	4	4	2	0.0011
N...(.....	0.73145	106	41	44	13	0.0014
N...+.....	-0.00434	107	13	5	2	0.0044
N.....	0.33554	108	61	60	23	0.0005
N...1.....	0.62412	109	6	4	6	0.0056
N...2.....	-0.28514	110	1	1	1	0.0056
N...3.....	4.39536	111	1	2	0	1.0000
N...=.....	4.85782	112	3	6	1	0.0012
N...C.....	-0.28336	113	17	17	4	0.0028
N...N.....	-4.38099	114	1	0	1	0.0056
O...(.....	0.32127	115	44	32	13	0.0018
O...-.....	-0.07885	116	13	5	2	0.0044
O.....	-0.22921	117	63	60	27	0.0002
O...1.....	1.61355	118	4	4	1	0.0026
O...2.....	-7.22730	119	1	3	0	1.0000
O...3.....	0.0	120	0	1	0	0.0000
O...=.....	-0.47872	121	33	22	11	0.0012
O...C.....	-0.60095	122	20	18	9	0.0005
O...N.....	0.0	123	0	1	0	0.0000
P...(.....	-1.62323	124	3	1	0	1.0000
P.....	-0.86061	125	3	1	0	1.0000
P...=.....	-0.87539	126	1	1	0	1.0000
P...O.....	-1.07201	127	2	0	0	1.0000
S...(.....	-0.54481	128	10	8	5	0.0011
S.....	0.75915	129	16	10	6	0.0006
S...1.....	-1.07269	130	1	2	1	0.0056
S...3.....	0.0	131	0	1	0	0.0000
S...=.....	-0.59863	132	7	1	1	0.0046
S...C.....	1.90488	133	7	5	4	0.0019
S...N.....	0.0	134	0	1	1	0.0000
S...O.....	0.0	135	0	1	0	0.0000
[...(.....	0.18259	136	13	5	2	0.0044
[...+.....	0.78926	137	13	5	2	0.0044
[...-.....	0.57280	138	13	5	2	0.0044
[.....	0.41126	139	13	5	2	0.0044
[...1.....	-0.66867	140	3	0	0	1.0000
[...N.....	-0.30207	141	13	5	2	0.0044
[...O.....	0.44875	142	13	5	2	0.0044
[...[.....	-2.91295	143	8	4	1	0.0050
c...(.....	0.23629	144	85	79	38	0.0004
c.....	0.25001	145	87	79	39	0.0004
c...1.....	0.54978	146	86	74	38	0.0003

c...2.....	0.48665	147	22	28	5	0.0030
c...3.....	-0.60696	148	5	12	0	1.0000
c...4.....	0.69776	149	1	2	0	1.0000
c...5.....	0.0	150	0	1	0	0.0000
c...6.....	0.0	151	0	1	0	0.0000
c...C.....	-0.39181	152	8	4	7	0.0047
c...F.....	0.61466	153	1	2	0	1.0000
c...Br.....	2.58774	154	1	0	0	1.0000
c...Cl.....	0.0	155	0	1	1	0.0000
c...N.....	-1.22965	156	16	11	7	0.0003
c...O.....	1.71069	157	18	19	15	0.0044
c...S.....	-1.21645	158	3	1	0	1.0000
c...c.....	-0.12906	159	80	76	38	0.0008
n...(.....	0.44617	160	12	9	3	0.0026
n.....	-0.57438	161	12	10	3	0.0026
n...1.....	0.64873	162	8	4	1	0.0050
n...2.....	0.0	163	0	2	0	0.0000
n...3.....	0.0	164	0	1	0	0.0000
n...c.....	0.22185	165	10	5	1	0.0055
o...(.....	2.09447	166	1	0	0	1.0000
o.....	-0.79365	167	1	0	0	1.0000
o...2.....	0.57608	168	1	0	0	1.0000
s...(.....	0.0	169	0	1	0	0.0000
s.....	0.0	170	0	1	0	0.0000

Comment for Table S13, S14, and S15:

The N1, N2, N3 are frequencies of SMILES attributes (SA) in training, invisible training, and calibration sets, respectively.

DEFECT[SAk] is defect of SAk calculated as the following:

$$DEFECT[SAk] = \frac{|P(SAk) - P'(SAk)|}{N(SAk) + N'(SAk)}$$

where P and P’ are probability of SAk in training and calibration sets, respectively;
N an N’ are frequencies of SAk in training and calibration sets, respectively.

Comment for Table S1, S2, and S3

The defectSMILES is calculated as the following

$$defectSMILES = \sum DEFECT[SAk]$$