Electronic Supplementary Material (ESI) for Green Chemistry.

Supporting Information

Modelling biodegradability based on OECD 301D data for the design of mineralising ionic liquids

Ann-Kathrin Amsel[a,b], Suman Chakravarti[c], Oliver Olsson[a], Klaus Kümmerer[a,b,]*

[a]Institute of Sustainable Chemistry, Leuphana University of Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany

[b]Research and Education Hub, International Sustainable Chemistry Collaborative Centre (ISC3), Leuphana University of Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany

[c]MultiCASE Inc., 5885 Landerbrook Dr. #210, Mayfield Heights, OH 44124 USA, United States

* e-mail corresponding author: klaus.kuemmerer@leuphana.de

1. Model building

The ILs are multicomponent and charged compounds. This prevents calculation of conventional descriptors that are used traditionally in building (Q)SAR models, *e.g.* LogP, water solubility. Formal charges also exclude availability of required parameters for calculating descriptors which are otherwise available for uncharged compounds. Hence, only fragment descriptors were used. For fragmentation, widely accepted extended connectivity fingerprint (ECFP) type circular fragments up to a bond depth of 5 from the center atom were used.[1] Fragments were generated from both the anion and cation parts of each chemical. A canonical representation of the fragments was used, which was applied in Chakravarti *et al.*[2] The canonical representation ensures unique representation of a particular substructure invariant from the atom ordering. A SMILES like syntax for easy interpretation was maintained. It contains more information on atomic level to prevent loss of structural environment context which happens with SMILES representation of fragments.[3] Atoms are annotated with hybridisation, number of hydrogens, formal charge, etc.[3] For example C2 represents sp2 hybridised non-ring carbon and C3H2 sp3 hybridised carbon with two hydrogens in this fragment [C2]-[C3H2]-[C3H]. The description of the atom types is available in Table S1.

Two types of fragment descriptors were used: i. fragments based on ECFP type circular fragments as described above and ii. element of fingerprint. A special continuous valued fingerprint developed by Chakravarti[3] was used on the individual components of an IL separately and combined to give a final fingerprint of 600 elements. The fingerprints are distributed representations of fragments. One fragment can potentially affect all the 600 elements of the fingerprint. Every element of the fingerprint is a decimal point number and can be positive as well as negative.[3]

The variable selection picked those fragments that were relevant to the biodegradability potential of the training chemicals. The L1 regularisation/Lasso regression via the "R" package *glmnet* was used for this purpose.[4,5] The specific criteria to assess the relevance of fragments to the biodegradability potential was mean squared error of the observed and predicted biodegradability. Variable selection was performed separately for binary classification (logistic) and continuous regression models. The selected fragments were considered to be the privileged substructures, *i.e.* relevant to biodegradability potential of ILs. In addition to the variable selection, L1 regularisation/Lasso regression produces regression coefficients for the selected fragments. Positive and negative regression coefficients are indicative of fragments that favour and decrease biodegradability, respectively.

1.1 Canonical representation of substructures

Table S1: Description of atom types.

FRAGMENT	DESCRIPTION
C1	sp hybridised non-ring carbon
C2	sp2 hybridised non-ring carbon
C2H	sp2 hybridised non-ring carbon with one hydrogen
C3	sp3 hybridised non-ring carbon
C3H	sp3 hybridised non-ring carbon with one hydrogen
C3H2	sp3 hybridised carbon with two hydrogens
C3H3	sp3 hybridised non-ring carbon with three hydrogens
O2	sp2 hybridised oxygen
O2-	sp2 hybridised oxygen negatively charged
O3	sp3 hybridised oxygen
O3H	sp3 hybridised oxygen with one hydrogen
N2-	sp2 hybridised non-ring nitrogen negatively charged
N2_1	sp2 hybridised nitrogen with one double bond
N3	sp3 hybridised nitrogen
N3+	sp3 hybridised nitrogen positively charged
N3H	sp3 hybridised nitrogen with one hydrogen
N3H2	sp3 hybridised nitrogen with two hydrogens
N3H2+	sp3 hybridised nitrogen with two hydrogens, positively charged
N3H3+	sp3 hybridised nitrogen with three hydrogens, positively charged
F3	sp3 hybridised fluorine
S3	sp3 hybridised sulphur
S3H	sp3 hybridised sulphur with one hydrogen
S3_2	sp3 hybridised sulphur with two double bonds
c	aromatic carbon
c.	aromatic carbon on a ring joint
cH	aromatic carbon with one hydrogen
n	aromatic nitrogen
n+	aromatic nitrogen positively charged
nH	aromatic nitrogen with one hydrogen
Cl3-	sp3 hybridised chlorine negatively charged

1.2 Descriptors used for the model IL_FP_cont

Table S2: Regression coefficients for the 61 elements of the fingerprint used for the model IL_FP_cont. Regression coefficient indicates contribution towards biodegradation (higher = more biodegradable).

VARIABLE_NAME	REGRESSION_ COEFFICIENT	VARIABLE_NAME	REGRESSION_ COEFFICIENT
intercept	42.4631	BIN_409	-0.7423
BIN_22	-0.9802	BIN_410	0.4150
BIN_76	-5.1607	BIN_421	0.1057
BIN_80	-6.8303	BIN_423	-0.1657
BIN_82	2.8263	BIN_425	0.0801
BIN_153	-2.5154	BIN_429	-0.0437
BIN_158	-0.4383	BIN_452	-0.1209
BIN_159	-0.6025	BIN_456	1.9117
BIN_161	4.0727	BIN_458	-0.4272
BIN_208	-1.4189	BIN_469	-0.2436
BIN_236	-5.0643	BIN_480	0.7530
BIN_253	0.0578	BIN_484	-0.3183
BIN_272	-2.2881	BIN_503	2.6264
BIN_284	-1.5153	BIN_512	-1.9509
BIN_293	1.7894	BIN_521	0.1557
BIN_305	1.0973	BIN_530	-0.5392
BIN_311	0.3026	BIN_531	-2.1854
BIN_331	1.9854	BIN_533	-0.0477
BIN_349	-9.3207	BIN_558	1.1936
BIN_351	-0.4624	BIN_562	-2.1410
BIN_355	0.7128	BIN_569	-0.1680
BIN_356	1.5193	BIN_570	-0.9162
BIN_358	-0.9837	BIN_572	-0.2513
BIN_363	0.8734	BIN_573	-2.0735
BIN_364	-0.4064	BIN_576	-3.7916
BIN_377	-0.3324	BIN_577	-1.1347
BIN_379	0.0826	BIN_582	0.7411
BIN_387	3.9337	BIN_583	-6.4103
BIN_394	-0.6601	BIN_594	0.0412
BIN_399	-1.8238	BIN_595	-0.1719
BIN_403	-2.3152	BIN_599	-1.3475

1.3 Fragments used for the models IL_Al_cont, IL_Al_class, ILNI_Al_cont and ILNI_Al_class

Table S3: Regression constant and alerts for the model IL_Al_cont and the related regression coefficients. Regression coefficient indicates contribution towards biodegradation (higher = more biodegradable). Canonical representation of the fragments was chosen to show hybridization, number of hydrogens, formal charge.

ID	ALERT	REGRESSION_ COEFFICIENT
	CONSTANT	24.0910
1	[O2]	2.1627
2	[c](:[cH])(:[cH])-[C2]	8.5377
3	[O2-]	6.5249
4	[N3H2]	5.5314
5	[S3H]	-1.9374
6	[C2]-[O2-]	0.2842
7	[C3H2]-[C3H2]-[n]	-2.1521
8	[cH](:[n]):[n+]	-3.5515
9	[C2](=[O2])(-[C3H])-[O2-]	0.2061
10	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	2.9806
11	[C3H3]-[n+]1:[cH]:[cH]:[n]:[cH]1	-8.1990
12	[C3H2]-[C3H2]-[O3]	3.0842
13	[C1]-[C3H2]-[n]	1.3233
14	[N2-](-[S3])-[S3]	10.3149
15	[C1]-[C3H2]-[n]1:[cH]:[cH]:[n+](:[cH]1)-[C3H2]	2.4854
16	[C3](-[F3])(-[F3])(-[F3])-[S3](=[O2])(=[O2])-[N2-]-[S3]	3.2426
17	[Cl3-]	9.0035
18	[c](:[cH])(:[cH])-[C3H3]	2.2331
19	[C3H2](-[C3H2]-[C3H3])-[C3H2]-[n+]	-9.4859
20	[C3H2]-[C3H2]-[C3H2]-[n+](:[cH]):[cH]	-0.5450
21	[c]:[cH]:[cH]:[n+](:[cH])-[C3H2]	-9.0554
22	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H3]	0.2878
23	[C3H2]-[C3H2]-[n+](:[cH]:[cH]):[cH]:[cH]	-5.2058
24	[C3H2]-[C3H2]-[N3+]	-0.2294
25	[C2](=[O2])(-[C3H2]-[N3+])-[N3H]-[C3H]	-3.6574
26	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H3])-[N3H]-[C2](=[O2])-[C3H2]-[N3+]	-1.6884
27	[C2]-[C3H2]-[C3H]	4.4449
28	[C3H]-[C3H2]-[C3H2]	0.6055
29	[C3H2]-[N3+](-[C3H3])(-[C3H3])-[C3H3]	33.1433
30	[C2](=[O2])(-[C3H2]-[n+]1:[cH]:[cH]:[n](:[cH]1)-[C3H3])-[N3H]	3.6090
31	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3])-[N3H]-[C2](=[O2])-[C3H2]-[n+]	5.6214
32	[C2]-[C3H](-[C3H2])-[N3H]-[C2](=[O2])-[C3H2]-[n+]1:[cH]:[cH]:[n](:[cH]1)-[C3H3]	0.1124
33	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	1.0807
34	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	4.9162

35	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-19.5037
36	[C2]-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-4.9086
37	[C3H]-[O3H]	1.0669
38	[C3H](-[C3H2])(-[C3H2])-[O3H]	2.6522
39	[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H2]	-14.6970
40	[C3H2]-[C3H2]-[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H2]	-4.5986
41	[C3H2]-[C3H2]-[N3+](-[C3H2]-[C3H2])(-[C3H2]-[C3H2])-[C3H2]-[C3H2]-[C3H2]-[C3H3]	-0.3843
42	[C3H]1-[C3H2]-[C3H2]-[C3H2]-[N3H]1	7.3152
43	[c](:[cH])(:[cH])-[O3]	-9.1273
44	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H3])-[N3H]-[C2](=[O2])-[C3H2]-[n+]	5.7458
45	[N2]-[O3H]	-3.1732
46	[C2]-[N3H]-[C3H2]	-10.9595
47	[C2H]=[N2]-[O3H]	-0.1178
48	[C2](=[O2])(-[C3H])-[N3H]-[C3H2]-[C3H2]	-2.1971
49	[C3H3]-[N3]	-1.1983
50	[C3H2](-[C3H3])-[n]	-10.2854
51	[C2]-[C3H2]-[N3]	32.0454
52	[C2]-[C3H2]-[C3H2]	2.0551
53	[C2](=[O2])(-[C3H2]-[N3])-[O2-]	0.3289
54	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2])-[N3]	1.1893
55	[C2]-[C3H2]-[N3](-[C2](=[O2])-[C3H2])-[C3H3]	0.8998
56	[C2]-[C3H](-[C3H2])-[N3+]	-2.0347
57	[C3H]1-[C3H2]-[C3H2]-[C3H2]-[N3+]1	-0.1403
58	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[C3H2]	3.2256
59	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[C3H2]	0.1154
60	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[C3H2]-[n+]	0.4979
61	[C2](=[O2])(-[C3H2]-[n])-[O3]-[C3H2]	1.3243
62	[C2](=[O2])(-[C3H2]-[n])-[O3]-[C3H2]-[C3H2]-[C3H2]	0.9451
63	[C2]-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	4.1352
64	[c]:[n+](:[cH])-[C3H3]	-4.7922
65	[c]-[C3H](-[C2])-[n+]	-7.2859
66	[c](:[cH])(:[cH])-[C3H]	-0.3177
67	[C3H]-[n+](:[cH]):[cH]	-0.1892
68	[c]1(:[cH]:[cH]:[cH]:[cH]:[cH]1)-[C3H]	-0.2379
69	[Al-]	-13.4799
70	[C3]-[C3H](-[C3])-[O3]	-0.5482

Table S4: Regression constant and alerts for the model IL_Al_class and the related regression coefficients. Regression coefficient indicates contribution towards biodegradation (higher = more biodegradable). Canonical representation of the fragments was chosen to show hybridization, number of hydrogens, formal charge.

ID	ALERT	REGRESSION_ COEFFICIENT
	CONSTANT	-3.1179
1	[C3H2]-[C3H3]	-0.0525
2	[C2]-[O3H]	0.2724
3	[c](:[cH])(:[cH])-[C2]	0.2820
4	[n]	-0.0839
5	[C3H3]-[n+]1:[cH]:[cH]:[n]:[cH]1	-0.3010
6	[C3H2]-[n]1:[cH]:[cH]:[n+](:[cH]1)-[C3H3]	-0.0223
7	[C3H2]-[C3H2]-[n+]	0.1855
8	[C1]-[C3H2]-[n]	1.0130
9	[N2-](-[S3_2])-[S3_2]	2.6090
10	[C3H2](-[C3H2]-[O3])-[n+](:[cH]):[cH]	0.0846
11	[C3](-[F3])(-[F3])(-[F3])-[S3_2](=[O2])(=[O2])-[N2-]	0.0483
12	[C1](#[N1])-[C3H2]-[n]1:[cH]:[cH]:[n+](:[cH]1)-[C3H2]	0.0157
13	[C3]-[S3_2](=[O2])(=[O2])-[N2-]-[S3_2](=[O2])(=[O2])-[C3](-[F3])(-[F3])-[F3]	0.0142
14	[C3](-[F3])(-[F3])(-[F3])-[S3_2](=[O2])(=[O2])-[N2-]-[S3_2](=[O2])(=[O2])-[C3](-[F3])(-[F3])-[F3]	0.0014
15	[C3H2]-[N3+](-[C3H3])(-[C3H3])-[C3H3]	5.7231
16	[C3H2]1-[C3H2]-[C3H2]-[N3+]-[C3H2]1	0.4060
17	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	0.4129
18	[N3H2+]	3.1941
19	[C2](=[O2])(-[C3H3])-[O2-]	0.0551
20	[C2]-[C3H2]-[N3]	3.3436
21	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2])-[N3]	0.0208
22	[C2]-[N3](-[C3H2]-[C2](=[O2])-[O2-])-[C3H3]	0.1072
23	[C2](=[O2])(-[C3H2]-[C3H2])-[N3](-[C3H2])-[C3H3]	0.0065
24	[C2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	0.0599
25	[C2](=[O2])(-[C3H2]-[C3H2])-[N3](-[C3H2]-[C2](=[O2])-[O2-])-[C3H3]	0.0037
26	[C3H2]-[N3+](-[C3H2])(-[C3H2]-[C3H2]-[C3H2]-[C3H3])-[C3H3]	0.3988
27	[C3H2]-[n+]1:[cH]:[cH]:[cH]:[cH]:[cH]1	0.1337
28	[C2](=[O2])(-[C3H2]-[n+])-[O3H]	0.6031
29	[c.]:[cH]:[cH]	-0.9396

Table S5: Regression constant and alerts for the model ILNI_Al_cont and the related regression coefficients. Regression coefficient indicates contribution towards biodegradation (higher = more biodegradable). Canonical representation of the fragments was chosen to show hybridization, number of hydrogens, formal charge.

ID	ALERT	REGRESSION_ COEFFICIENT
	REGRESSION_CONSTANT	48.7302
1	[C3H2]-[C3H2]-[N3H]-[C3H2]-[C3H2]	37.1426
2	[C3H2]-[N3+](-[C3H3])(-[C3H3])-[C3H3]	34.5821
3	[c](:[cH])(:[cH]:[cH]:[cH])-[C2]	31.8354
4	[C3H3]-[n+](:[cH]):[cH]:[cH]:[cH]	22.6716
5	[C2]-[C3H2]-[N3]	16.4266
6	[C3H2]-[C3H2]-[C3H2]-[O3]-[S3_2]	15.0042
7	[N3H3+]	14.7954
8	[c](:[cH])(:[cH]:[cH]:[cH])-[O3H]	14.5016
9	[C3H]-[C3H2]-[N3H]	13.7104
10	[C2]-[C3H](-[C3H2])-[N3H3+]	13.5572
11	[C3H]1-[C3H2]-[C3H2]-[C3H2]-[N3H]1	12.6291
12	[C2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	11.8645
13	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2])-[N3H]-[C2]	11.6728
14	[C3H2]-[C3H2]-[N3H2]	10.0375
15	[C2]-[C3H2]-[C3H]	9.9845
16	[C2]-[C3H2]-[C3H2]	9.9586
17	[C2](=[O2])(-[C3H2]-[n+]1:[cH]:[cH]:[n](:[cH]1)-[C3H3])-[N3H]	9.4466
18	[C2]-[C3H3]	9.0880
19	[C3](-[F3])(-[F3])(-[F3])-[S3_2]	8.9476
20	[C3H2](-[C3H2]-[C3H3])-[C3H2]-[O3]	8.2691
21	[C2]-[C3H]1-[C3H2]-[C3H2]-[C3H2]-[N3H]1	7.7071
22	[C2](=[O2])(-[C3H2]-[n+])-[N3H]-[C3H]	7.6261
23	[C2]-[O3H]	6.9449
24	[nH]	6.2442
25	[C2]-[N3H]-[C3H2]-[C3H2]-[C3H2]	6.1970
26	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H3])-[N3H]-[C2](=[O2])-[C3H2]-[n+]	6.0009
27	[C3H2]-[C3H2]-[O3]	5.7721
28	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3])-[N3H]-[C2](=[O2])-[C3H2]-[N3+](-[C3H2]-[C3H2])(-[C3H3])-[C3H3]	5.2437
29	[C3H]-[C3H3]	5.1976
30	[C3H2](-[C3H2]-[O3H])-[N3]	5.1923
31	[C3H2]-[C3H2]-[C3H2]-[O3]-[C3H2]	4.9450
32	[C3H]-[N3H2]	4.7824
33	[C2](=[O2])(-[C3H])-[O3H]	4.4935
34	[C2]-[O2-]	3.2180
35	[c]-[O3H]	3.0171
36	[C2]-[C3H](-[C3H2])-[N3H]	2.7625
37	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[S3_2]	2.7288
38	[c]-[C3H](-[C2](=[O2])-[O3]-[C3H2])-[n+]	2.3036
39	[c](:[cH])(:[cH])-[C3H](-[C2](=[O2])-[O3])-[n+](:[cH]):[cH]	2.2482

40	[C2](=[O2])(-[C3H2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	1.9281
41	[C2](=[O2])(-[C3H2])-[O2-]	1.6421
42	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[S3_2](=[O2])(=[O2])-[O2-]	1.6211
43	[O2-]	1.5824
44	[c]-[C3H2]-[C3H]	1.4427
45	[C2]-[C3H](-[C3H])-[N3H2]	1.0532
46	[C3H](-[C3H2])(-[C3H3])-[C3H3]	0.9822
47	[O3H]	0.9663
48	[C2]-[O3]-[C3H2]	0.9394
49	[C2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	0.7255
50	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3])-[N3H3+]	0.6869
51	[O2]	0.5942
52	[C2]-[C3H2]-[n+]1:[cH]:[cH]:[cH]:[cH]:[cH]1	0.5225
53	[C3H](-[C3H2])(-[C3H2])-[O3H]	0.4354
54	[c](:[cH])(:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2])-[N3H3+]	0.4041
55	[C2](=[O2])(-[C3H2]-[n+]1:[cH]:[cH]:[n](:[cH]1)-[C3H3])-[N3H]-[C3H]	0.4007
56	[C2](=[O2])(-[C3H2]-[N3])-[O2-]	0.3278
57	[C2H]	0.2250
58	[C3H3]-[N3]	0.1723
59	[C2]-[N3](-[C3H2])-[C3H3]	0.0970
60	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3](-[C3H2])-[C3H3]	0.0933
61	[C2](=[O2])(-[C3H2])-[N3]	0.0878
62	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]-[S3_2](=[O2])(=[O2])-[O2-]	0.0785
63	[C2](=[O2])(-[C3H2]-[C3H2])-[N3](-[C3H2]-[C2](=[O2])-[O2-])-[C3H3]	0.0749
64	[c](:[cH])(:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2])-[N3H]-[C2](=[O2])-[C3H2]-[n+]1:[cH]:[cH]:[n](:[cH]1)-[C3H3]	0.0287
65	[C2](=[O2])(-[C3H])-[O3]-[C3H2]-[C3H2]	0.0168
66	[c](:[cH])(:[cH])-[C3H2]-[C3H](-[C2])-[N3H3+]	0.0146
67	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2])-[N3]	0.0045
68	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3]	0.0025
69	[c]1(:[cH]:[cH]:[cH]:[cH]:[cH]1)-[C3H](-[C2](=[O2])-[O3])-[n+](:[cH]):[cH]	0.0015
70	[cH]:[cH]:[cH]:[cH]:[cH]	0.0006
71	[C3H2]-[C3H2]-[N3+](-[C3H2]-[C3H2])(-[C3H2]-[C3H2])-[C3H2]-[C3H2]-[C3H2]-[C3H3]	-0.0003
72	[C2](=[C2H])(-[C3H3])-[O3]	-0.0010
73	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3H2]	-0.0021
74	[c]1:[cH]:[cH]:[cH]:[n+](:[cH]1)-[C3H2]-[C2]	-0.0096
75	[S3H]	-0.0436
76	[F3]	-0.0848
77	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3H2]	-0.0859
78	[c](:[cH])(:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3H2]	-0.1423
79	[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H3]	-0.2261
80	[C2]-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-0.3859
81	[C3H2]-[C3H2]-[C3H2]-[N3+](-[C3H2]-[C3H2]-[C3H2])(-[C3H2]-[C3H2]-[C3H2])-[C3H2]-[C3H2]-[C3H2]	-0.4731

82	[C3H2]-[C3H2]-[C3H2]-[N3+](-[C3H2]-[C3H2]-[C3H2])(-[C3H2]-[C3H2]-[C3H2])-[C3H2]-[C3H2]-[C3H2]-[C3H3]	-0.8327
83	[c]:[cH]:[n+](:[cH]:[cH])-[C3H2]-[C2]	-0.9341
84	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3]	-0.9760
85	[c]1:[c](:[cH]:[cH]:[cH]:[cH]1)-[C2]	-1.0763
86	[C3H]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-1.1152
87	[C2](=[O2])(-[C3H2]-[n](:[cH]):[cH])-[O3]-[C3H2]-[C3H2]-[C3H3]	-1.1682
88	[C3H]-[C3H2]-[C3H2]	-1.5838
89	[C3H2](-[C3H3])-[N3+]	-1.7078
90	[C3H2]-[n](:[cH]):[cH]	-1.7293
91	[C3H3]-[n+]	-1.7560
92	[C3]-[F3]	-2.1666
93	[C3H2]-[C3H2]-[N3+]	-2.6015
94	[I3-]	-2.7713
95	[C3H2](-[C3H2]-[C3H2]-[C3H2]-[C3H3])-[C3H2]-[C3H2]-[C3H2]-[O3]	-2.8595
96	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H3])-[N3H]-[C2](=[O2])-[C3H2]-[N3+]	-3.2251
97	[C2]-[N2-]-[S3_2](=[O2])(=[O2])-[O3]-[C2]	-3.5590
98	[c]:[c](:[cH]:[cH]:[cH])-[S3_2]	-3.8670
99	[Br3-]	-4.3447
100	[C3H2]-[N3H2]	-4.7930
101	[Cl3-]	-4.9373
102	[c]:[c](:[cH])-[S3_2]	-5.0178
103	[C2](=[O2])(-[C3H](-[C3H2])-[N3H2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-5.1341
104	[c]:[cH]:[cH]:[cH]:[n+]	-5.3127
105	[C3H2]-[C3H2]-[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H2]	-5.4589
106	[C3H2](-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3])-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]	-6.2131
107	[N2_1]	-6.3064
108	[C2]-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-6.7569
109	[C3H2]-[C3H2]-[O3H]	-7.1681
110	[C3H2]-[C3H2]-[n]	-7.3933
111	[N2-]-[S3_2](=[O2])(=[O2])-[O3]	-7.8923
112	[n]	-8.7230
113	[C2]-[O3]-[S3_2]	-9.7028
114	[N3H]	-10.2354
115	[c]:[cH]:[cH]:[n+](:[cH])-[C3H2]	-10.6381
116	[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-11.1189
117	[C2]-[C2H]=[C2](-[C3H3])-[O3]-[S3_2]	-11.2764
118	[c]:[c](:[cH]:[cH]:[cH])-[C2]	-12.5455
119	[C3H2](-[C3H3])-[n]	-13.5779
120	[C3H2](-[C3H2]-[C3H3])-[C3H2]-[n+]	-13.7240
121	[c](:[cH])(:[cH])-[O3]	-13.7429
122	[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H2]	-14.9098
123	[C3H2]-[C3H2]-[N3]	-15.6074
124	[N3+]	-16.1439

125	[n+]	-17.1591
126	[c](:[cH])(:[cH])-[C3H]	-17.8286
127	[C2](=[O2])(-[C3H](-[C3H2])-[N3H2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-18.6859
128	[c](:[cH])(:[cH])-[C2]	-20.6146
129	[C3H2]-[N3H]-[C3H2]	-22.5624
130	[N3H2+]	-25.6157

Table S6: Regression constant and alerts for the model ILNI_Al_class and the related regression coefficients. Regression coefficient indicates contribution towards biodegradation (higher = more biodegradable). Canonical representation of the fragments was chosen to show hybridization, number of hydrogens, formal charge.

ID	ALERT	REGRESSION_ COEFFICIENT
	CONSTANT	-0.3018
1	[C3H2]-[N3+](-[C3H3])(-[C3H3])-[C3H3]	3.0851
2	[C2]-[C3H2]-[C3H2]	1.4679
3	[C2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	1.3265
4	[N3H3+]	1.3122
5	[C3H3]-[n+](:[cH]):[cH]:[cH]:[cH]	1.2936
6	[C3](-[F3])(-[F3])(-[F3])-[S3_2]	0.8861
7	[C2]-[O3H]	0.7366
8	[C2]-[C3H3]	0.7162
9	[C3H2]-[C3H2]-[O3]-[C3H2]-[C3H2]	0.7126
10	[C2]-[C3H](-[C3H])-[N3H2]	0.6696
11	[C2](=[O2])(-[C3H2])-[N3]	0.6201
12	[C2]-[C3H](-[C3H2])-[N3H3+]	0.5443
13	[c](:[cH])(:[cH]:[cH]:[cH])-[C2]	0.4417
14	[C3H2]-[C3H2]-[N3H]-[C3H2]-[C3H2]	0.2922
15	[C3H](-[C3H2])(-[C3H3])-[C3H3]	0.2303
16	[C3H2](-[C3H2]-[C3H3])-[C3H2]-[O3]	0.2293
17	[C2]-[C3H2]-[C3H]	0.1792
18	[C3H2]-[C3H2]-[N3H2]	0.0949
19	[nH]	0.0687
20	[C2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	0.0585
21	[C2]-[N3](-[C3H2])-[C3H3]	0.0430
22	[C2](=[O2])(-[C3H2]-[C3H2])-[N3](-[C3H2]-[C2](=[O2])-[O2-])-[C3H3]	0.0326
23	[C2](=[O2])(-[C3H2]-[N3])-[O2-]	0.0102
24	[C2](=[O2])(-[C3H](-[C3H2])-[N3H3+])-[O3]-[C3H2]	0.0016
25	[C2](=[O2])(-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3]	0.0009
26	[c]-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3H2]	-0.0002
27	[C2]-[C3H](-[C3H2]-[S3H])-[N3H2]	-0.0002
28	[C2]-[C2H]=[C2](-[C3H3])-[O3]-[S3_2]	-0.0003
29	[N2-]-[S3_2](=[O2])(=[O2])-[O3]	-0.0013
30	[c]	-0.0110
31	[C2](=[C2H])(-[C3H3])-[O3]	-0.0232
32	[c](:[cH])(:[cH])-[O3]	-0.0295

33	[c]:[cH]:[cH]	-0.0323
34	[C3H2]-[C3H2]-[N3+]	-0.0376
35	[C3H2](-[C3H3])-[n]	-0.0547
36	[C3H2]-[N3+](-[C3H2])(-[C3H3])-[C3H3]	-0.0610
37	[c](:[cH]:[cH])(:[cH]:[cH])-[C3H2]-[C3H](-[C2](=[O2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2])-[N3H2]	-0.0627
38	[C2]-[O3]-[C3H2]-[C3H3]	-0.0714
39	[O3]	-0.0767
40	[C3H2]	-0.1017
41	[C3H2]-[C3H3]	-0.1561
42	[N3H]	-0.1926
43	[C3H2]-[C3H2]-[N3H2+]	-0.2359
44	[C3H3]-[O3]	-0.2799
45	[Br3-]	-0.3028
46	[c]:[cH]:[cH]:[cH]:[n+]	-0.3031
47	[C2]-[O3]-[S3_2]	-0.3207
48	[C3H2]-[C3H2]-[N3]	-0.3741
49	[c](:[cH])(:[cH])-[C3H]	-0.4450
50	[C2](=[O2])(-[C3H](-[C3H2])-[N3H])-[O3]-[C3H2]-[C3H3]	-0.5287
51	[N3H2+]	-0.6133
52	[C3H3]-[n+]1:[cH]:[cH]:[n]:[cH]1	-0.6323
53	[C2]-[N2-]-[S3_2]	-0.6451
54	[S3H]	-0.6567
55	[N3+]	-0.6766
56	[n]	-0.8423
57	[C3H2](-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H3])-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[O3]	-0.9708
58	[C2](=[O2])(-[C3H](-[C3H2])-[N3H2])-[O3]-[C3H2]-[C3H2]-[C3H2]-[C3H2]-[C3H2]	-1.0219
59	[C3H2]-[N3+](-[C3H2])(-[C3H2])-[C3H2]	-1.0421
60	[n+]	-1.2691

2. Model validation

The validation of QSBRs was divided in internal and external validation as proposed by OECD to assess the goodness-of-fit and the predictivity.[6] The goodness-of-fit indicates the model's ability to reproduce the experimental data in the training set.[7] The predictivity is the ability to predict reliably biodegradability for chemicals in the test set.[7] To assess the goodness-of-fit in the internal validation *train_set_IL* was used for the models IL_FP_cont, IL_Al_cont, IL_Al_class and *train_set_ILNI* for the models ILNI_Al_cont and ILNI_Al_class. In the external

validation the model's predictivity was evaluated using *test_set_IL* for the models IL_FP_cont, IL_Al_cont, IL_Al_class and *test_set_ILNI* for the models ILNI_Al_cont and ILNI_Al_class.

Typical performance measures for classification models are the accuracy (equation 1, Table S7), the sensitivity (true positive rate (TPR), equation 2, Table S7), the specificity (true negative rate (TNR), equation 3, Table S7) and the area under the curve (AUC).[6,8] The accuracy is the proportion of correctly classified chemicals as positive (biodegradable) or negative (non-biodegradable) (equation 1). If a model has a higher value for the accuracy compared to another model, it is more reliable.[6] The sensitivity and specificity measure the correct classification of chemicals as positive and as negative, respectively (equations 2 and 3, Table S7). A high sensitivity corresponds to a high true positive rate indicating a high probability that the test chemical is correctly predicted as positive.[6] A high specificity or true negative rate is related to a high probability that a chemical is correctly predicted as negative.[6] The AUC is defined as the area under the receiver operating characteristic (ROC) curve and can have values from 0.0 to 1.0, while 1.0 indicates that the model's predictions are all correct and 0.5 that the model cannot discriminate between the two classes. The ROC and AUC show the relationship between sensitivity and specificity of a model.[6,8] TPR, TNR, the accuracy and AUC were calculated based on the training set and the test set to assess the goodness-of-fit and the predictivity, respectively, of the models IL_Al_class and ILNI_Al_class. Several performance measures were evaluated for classification models since accuracy, sensitivity, specificity and AUC help to understand the model's performance in predicting both classes, biodegradable and non-biodegradable ILs.

For OLS models a commonly used performance measure is the squared correlation coefficient R^2, which ranges from 0.0 to 1.0.[6,8] A R^2 of 1.0 means that the predictions correspond to the experimental values for given a property. R^2 was calculated based on the training set to assess

the goodness-of-fit and on the test set to assess the predictivity for IL_FP_cont, IL_Al_cont and ILNI_Al_cont (equation 4, Table S7).

Table S7: Performance measures and equations.

No.	Performance measure	equation	explanations
1	Accuracy	$\text{accuracy}=\frac{TP+TN}{TP+FP+TN+FN}\times 100$	TP, true positives TN, true negatives FP, false positives FN, false negatives TPR, true positive rate TNR, true negative rate
2	TPR	$TPR=\frac{TP}{TP+FN}\times 100$	
3	TNR	$TNR=\frac{TN}{TN+FP}\times 100$	
4	R^2	$R^2=1-\frac{\sum(y_{exp}-y_{pred})^2}{\sum(y_{exp}-\bar{y}_{exp})^2}$	y_{exp}, experimental value y_{pred}, predicted value $\bar{y}_{exp}$, mean experimental value

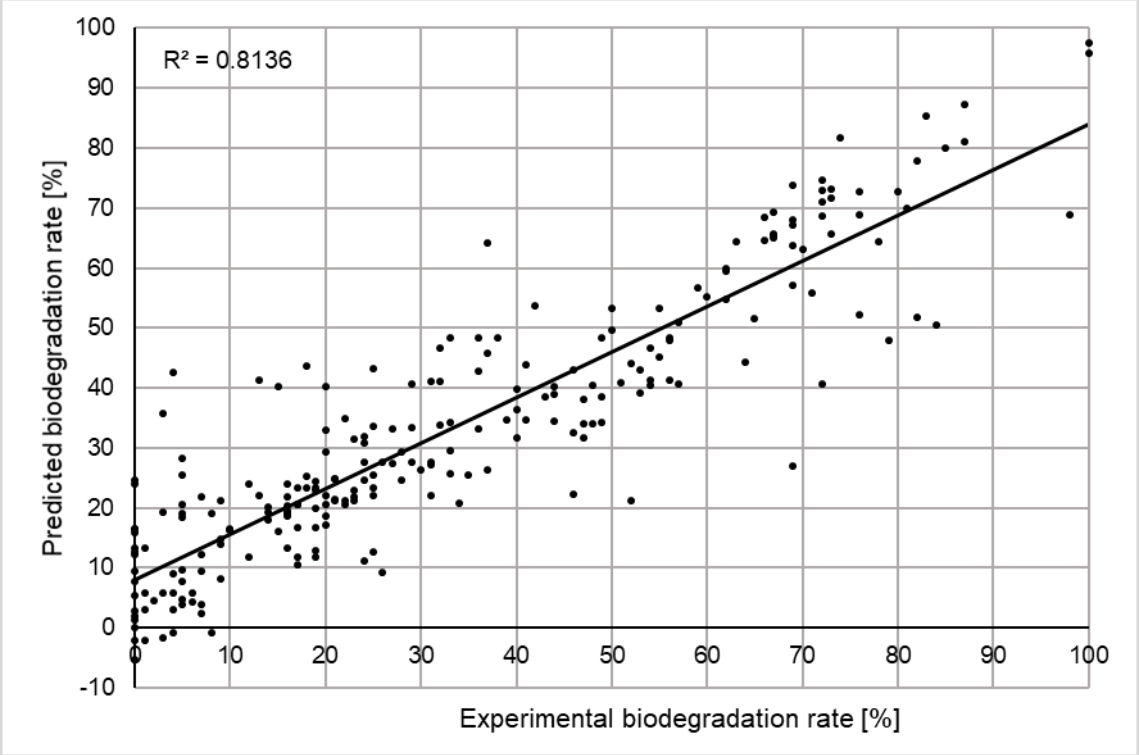

Figure S1: Predicted vs. experimental biodegradation rates for the IL_train_set of the model IL_FP_cont.

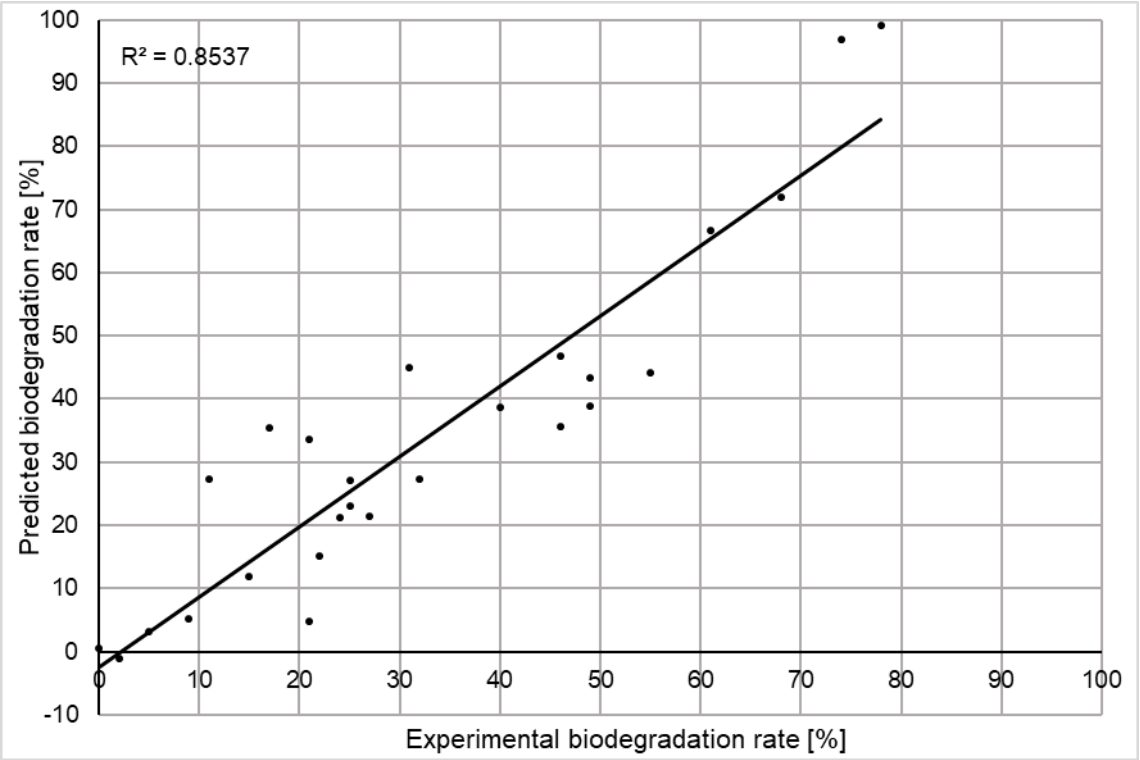

Figure S2: Predicted vs. experimental biodegradation rates for the IL_test_set of the model IL_FP_cont.

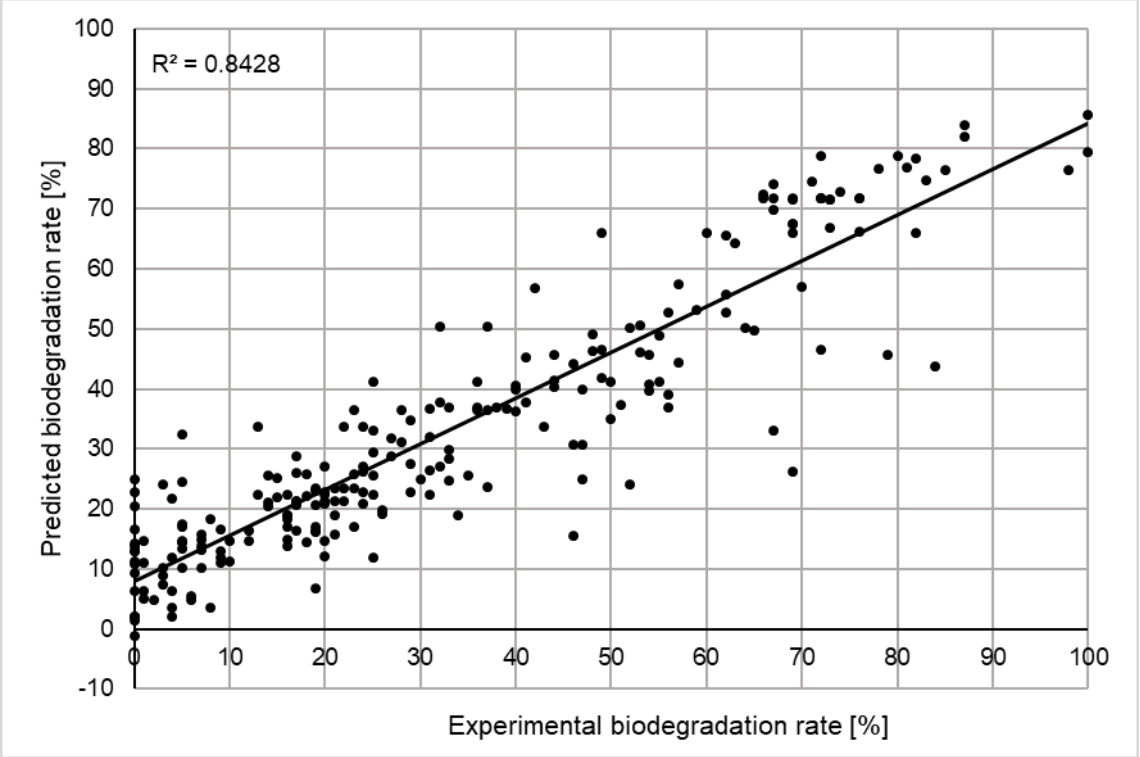

Figure S3: Predicted vs. experimental biodegradation rates for the IL_train_set of the model IL_Al_cont.

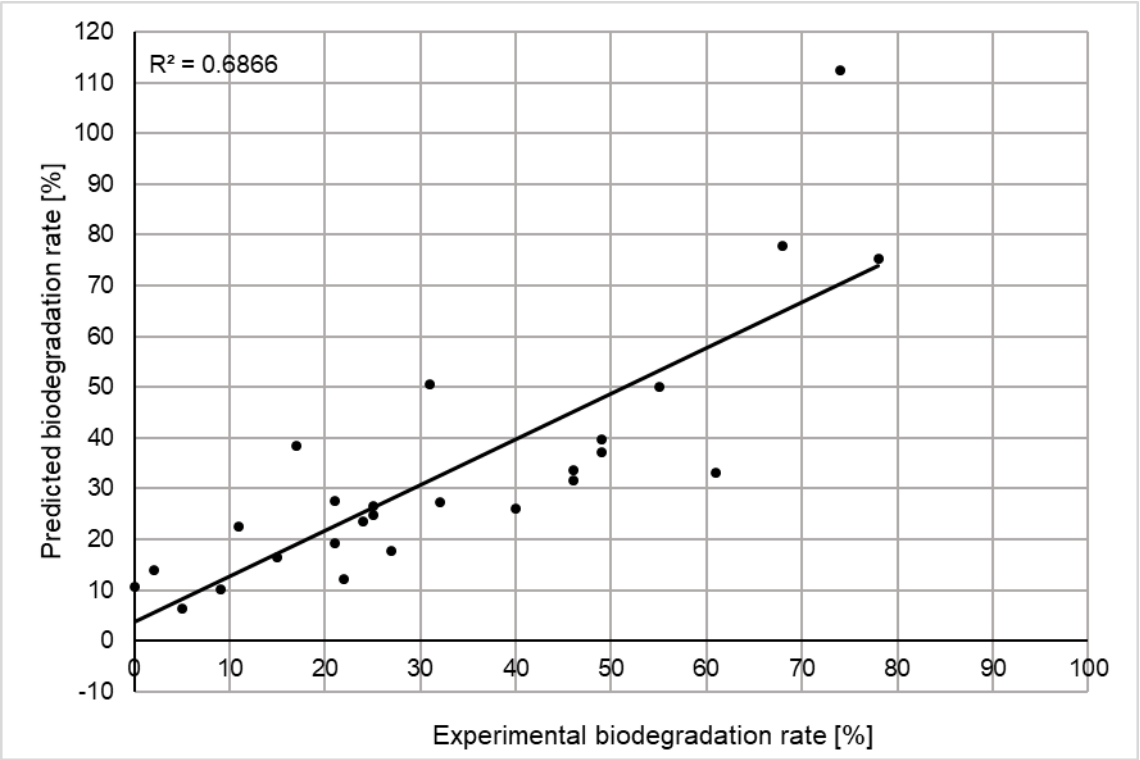

Figure S4: Predicted vs. experimental biodegradation rates for the IL_test_set of the model IL_Al_cont.

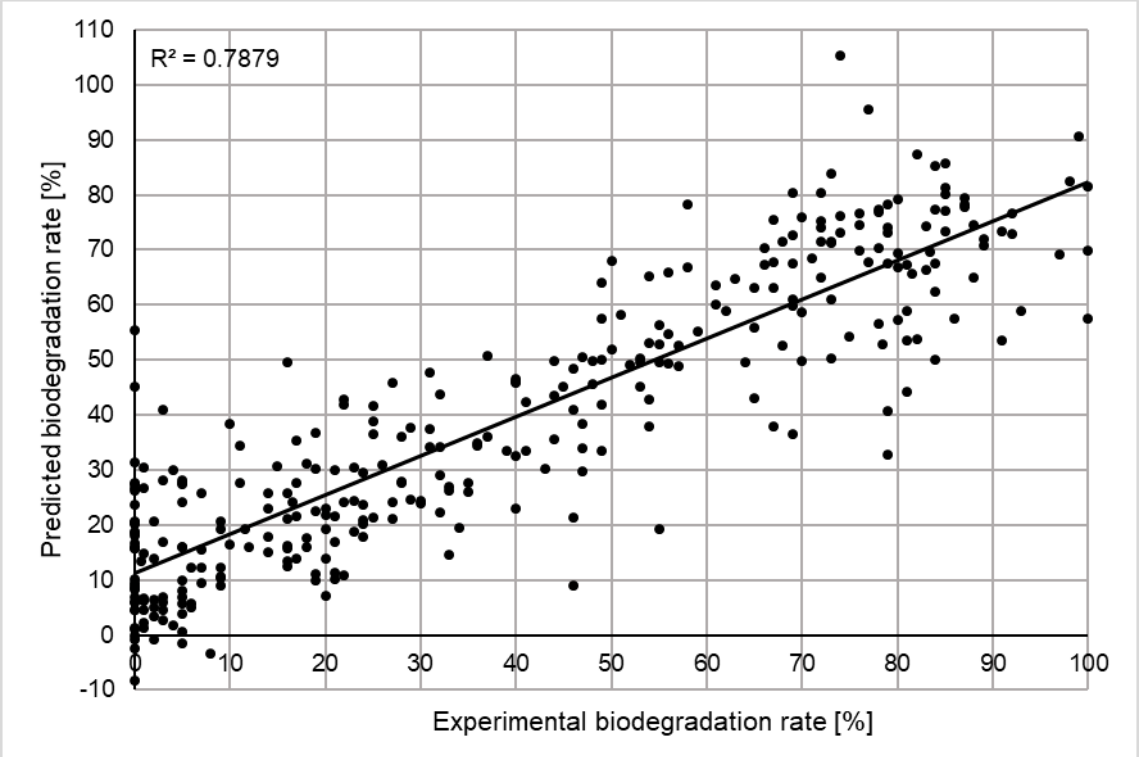

Figure S5: Predicted vs. experimental biodegradation rates for the IL_train_set of the model ILNI_Al_cont.

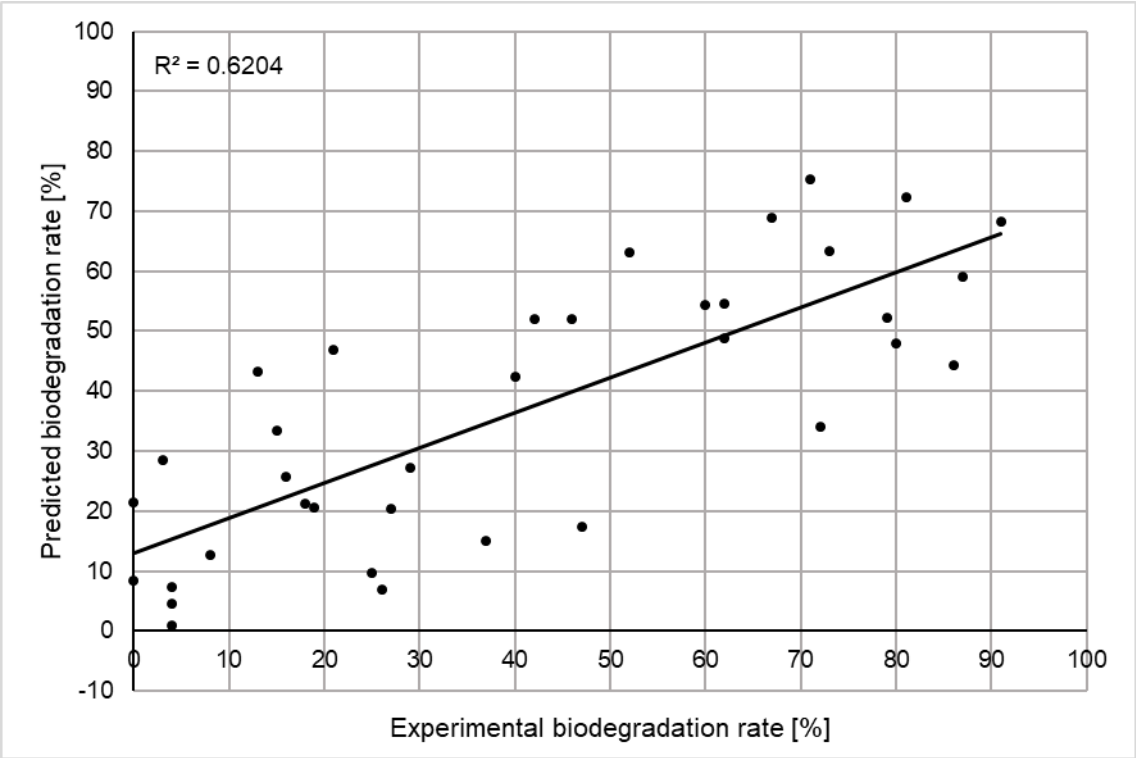

Figure S6: Predicted vs. experimental biodegradation rates for the IL_test_set of the model ILNI_Al_cont.

Table S8: Compliance of models with OECD principles for validating (Q)SAR models.

principle	**model 1** **IL_FP_cont**	**model 2** **IL_Al_cont**	**model 3** **IL_Al_class**	**model 4** **ILNI_Al_cont**	**model 5** **ILNI_Al_class**
1. a defined endpoint	✓	✓	✓	✓	✓
2. an unambiguous algorithm	✓	✓	✓	✓	✓
3. a defined domain of applicability	✓	✓	✓	✓	✓
4. appropriate measures of goodness-of–fit, robustness and predictivity	goodness-of–fit predictivity ✓	goodness-of–fit predictivity ✓	goodness-of–fit predictivity ✓	goodness-of–fit predictivity ✓	goodness-of–fit predictivity ✓
5. a mechanistic interpretation, if possible	✓	✓	✓	✓	✓

3. Clustering

Table S9: tSNE coordinates for tSNE plot train_set_IL.

Name	SMILES	TSNE_1	TSNE_2
Piperi_2_100	C1CC[N+]CC1.O=C([O-])CN(C)C(=O)CCCCCCCCCCC	3.70766973495483	1.07376456260681
Morph_3_100	C1COCC[N+]1.O=C([O-])CN(C)C(=O)CCCCCCCCCCC	3.70271754264831	1.09090626239776
QAC_43_98	C[N+](C)(C)CC(O)CC(O)=O.[Cl-]	-4.43112134933471	17.770824432373
Chol_13_87	C[N+](C)(C)CCO.O=C([O-])C(N)CCC(N)=O	-6.2032356262207	12.8914775848388
Chol_12_87	C[N+](C)(C)CCO.O=C([O-])C(N)CC(N)=O	-6.17524480819702	13.2271690368652
Chol_11_85	C[N+](C)(C)CCO.OC(=O)CCC(N)C([O-])=O	-5.27482938766479	13.03931999206540
Pyri_59_84	C[n+]1cccc(c1)C(O)=O.[Cl-]	13.7053985595703	-13.980167388916
Chol_30_83	C[N+](C)(C)CCO.[O-]C(=O)CCC(C)C1CC[C@H]2[C@@H]3CC[C@@H]4C[C@H](O)CC[C@]4(C)[C@H]3CCC12C	-18.0120487213134	4.72933959960937
Chol_10_82	C[N+](C)(C)CCO.OC(=O)CC(N)C([O-])=O	-5.31530332565307	13.4411344528198
Chol_1_82	C[N+](C)(C)CCO.[Cl-]	-5.16975688934326	18.4372158050537
QAC_44_81	C[N+](C)(C)CCOC(N)=O.[Cl-]	-5.0218768119812	19.5310478210449
Pyri_63_80	CCCCCCCC[n+]1ccc(C)c(C)c1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	0.650794327259063	-13.0990304946899
Pyrr_1_79	CCCC[N+]1(C)CCCC1.C(F)(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.3414325714111	2.7429928779602
Pyri_62_78	CCCCCC[n+]1ccccc1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	0.917508006095886	-12.7225151062011
Chol_18_76	C[N+](C)(C)CCO.NC(Cc1cnc2ccccc12)C([O-])=O	-3.88214564323425	11.4677810668945
Chol_7_76	C[N+](C)(C)CCO.OCC(N)C([O-])=O	-5.5636134147644	15.1276664733886
QAC_42_76	C[N+](C)(C)CC([O-])=O	-6.22535133361816	18.5007190704345
Chol_8_74	C[N+](C)(C)CCO.[O-]C(=O)C(N)C(C)O	-5.68957757949829	14.5134315490722
Chol_2_73	C[N+](C)(C)CCO.O=C([O-])CN	-5.39975976943969	16.3803367614746
Chol_16_73	C[N+](C)(C)CCO.O=C([O-])C(N)CCCNC(=N)N	-5.99325132369995	11.6799812316894
Chol_24_73	C[N+](C)(C)CCO.O=C([O-])CC1CCCCC1	-3.68059182167053	15.1449298858642
Pyrr_3_72	CCCC[N+]1(C)CCCC1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	-16.0947036743164	4.24710750579833
Chol_5_72	C[N+](C)(C)CCO.O=C([O-])C(N)CC(C)C	-4.79062891006469	14.2884378433227
Chol_3_72	C[N+](C)(C)CCO.O=C([O-])C(C)N	-5.52838993072509	15.6727571487426
Chol_6_72	C[N+](C)(C)CCO.O=C([O-])C(N)C(C)CC	-4.82197189331054	14.2499446868896
Pyri_14_72	CCCCCCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	3.50151777267456	-14.6842365264892
Chol_25_71	C[N+](C)(C)CCO.[O-]C(=O)c1ccccc1	-3.2891616821289	13.4801225662231
Chol_32_70	C[N+](C)(C)CCO.Cl[Zn-](Cl)Cl	-5.48714685440063	17.3066253662109
Chol_22_69	C[N+](C)(C)CCO.[O-]C(=O)CC1CCCC1	-3.11094474792480	16.3323173522949
Chol_4_69	C[N+](C)(C)CCO.O=C([O-])C(N)C(C)C	-5.02757263183593	14.7509088516235

Pyri_1_69	OC(=O)C[n+]1ccccc1.[Br-]	12.9292669296264	-14.5266304016113
Chol_23_69	C[N+](C)(C)CCO.[O-]C(=O)C1CCCCC1	-3.80225753784179	15.2759160995483
Chol_21_69	C[N+](C)(C)CCO.[O-]C(=O)C1CCCC1	-3.36309957504272	16.3540287017822
Chol_19_69	C[N+](C)(C)CCO.O=C([O-])C	-5.2176284790039	17.163646697998
QAC_16_67	CC[N+]CC.[O-]C(C)=O	-8.2155475616455	10.7281951904296
Chol_17_67	C[N+](C)(C)CCO.[O-]C(=O)C1CCCN1	-7.33045291900634	16.216438293457
Chol_15_67	C[N+](C)(C)CCO.NC(Cc1cncn1)C([O-])=O	-2.05229449272155	14.4289722442626
Chol_31_67	C[N+](C)(C)CCO.[O-]C(=O)C(N)CS	-6.25245428085327	15.0045413970947
Chol_9_66	C[N+](C)(C)CCO.[O-]C(=O)C(N)CCSC	-6.43536710739135	14.1228103637695
Chol_20_66	C[N+](C)(C)CCO.NC(Cc1ccccc1)C([O-])=O	-3.79342484474182	12.6000175476074
Imid_75_65	CCOCC[n+]1cc[n](CC#N)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.8988800048828	0.837229132652282
Pyri_31_64	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)OCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.13777637481689	-6.32551288604736
Pyri_61_63	CCCC[n+]1ccccc1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	0.975685715675354	-12.5536394119262
Imid_76_62	COCCOCC[n+]1cc[n](CC#N)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.8693513870239	0.500636935234069
Imid_84_62	CCCC[n]1cc[n+](C)c1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	5.74297332763671	-2.08301901817321
Imid_77_62	COCCOCCOCC[n+]1cc[n](CC#N)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.8274383544921	0.216617226600646
Chol_26_60	C[N+](C)(C)CCO.O=C([O-])c1ccccc1O	-3.17580676078796	13.2621307373046
Pyri_36_59	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)OCCOCCOCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-3.51137590408325	-6.00865173339843
Imid_85_57	CC[n]1cc[n+](C)c1.O=C([O-])CN(C)C(=O)CCCCCCCCCCC	5.81585311889648	-2.02884793281555
Imid_3_57	CCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	4.48532533645629	-5.79037475585937
Pyri_29_56	CCOC(=O)c1ccc[n+](CC(=O)NC(Cc2ccccc2)C(O)=O)c1.[Br-]	5.67592811584472	-9.65869426727294
Imid_36_56	CCCCCOC(=O)C[n]1cc[n+](C)c1C.[O-]S(=O)(=O)OCCCCCCCC	7.1101369857788	-3.05688667297363
QAC_27_56	C[N+](C)(C)C.[O-]C(=O)[C@H]1C[C@@H](O)CN1	-8.45467090606689	16.9133701324462
QAC_18_55	C[N+](C)(C)C.[O-]C(=O)[C@@H]1CCCN1	-7.56101369857788	17.4599590301513
Imid_34_55	CCCOC(=O)C[n]1cc[n+](C)c1C.[O-]S(=O)(=O)OCCCCCCCC	7.48558759689331	-3.01871514320373
Imid_35_54	CCCCCOC(=O)C[n]1cc[n+](C)c1.[O-]S(=O)(=O)OCCCCCCCC	7.10984754562377	-2.61270141601562
Pyri_56_54	CCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C(Cc2ccccc2)NC(=O)C[n+]3ccccc3.[Br-]	0.783579647541046	-7.07958316802978
Pyri_20_54	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2.[Br-]	5.53487205505371	-8.89069843292236
Imid_11_53	C[n]1cc[n+](CC(=O)NC(Cc2ccccc2)C(=O)OCCCCCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4cc[n](C)c4)c1.[Br-]	-1.44878304004669	-4.52647304534912
Pyri_23_53	CCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2.[Br-]	2.97736310958862	-8.0173110961914
Pyri_33_52	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)OCCCCCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.53383922576904	-6.12260437011718
Pyri_52_52	CC[n+]1ccccc1.[Br-]	13.1154251098632	-13.5773162841796
Imid_2_51	CCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	5.08151483535766	-5.9729299545288
Phos_1_50	CCCC[P+](C)(CCCC)CCCC.COS(=O)(=O)[O-]	3.92416477203369	2.62208580970764

QAC_19_50	C[N+](C)(C)C.[O-]C(=O)[C@H]1CCCN1	-7.56016159057617	17.4596614837646
Pyri_8_49	CCCCCCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	3.69155597686767	-14.2212114334106
Chol_28_49	C[N+](C)(C)CCO.[O-]C(=O)c1c2ccccc2cc2ccccc21	-2.25006556510925	11.3612747192382
Imid_26_49	CCCOC(=O)C[n]1cc[n+](C)c1.[O-]S(=O)(=O)OCCCCCCCC	7.53869915008544	-2.58106994628906
Imid_4_48	CCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	3.86241245269775	-5.61261940002441
Pyri_18_48	CCCCCCCCCCCCOC[n+]1cccc(O)c1.[Cl-]	7.98306846618652	-14.2069120407104
Imid_1_47	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	5.75005006790161	-6.25369119644165
Imid_65_47	CCCC[n]1cc[n+](C)c1.CC(C)C(N)C([O-])=O	14.8855390548706	-4.56465530395507
Pyri_26_47	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cccc(OC)c2.[Br-]	5.81974220275878	-8.58088684082031
Pyri_43_46	CCOC(=O)C(Cc1ccc(O)cc1)NC(=O)C[n+]2cccc(OC)c2.[Br-]	5.95815086364746	-8.30684661865234
Imid_51_46	CC[n]1cc[n+](C)c1.CC1=CC=C(C=C1)S(=O)(=O)[O-]	14.3968429565429	-7.02977752685546
Pyri_22_46	CCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2.[Br-]	3.63105058670043	-8.29030323028564
Imid_9_44	C[n]1cc[n+](CC(=O)NC(Cc2ccccc2)C(=O)OCCCCCCCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4cc[n](C)c4)c1.[Br-]	-1.48044347763061	-4.51763200759887
Pyri_17_44	CCCCCCCOC[n+]1cccc(O)c1.[Cl-]	8.6590871810913	-14.1882972717285
Imid_5_44	CCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	3.33737015724182	-5.48866653442382
Pyri_10_43	CCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	5.94947290420532	-15.4586467742919
Chol_27_42	C[N+](C)(C)CCO.[O-]C(=O)COc1ccc2ccccc2c1	-2.32754230499267	11.5318698883056
Pyri_7_41	CCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	5.00788593292236	-14.1980791091918
Pyri_34_41	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)OCCCCCCCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.68101978302001	-5.88569068908691
Imid_8_40	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n]2cc[n+](CC(=O)NC(Cc3ccccc3)C(=O)OCC)c2.[Br-]	0.955089688301086	-5.82665395736694
Pyri_16_40	CCCOC[n+]1cccc(O)c1.[Cl-]	10.2421340942382	-14.124285697937
Imid_78_40	CCOC(=O)C(Cc1ccc(O)cc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	5.84905672073364	-6.38624048233032
Pyri_6_39	CCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	5.41390371322631	-14.2299604415893
QAC_28_38	C[N+](C)(C)C.[O-]C(=O)C1C[C@H](O)CN1	-8.52370262145996	17.5453300476074
Pyrr_2_37	CC[N+]1(C)CCCC1.SCC(N)C([O-])=O	-8.10406589508056	14.1254301071166
Pyri_28_37	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cccc(c2)C(O)=O.[Br-]	5.24513435363769	-9.37524700164794
Imid_54_37	CCCC[n]1cc[n+](C)c1.CC2=CC=C(S(=O)([O-])=O)C=C2	13.9493112564086	-6.47571039199829
Chol_37_36	CCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.21300792694091	2.90924763679504
QAC_26_36	C[N+](C)(C)C.[O-]C(=O)[C@@H]1C[C@@H](O)CN1	-8.53542518615722	16.9212646484375
Pyri_24_36	CCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2.[Br-]	2.46674966812133	-7.74947595596313
QAC_22_35	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)[C@@H]1CCCN1	-13.9218921661376	9.15426158905029
Imid_29_34	CCCOC(=O)C[n]1cc[n+](C)c1.N#C\N=C=[N-]	20.2988662719726	-3.80323529243469
QAC_29_33	C[N+](C)(C)C.[O-]C(=O)[C@@H]1C[C@H](O)CN1	-8.67232704162597	17.3657150268554
Pyri_30_33	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccc(cc2)N(C)C.[Br-]	4.85878705978393	-8.75059223175048

Imid_33_33	CCCCCOC(=O)C[n]1cc[n+](C)c1C.[Br-]	19.5842876434326	-2.5360631942749
Imid_66_33	CCCC[n]1cc[n+](C)c1.[O-]C(=O)CN	15.9784984588623	-5.40444898605346
Pyri_13_32	CCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	4.6415991783142	-15.0310678482055
Pyri_27_32	CCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cccc(c2)C(=O)OCC.[Br-]	6.187735080719	-9.39742755889892
Pyri_9_32	CCCCCCCCCCCCCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	1.6937916278839	-14.5220041275024
Pyri_12_31	CCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	5.0219097137451	-15.1908893585205
QAC_30_31	CC[N+](CC)(CC)CC.[O-]C(=O)[C@H]1C[C@H](O)CN1	-11.4784231185913	13.0544080734252
Pyri_50_31	CCCCOC(=O)C(c1ccccc1)[n+]2cccc(c2)C(=O)OCC.[Br-]	7.901930809021	-10.2107591629028
QAC_20_31	CC[N+](CC)(CC)CC.[O-]C(=O)[C@@H]1CCCN1	-10.5960483551025	13.3093738555908
Imid_69_30	CCCC[n]1cc[n+](C)c1.[O-]C(=O)C(N)Cc1ccccc1	13.4464435577392	-5.69276809692382
Imid_32_29	CCCCCOC(=O)C[n]1cc[n+](C)c1.[Br-]	19.1398525238037	-3.37090444564819
Imid_73_29	CCOCC[n+]1cc[n](CC#N)c1.N#C[N-]C#N	16.6803817749023	-9.47973823547363
Chol_36_29	CCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-5.86797857284545	3.37457299232482
Piperi_1_28	CC[N+]1(C)CCCCC1.[O-]C(=O)C(N)CS	-8.28321838378906	13.6789846420288
Chol_42_28	C[N+](C)(CCO)CC(=O)NC(Cc1ccccc1)C(=O)OCCCCCCCCOC(=O)C(Cc2ccccc2)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.2668137550354	-1.10086703300476
Chol_38_27	CCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.45601320266723	2.52527904510498
Imid_21_27	CCCCCCCCOC(=O)C[n]1cc[n+](C)c1.[Br-]	18.39888381958	-2.56711840629577
Imid_19_26	CCCCOC(=O)C[n]1cc[n+](C)c1.[Br-]	19.3928279876708	-3.72734451293945
Imid_40_26	CCCCCCCCCC[n]1cc[n+](C)c1.[Br-]	16.6098480224609	-3.58249330520629
Pyri_19_25	CCCCCCCCCCCCCCCCCCOC[n+]1cccc(O)c1.[Cl-]	2.02470588684082	-13.635495185852
Imid_13_25	CCCC[n]1cc[n+](C)c1.CCCCCCCCOS(=O)(=O)[O-]	7.74207830429077	-1.78488612174987
Imid_30_25	CCCOC(=O)C[n]1cc[n+](C)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.8352842330932	1.503093957901
Pyri_68_25	CCCCCCCCNC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccc(cc2)C=NO.[Br-]	1.59297549724578	-9.14685153961181
QAC_33_25	CC[N+](CC)(CC)CC.[O-]C(=O)[C@H]1CC(O)CN1	-11.2249193191528	12.7484045028686
QAC_23_25	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)[C@H]1CCCN1	-13.9218406677246	9.15444660186767
Chol_34_24	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-4.6269416809082	5.46800279617309
Pyri_4_24	CCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	6.23346948623657	-14.6206903457641
Imid_20_24	CCCCCCOC(=O)C[n]1cc[n+](C)c1.[Br-]	18.82590675354	-3.01911234855651
Chol_40_24	C[N+](C)(CCO)CC(=O)NC(Cc1ccccc1)C(=O)OCCCCCCCCCCOC(=O)C(Cc2ccccc2)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.26619100570678	-1.12055337429046
Pyri_66_24	CCCCCCCCNC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2C=NO.[Br-]	1.30206203460693	-8.72602558135986
QAC_6_23	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H]1[C@H](O)[C@H](O)[C@@H](O)C(O)O1	-17.9732437133789	6.27849817276
Imid_31_23	CCCOC(=O)C[n]1cc[n+](C)c1C.[Br-]	20.0255737304687	-2.9039876461029
Imid_74_23	CCOCC[n+]1cc[n](CC#N)c1.FC(F)(F)C([O-])=O	16.6586875915527	-9.5876932144165
Prol_5_23	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+]2(C)CCCC2C(=O)OCC.[Br-]	-6.12547540664672	4.93962574005126

QAC_37_22	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)[C@@H]1C[C@H](O)CN1	-14.7729158401489	9.16037464141845
Pyri_5_22	CCCCOC[n+]1cccc(O)c1.[N-]1C(=O)C=C(C)OS1(=O)=O	5.96579504013061	-14.4321346282958
Prol_4_22	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+]2(C)CCCC2C(=O)OCC.[Br-]	-6.27430915832519	5.0576696395874
Imid_25_21	CCCOC(=O)C[n]1cc[n+](C)c1.[Br-]	19.6348304748535	-4.1734185218811
Imid_18_21	CCOC(=O)C[n]1cc[n+](C)c1.[Br-]	19.8564758300781	-4.81188011169433
Prol_6_21	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+]2(C)CCCC2C(=O)OCC.[Br-]	-6.27458429336547	5.06069374084472
Pyri_47_21	CCCCOC(=O)C(c1ccccc1)[n+]2ccccc2.[Br-]	9.24876022338867	-10.7573413848876
Chol_39_20	CCCCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.59240293502807	2.26763272285461
QAC_32_20	CC[N+](CC)(CC)CC.[O-]C(=O)[C@H]1C[C@@H](O)CN1	-11.2249193191528	12.7484045028686
QAC_36_20	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)C1C[C@H](O)CN1	-14.9042158126831	9.65708923339843
Chol_43_20	CCOC(=O)C(Cc1ccc(O)cc1)NC(=O)C[N+](C)(C)CCO.[Br-]	-4.38632345199584	5.23766088485717
Imid_80_20	CCOC(=O)C(c1ccccc1)[n+]2cc[n](C)c2.[Br-]	12.6929006576538	-7.9522738456726
Pyri_15_20	CCCCCCCCCCCCCCCCCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	1.66072750091552	-14.5969476699829
Pyri_67_20	CCCCCCCCNC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cccc(C=NO)c2.[Br-]	1.55907976627349	-8.95856475830078
QAC_41_19	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)C1C[C@H](O)CN1	-16.1742362976074	7.18169260025024
QAC_24_19	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H]1CCCN1	-15.1174163818359	7.31055164337158
QAC_5_19	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H]1[C@@H](O)[C@H](O)[C@@H](O)C(O)O1	-17.9735393524169	6.27843809127807
Imid_67_19	CCCC[n]1cc[n+](C)c1.[O-]C(=O)C(N)CS	14.819200515747	-5.0198802947998
QAC_15_19	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)C(N)CS	-15.2551479339599	6.3149185180664
Chol_41_19	C[N+](C)(CCO)CC(=O)NC(Cc1ccccc1)C(=O)OCCCCCCCCCCCCCCOC(=O)C(Cc2ccccc2)NC(=O)C[N+](C)(C)CCO.[Br-]	-6.2657399177551	-1.13530910015106
Prol_2_19	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+]2(C)CCCC2C(O)=O.[Br-]	-5.6663537025452	5.69187021255493
Imid_63_18	COC(=O)C[n]1cc[n+](C)c1.[Br-]	19.7156372070312	-5.20225620269775
QAC_14_18	CC[N+](CC)(CC)CC.[O-]C(=O)C(N)CS	-9.1321821212769	13.5283708572387
Pyri_35_18	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)OCCCCCCCCCCCCCOC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.7738959789276	-5.60225629806518
Imid_72_17	CCCCCC[n]1cc[n+](C)c1.[O-]C(=O)C(N)CS	14.3366813659667	-4.59002780914306
Morph_2_17	CCOC(=O)C(Cc1ccccc1)NC(=O)C[N+]2(C)CCOCC2.[Br-]	-4.9516863822937	4.82453870773315
QAC_35_17	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)[C@H]1C[C@@H](O)CN1	-14.5710945129394	9.73230743408203
QAC_25_17	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@H]1CCCN1	-15.1174163818359	7.31055164337158
QAC_34_17	CCC[N+](CCC)(CCC)CCC.[O-]C(=O)[C@@H]1C[C@@H](O)CN1	-15.0354957580566	9.32922649383544
QAC_39_17	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@H]1C[C@@H](O)CN1	-16.2392883300781	7.41394090652465
QAC_46_16	CCCCCCCCCCCCCC[N+](C)(C)Cc1ccccc1.[Cl-]	-13.5167140960693	4.34703874588012
Imid_27_16	CCCOC(=O)C[n]1cc[n+](C)c1.F[B-](F)(F)F	21.3700885772705	-6.50650644302368
QAC_8_16	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H](O)[C@H](O)C(O)=O	-16.3372535705566	5.82237386703491
QAC_7_16	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@H](C)O	-14.7219638824462	6.17397880554199

Imid_28_16	CCCOC(=O)C[n]1cc[n+](C)c1.F[P-](F)(F)(F)(F)F	21.5415592193603	-10.3002576828002
Pyri_46_16	CCOC(=O)C(c1ccccc1)[n+]2ccccc2.[Br-]	9.88002300262451	-10.8908939361572
Imid_71_16	CC[n]1cc[n+](C)c1.SCC(N)C([O-])=O	15.1960134506225	-5.69440650939941
Prol_3_15	CCOC(=O)C1CCC[N+]1(C)CC(=O)NC(Cc2ccccc2)C(O)=O.[Br-]	-5.70781278610229	5.83898496627807
Imid_6_15	CCCCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2cc[n](C)c2.[Br-]	3.00310373306274	-5.39265632629394
QAC_45_14	CN(C)C(=O)Oc1cccc(c1)[N+](C)(C)C.COS(=O)(=O)[O-]	-0.667103886604309	12.6925706863403
QAC_9_14	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H](O)CC(=O)O	-15.8717374801635	5.66511440277099
Imid_53_14	CCCC[n]1cc[n+](C)c1.O=S(OC)([O-])=O	15.8279113769531	-6.12765312194824
QAC_31_13	CC[N+](CC)(CC)CC.[O-]C(=O)[C@@H]1C[C@H](O)CN1	-11.3693714141845	13.1874265670776
Pyri_11_13	CCCCOC[n+]1cccc(O)c1.[N-]1C(=O)c2ccccc2S1(=O)=O	5.57424306869506	-15.3808050155639
QAC_12_12	CCCC[N+](CCCC)(CCCC)CCCC.CC(=O)C([O-])=O	-14.3286228179931	5.6256628036499
QAC_38_12	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H]1C[C@@H](O)CN1	-15.9990396499633	7.69440460205078
QAC_11_10	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)CC(O)=O	-14.9612131118774	5.48898363113403
Pyri_60_10	CCCC[n+]1ccc(C)cc1.[Cl-]	12.1130313873291	-13.4023532867431
Pyri_65_9	CCCCCCCCNC(=O)C(C)NC(=O)C[n+]1ccc(cc1)C=NO.[Br-]	2.22853279113769	-9.94932651519775
Imid_55_9	CC[n]1cc[n+](C)c1.O=C([O-])C	17.4313564300537	-6.75688934326171
Imid_44_9	CCCCCC[n]1c[n+](C)c(C)c1C.[I-]	16.4161834716796	-1.8695763349533
Imid_48_9	CC[n]1cc[n+](C)c1.COS(=O)(=O)[O-]	16.277997970581	-6.96875619888305
QAC_4_8	CCCC[N+](CCCC)(CCCC)CCCC.[Br-]	-13.2546215057373	6.13899564743041
Pyri_45_8	COC(=O)C(c1ccccc1)[n+]2ccccc2.[Br-]	10.0853462219238	-10.9872941970825
Imid_15_7	CCCC[n]1cc[n+](C)c1.F[P-](F)(F)(F)(F)F	21.5092048645019	-10.3088836669921
Pyri_3_7	CN(C)c1cc[n+](CC(O)=O)cc1.[Br-]	12.3336420059204	-15.1538877487182
Imid_38_7	CCCCCC[n]1cc[n+](C)c1.[Br-]	17.4007606506347	-4.89733362197875
Pyri_41_7	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)NCCOCCOCCNC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-3.61369204521179	-6.94945240020751
Imid_41_7	CC[n]1c(C)c(C)[n+](C)c1C(C)C.[I-]	16.2888088226318	-1.25412285327911
Imid_43_6	CC[n]1c[n+](C)c(C)c1C.[I-]	16.3416099548339	-1.46455156803131
Imid_57_6	C[n]1cc[n+](CC=C)c1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	-4.46640443801879	-14.6331701278686
Imid_12_5	CCCC[n]1cc[n+](C)c1.[Br-]	17.9534454345703	-5.63580560684204
Chol_33_5	C[N+](C)(CCO)CC(=O)NC(Cc1ccccc1)C(O)=O.[Br-]	-4.31783580780029	6.47932863235473
Imid_17_5	CCCC[n]1cc[n+](C)c1.C(F)(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	11.7825717926025	1.67416656017303
Imid_24_5	CCN(CC)OC(=O)C[n]1cc[n+](C)c1.[Br-]	21.8344421386718	-3.8577971458435
Imid_23_5	CCCCN(C)OC(=O)C[n]1cc[n+](C)c1.[Br-]	21.5011959075927	-3.80479884147644
Pyri_51_5	CC[n+]1ccc(C)cc1.[Br-]	12.9039916992187	-13.3069705963134
Pyri_2_5	COc1ccc[n+](CC(O)=O)c1.[Br-]	13.1469249725341	-14.8747549057006

Imid_22_5	CCCCNOC(=O)C[n]1cc[n+](C)c1.[Br-]	18.4880065917968	-4.11404275894165
Imid_79_5	COC(=O)C(c1ccccc1)[n+]2cc[n](C)c2.[Br-]	12.7063598632812	-7.97579908370971
Pyri_64_4	CCCCCCCCNC(=O)C[n+]1ccc(cc1)C=NO.[Br-]	2.6361756324768	-10.4159469604492
Pyri_25_4	CCCCCCCCCCCCOC(=O)C(Cc1ccccc1)NC(=O)C[n+]2ccccc2.[Br-]	2.15518069267272	-7.55117464065551
Imid_39_4	CC[n]1cc[n+](C)c1.[Br-]	18.5165653228759	-6.71830892562866
QAC_3_4	CCCC[N+](CCCC)(CCCC)CCCC.[OH-]	-13.2554597854614	6.13900661468505
Pyri_37_4	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)NCCCCCCNC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-3.07766270637512	-7.37985277175903
Imid_16_3	CCCC[n]1cc[n+](C)c1.N#CN=C=[N-]	16.742073059082	-5.8661789894104
QAC_1_3	C[N+](C)(C)C.[Br-]	-5.9220643043518	18.9447154998779
Imid_59_3	CCCCCC[n+]1cc[n](C)c1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	-4.6509952545166	-14.306616783142
Pyri_53_3	CCCC[n+]1ccccc1.[Br-]	12.3013162612915	-13.7580137252807
Imid_60_2	C[n]1cc[n+](CCO)c1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	-4.57549285888671	-14.6896724700927
Pyri_57_1	CN(C)C(=O)Oc1ccc[n+](C)c1.[Br-]	14.1724615097045	-15.0681076049804
Pyri_48_1	CCCCNC(=O)C(c1ccccc1)[n+]2ccccc2.[Br-]	9.3696060180664	-10.5075769424438
Imid_37_1	CC[n]1cc[n+](C)c1.[Cl-]	18.5165653228759	-6.71830892562866
Pyri_38_1	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)NCCCCCCCCNC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.8611762523651	-7.26053094863891
Pyri_54_0	CCCC[n+]1ccc(C)cc1.[Cl-]	12.1137971878051	-13.4016714096069
Imid_61_0	CC[n]1cc[n+](C)c1.O=S(C(F)(F)F)([O-])=O	15.4600315093994	-7.84767150878906
Imid_56_0	C[n]1cc[n+](CC(O)=O)c1.[Br-]	18.9134082794189	-7.47907638549804
Imid_45_0	CC[n]1cc[n+](C)c1.[B-](F)(F)(F)F	20.912124633789	-7.08776330947875
QAC_10_0	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)CCC(=O)O	-15.3554134368896	5.36006641387939
Imid_62_0	CCC[n+]1cc[n](C)c1.[I-]	18.6227226257324	-6.21920871734619
Imid_58_0	CCCC[n+]1cc[n](C)c1C.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	-4.5466704368591	-14.3162717819213
QAC_13_0	CCCC[N+](CCCC)(CCCC)CCCC.OS(=O)(=O)[O-]	-14.1138019561767	6.3174729347229
Imid_50_0	CC[n]1cc[n+](C)c1.C(=[N-])=NC#N	17.245132446289	-7.07140016555786
Imid_52_0	CC[n]1cc[n+](C)c1.F[P-](F)(F)(F)(F)F	21.5003280639648	-10.3066329956054
Imid_49_0	CC[n]1cc[n+](C)c1.CCOS(=O)(=O)[O-]	16.0023708343505	-7.01844644546508
Pyri_58_0	C[n+]1ccccc1C=NO.[Cl-]	13.9876689910888	-13.4650001525878
Imid_82_0	CCCCNC(=O)C(c1ccccc1)[n+]2cc[n](C)c2.[Br-]	12.2079467773437	-7.27972793579101
QAC_2_0	CC[N+](CC)(CC)CC.[Br-]	-3.8563106060028	18.979694366455
Pyri_39_0	O=C(C[n+]1ccccc1)NC(Cc2ccccc2)C(=O)NCCCCCCCCCCNC(=O)C(Cc3ccccc3)NC(=O)C[n+]4ccccc4.[Br-]	-2.6003561019897	-7.26533126831054
Imid_46_0	CC[n]1cc[n+](C)c1.COP(=O)([O-])OC	16.4550170898437	-7.60517311096191
Morph_1_0	CCCC[N+]1(C)CCOCC1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	-5.05876398086547	-14.4798145294189

Table S10: tSNE coordinates for tSNE plot train_set_ILNI.

Name	SMILES	TSNE_1	TSNE_2
Piperi_2_100	O=C([O-])CN(C)C(=O)CCCCCCCCCCC.[NH2+]1CCCCC1	-4.168165521	1.1301099
Morph_3_100	O=C([O-])CN(C)C(=O)CCCCCCCCCCC.C1C[NH2+]CCO1	-4.202221023	1.090324761
NI_11_100	OCCN(C)CC(O)=O	41.51184064	15.50894864
Quat_31_99	[Cl-].O=C(OCCCC)[C@@H]([NH3+])Cc1ccccc1	5.269448687	8.324394137
Chol_46_98	C[N+](C)(C)CC(CC(=O)O)O.[Cl-]	2.57023505	43.41399206
NI_9_97	C(C[C@@H](C(=O)O)N)CNC(=N)O	28.1610961	14.17622529
NI_37_93	c1cc(cnc1)O	23.3456031	-0.036890565
Quat_30_92	[Cl-].O=C(OCC)[C@@H]([NH3+])Cc1ccccc1	6.598904692	8.247170359
NI_81_92	C(CN)C(=O)O	31.33391424	7.281841396
NI_31_91	CCCCNCCCC	45.16599879	9.621252609
NI_41_91	CCCCOCCOCCO	41.32966204	0.134451478
NI_56_91	C1=CC=C(C=C1)CC(C(=O)O)N	8.242929697	8.481162712
NI_18_89	CCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	5.468820863	7.885792754
NI_57_89	CC(C(=O)O)N	31.27602821	6.382873337
NI_6_88	C1=CC=C2C(=C1)C(=CN2)CC(C(=O)O)N	7.673508579	31.13124811
Anion_3_88	CCCCCCCCCCCCOS(=O)(=O)[O-].[Na+]	-0.4926201	1.057384128
Chol_12_87	O=C([O-])C(N)CC(N)=O.C[N+](C)(C)CCO	2.02949084	37.88910828
Chol_13_87	O=C([O-])C(N)CCC(N)=O.C[N+](C)(C)CCO	1.670028741	36.63436067
NI_8_87	CC(C)C(C(=O)O)N	29.79527255	10.60055196
NI_59_87	C(CC(C(=O)O)N)CN=C(N)N	26.32837341	14.2746257
NI_36_86	c1ccnc(c1)O	23.76381236	0.171653663
NI_78_86	CCC(=O)O	32.77259579	6.316309051
Chol_11_85	OC(=O)CCC(N)C([O-])=O.C[N+](C)(C)CCO	1.24182741	36.8501164
Quat_32_85	[Cl-].O=C(OCCCCCC)[C@@H]([NH3+])Cc1ccccc1	4.05589574	8.045497379
NI_3_85	C1CC(NC1)C(=O)O	33.0946836	12.95110173
NI_44_85	CC(=CC(=O)C)C	36.28107237	5.759744649
NI_45_85	CC(C)CC(=O)C	35.06578508	6.573941457
Pyri_59_84	C[n+]1cccc(c1)C(=O)O.[Cl-]	-40.26136366	2.017449976
NI_40_84	CCCCOCCO	40.67938736	0.694121437
NI_58_84	C(C(C(=O)O)N)C(=O)O	27.51422859	10.99560066
NI_60_84	C(C(C(O)=O)N)C(N)=O	27.03471316	11.19352963
NI_70_84	CO	33.55565768	2.940083208

NI_65_83	OC1CNC(C1)C(O)=O	33.09447484	12.99665431
Chol_30_83	C[N+](C)(C)CCO.[O-]C(=O)CCC(C)C1CC[C@H]2[C@@H]3CC[C@@H]4C[C@H](O)CC[C@]4(C)[C@H]3CCC12C	-17.57641036	18.42680453
NI_52_83	CC(C)CNCC(C)C	43.50832784	8.549003297
Chol_1_82	C[N+](C)(C)CCO.[Cl-]	1.199156784	44.61068176
Chol_10_82	OC(=O)CC(N)C([O-])=O.C[N+](C)(C)CCO	1.708260704	38.2605536
Anion_2_82	c1ccc(cc1)C(=O)[O-].[Na+]	20.76170584	2.59423968
Chol_47_81	C[N+](C)(C)CCOC(=O)N.[Cl-]	2.212917268	45.14125584
NI_7_81	C1=CC(=CC=C1CC(C(=O)O)N)O	8.512903098	8.875362125
NI_16_81	C1CCNCC1	44.91841222	9.245235378
QUAT_40_81	[Cl-].C[NH+](C)C	34.13504017	3.293833198
QUAT_43_0	[Cl-].[NH3+]CCCC	39.59521504	8.5964304
Pyri_62_80	Cc1cc[n+](CCCCCCCC)cc1C.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	-18.0238446	-1.607312931
NI_29_80	C(CCN)CC(C(=O)O)N	27.45796344	13.81870216
NI_39_80	c1cc(nc(c1)C(=O)O)C(=O)O	21.18397861	0.652622963
NI_50_80	CCCCCCCCCC(=O)O	-1.27127177	-0.031598491
NI_73_80	c1ccc(c(c1)C(=O)O)O	20.79354402	2.527505082
Pyrr_1_79	CCCC[N+]1(CCCC1)C.C(F)(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	19.03504267	-22.65445787
Quat_33_79	[Cl-].O=C(OCCCCCCCC)[C@@H]([NH3+])Cc1ccccc1	2.81446694	7.435884507
NI_5_79	CC(C(C(=O)O)N)O	29.35792858	9.530025075
NI_34_79	CN(C)CCO	40.58101017	14.94170989
NI_49_79	C1=CC(=CN=C1)C(=O)N	22.74549639	0.587898122
NI_69_79	CCC(C)C(C(=O)O)N	29.81484118	11.18736239
NI_79_79	CCCCC(=O)O	33.86596455	8.159313522
NI_55_78	C(C(=O)O)N	31.54588287	5.886906335
Pyri_61_78	CCCCCCCCCCCC(=O)N(C)CC([O-])=O.CCCCCC[n+]1ccccc1	-17.91434971	-2.129320559
Chol_29_78	C[N+](C)(C)CCO.[O-]C(=O)CC[C@H](C)[C@H]1CC[C@H]2[C@@H]3CC[C@@H]4C[C@H](O)CC[C@]4(C)[C@H]3C[C@H](O)[C@@]21C	-17.57759708	18.42773625
Pyri_74_78	C[n+]1cccc(c1)C(=O)N.[I-]	-40.03256672	1.880176846
QUAT_41_78	[Cl-].CC(C)C[NH3+]	33.05759021	4.884342952
NI_4_77	C(C(C(=O)O)N)O	29.8246119	8.746102487
Anion_4_77	CCCCCCCCCCCC(=O)N(C)CC(=O)[O-].[Na+]	-2.723434244	0.784262199
Chol_7_76	OCC(N)C([O-])=O.C[N+](C)(C)CCO	0.890483982	40.30011027
Chol_18_76	C[N+](C)(C)CCO.[O-]C(=O)C(N)Cc1c[NH]c2ccccc21	7.47770161	31.53696821
Chol_45_76	C[N+](C)(C)CC(=O)[O-]	2.740276481	41.0829872

NI_51_75	CCCCCN	39.59979337	8.623558841
Chol_8_74	OCC[N+](C)(C)C.[O-]C(=O)C(N)C(C)O	0.752744482	39.50911312
Chol_19_74	O=C([O-])C.C[N+](C)(C)CCO	1.434515593	43.61608054
Chol_44_74	O=C([O-])CN(C)C(=O)CCCCCCCCCCC.C[N+](C)(C)CCO	-5.888977819	2.194721849
Chol_2_73	O=C([O-])CN.C[N+](C)(C)CCO	1.29433192	42.7364383
Chol_16_73	O=C([O-])C(N)CCCNC(=N)N.C[N+](C)(C)CCO	1.713930548	35.00732014
Chol_24_73	O=C([O-])CC1CCCCC1.C[N+](C)(C)CCO	-4.35956756	34.96007924
NI_47_73	C=CCO	32.9630534	2.686957114
NI_48_73	CCOCCOCCOC(=O)C	41.68063077	-0.17069442
NI_62_73	C(CC(=O)O)C(C(=O)O)N	27.71636955	12.38512992
Pyrr_4_72	CCCC[N+]1(C)CCCC1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	-6.10492531	2.362577076
Pyri_14_72	[n+]1(cc(O)ccc1)COCCCCCCCCCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-22.56817478	1.439575527
Chol_3_72	O=C([O-])C(C)N.C[N+](C)(C)CCO	1.199310826	41.84391399
Chol_5_72	O=C([O-])C(N)CC(C)C.C[N+](C)(C)CCO	0.074806166	38.07138724
Chol_6_72	O=C([O-])C(N)C(C)CC.C[N+](C)(C)CCO	-0.322792228	38.57175734
NI_17_72	C1COCCN1	43.10957476	4.856427009
Chol_25_71	[O-]C(=O)c1ccccc1.C[N+](C)(C)CCO	6.567566919	37.40049527
NI_42_71	CC(=O)CC(=O)C	35.69919398	6.448567042
Chol_32_70	Cl[Zn-](Cl)Cl.C[N+](C)(C)CCO	0.585290898	44.35613748
NI_61_70	C(CC(=O)N)C(C(=O)O)N	27.23998652	12.44185321
Pyri_77_70	[Cl-].Oc1c(C=O)c(c[nH+]c1C)CO	23.10803907	-1.396034225
Pyri_1_69	[Br-].OC(=O)C[n+]1ccccc1	-38.22417345	0.876527488
Chol_4_69	O=C([O-])C(N)C(C)C.C[N+](C)(C)CCO	-0.183781245	39.33730317
Chol_21_69	OCC[N+](C)(C)C.[O-]C(=O)C1CCCC1	-3.773203509	36.69320823
Chol_22_69	[O-]C(=O)CC1CCCC1.C[N+](C)(C)CCO	-3.78178668	36.58471684
Chol_23_69	[O-]C(=O)C1CCCCC1.C[N+](C)(C)CCO	-4.40051879	34.91625484
NI_54_69	CCCC(=O)O	33.71704278	7.676122083
Pyri_70_68	C[N+]1=CC=CC=C1.F[P-](F)(F)(F)(F)F	-0.042069124	-45.53569255
Chol_14_68	O=C([O-])C(N)CCCCN.C[N+](C)(C)CCO	0.711643751	35.64410495
Chol_15_67	NC(Cc1cnc[NH]1)C([O-])=O.C[N+](C)(C)CCO	3.32977713	38.41834279
Chol_17_67	OCC[N+](C)(C)C.[O-]C(=O)C1CCCN1	-1.460442534	39.6679231
Chol_31_67	OCC[N+](C)(C)C.[O-]C(=O)C(N)CS	0.356701199	41.04710224
Quat_15_67	CC[NH2+]CC.[O-]C(C)=O	42.85762445	6.785198309
Quat_36_67	[Cl-].O=C(OCCCCCCCCCCCCCC)[C@@H]([NH3+])Cc1ccccc1	-0.66241852	6.730428619

Chol_9_66	[O-]C(=O)C(N)CCSC.C[N+](C)(C)CCO	-0.36649246	36.81434236
Chol_20_66	NC(Cc1ccccc1)C([O-])=O.C[N+](C)(C)CCO	5.397162473	37.13818203
Imid_75_65	N#CCn1cc[n+](CCOCC)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	19.00377571	-24.97323987
Quat_35_65	[Cl-].O=C(OCCCCCCCCCCCC)[C@@H]([NH3+])Cc1ccccc1	0.430862762	6.629280083
NI_38_65	c1c[nH]ccc1=O	24.72186898	0.098037162
Pyri_31_64	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-43.58305223	-19.7519055
Pyri_60_63	CCCCCCCCCCCC(=O)N(C)CC([O-])=O.CCCC[n+]1ccccc1	-17.77728035	-2.404233095
Imid_76_62	N#CCn1cc[n+](CCOCCOC)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	19.08691281	-25.48869526
Imid_77_62	N#CCn1cc[n+](CCOCCOCCOC)c1.FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	19.1944681	-25.98047827
Imid_84_62	CCCCn1cc[n+](C)c1.CCCCCCCCCCCC(=O)N(C)CC([O-])=O	-6.82605472	-5.209869836
Quat_16_61	CC[NH+](CC)CC.[O-]C(C)=O	40.32923626	13.89412561
Quat_38_61	[O-]C=O.[NH3+]CCO	31.3219585	4.256073238
Chol_26_60	O=C([O-])c1ccccc1O.C[N+](C)(C)CCO	6.747981418	37.41073098
Pyri_36_59	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCOCCOCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-44.74463111	-19.43532084
NI_64_58	CC(C)CC(C(=O)O)N	29.39647522	11.72904005
NI_68_58	C(C(C(=O)O)N)S	29.62093242	8.128757988
Imid_3_57	O=C([C@H](CC1=CC=CC=C1)NC(C[N+]2=CN(C=C2)C)=O)OCCCCCC.[Br-]	-17.38958783	-10.44414844
Imid_85_57	O=C([O-])CN(C)C(=O)CCCCCCCCCCC.C[n+]1ccn(CC)c1	-6.80887768	-5.223367701
Imid_36_56	[O-]S(=O)(=O)OCCCCCCCC.O=C(OCCCCC)Cn1cc[n+](C)c1C	-4.735984029	-8.809469218
Pyri_29_56	CCOC(C1=CC=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(O)=O)=O)=C1)=O.[Br-]	-23.19388295	-16.15665378
NI_63_56	C1=C(NC=N1)CC(C(=O)O)N	25.78627052	10.78510798
Imid_34_55	[O-]S(=O)(=O)OCCCCCCCC.O=C(OCCC)Cn1cc[n+](C)c1C	-5.259866633	-9.318693449
Pyri_32_55	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-43.6644138	-20.30375969
NI_19_55	CCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	4.295309964	7.497153182
NI_24_55	CCCCCCCCCCCCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	-1.367135191	5.927152546
Imid_35_54	[O-]S(=O)(=O)OCCCCCCCC.O=C(OCCCCC)Cn1cc[n+](C)c1	-4.476823692	-9.448249994
Pyri_20_54	O=C(N[C@H](C(OCC)=O)CC1=CC=CC=C1)C[N+]2=CC=CC=C2.[Br-]	-22.65289676	-14.62405173
Pyri_56_54	O=C(N[C@@H](CC1=CC=CC=C1)C(N[C@@H](CC2=CC=CC=C2)C(OCCCCCC)=O)=O)C[N+]3=CC=CC=C3.[Br-]	-27.67328889	-12.80167985
NI_30_54	CSCCC(C(=O)O)N	29.47376096	12.89084705
Imid_11_53	CN1C=C[N+](CC(N[C@H](C(OCCCCCCCCCOC([C@@H](NC(C[N+]2=CN(C)C=C2)=O)CC3=CC=CC=C3)=O)=O)CC4=CC=CC=C4)=O)=C1.[Br-].[Br-]	-38.73454629	-20.46722372
Pyri_23_53	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCC)=O)C[N+]2=CC=CC=C2.[Br-]	-23.53541428	-11.42386176
Quat_17_53	[O-]C(=O)[C@@H]1CCCN1.C[N+](C)(C)C	6.833841692	44.29781825
Pyri_33_52	O=C(N[C@H](C(OCCCCCCCCCOC([C@H](CC1=CC=CC=C1)NC(C[N+]2=CC=CC=C2)=O)=O)=O)=O)CC3=CC=CC=C3)C[N+]4=CC=CC=C4.[Br-].[Br-]	-43.30620164	-20.8751982
Quat_37_52	[Cl-].O=C(OCCCCCCCCCCCCCCCCC)[C@@H]([NH3+])Cc1ccccc1	-1.473229125	6.52364833

Imid_2_51	O=C([C@H](CC1=CC=CC=C1)NC(C[N+]2=CN(C=C2)C)=O)OCCCC.[Br-]	-16.77989738	-11.09013709
Phos_1_50	CCCC[P+](C)(CCCC)CCCC.COS(=O)(=O)[O-]	0.487945641	1.041298075
Quat_34_50	[Cl-].O=C(OCCCCCCCCCC)[C@@H]([NH3+])Cc1ccccc1	1.614662027	6.904820193
Imid_26_49	[O-]S(=O)(=O)OCCCCCCCC.O=C(OCCC)Cn1cc[n+](C)c1	-4.842853649	-9.879236898
Pyri_8_49	[n+]1(cc(O)ccc1)COCCCCCCCCCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-24.95202156	-0.999018586
Pyri_21_49	CCCCOC([C@@H](NC(C[N+]1=CC=CC=C1)=O)CC2=CC=CC=C2)=O.[Br-]	-22.99509002	-13.62750997
Pyri_42_49	CCOC([C@H](CC1=CC=C(O)C=C1)NC(C[N+]2=CC=CC=C2)=O)=O.[Br-]	-21.61486099	-15.42439896
Chol_28_49	C[N+](C)(C)CCO.[O-]C(=O)c1c2ccccc2cc2ccccc21	6.703150317	33.36526009
Imid_4_48	CN1C=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCCCCCCCCC)=O)=O)=C1.[Br-]	-18.03928385	-9.859189084
Pyri_18_48	[n+]1(cc(O)ccc1)COCCCCCCCCCCC.[Cl-]	-30.12231549	-0.155290331
Imid_1_47	[Br-].O=C(C[n+]1ccn(C)c1)N[C@@H](Cc1ccccc1)C(=O)OCC	-16.25170812	-11.73193942
Imid_65_47	CCCCn1cc[n+](C)c1.CC(C)C(N)C([O-])=O	-1.384168007	-30.55317564
Pyri_26_47	COC1=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCC)=O)=O)=CC=C1.[Br-]	-21.75436565	-14.38169874
Pyri_69_47	CCOC(=O)C[N+]1=CC=CC=C1.[Br-]	-36.9408144	1.791706699
NI_53_47	CC(CN)O	32.58149099	4.604822398
Imid_51_46	CCN1C=C[N+](=C1)C.CC1=CC=C(C=C1)S(=O)(=O)[O-]	-4.04837953	-35.39814837
Imid_70_46	CCCCn1cc[n+](C)c1.[O-]C(=O)C(N)CC(=O)O	-2.20602047	-30.62752878
Pyri_22_46	O=C([C@H](CC1=CC=CC=C1)NC(C[N+]2=CC=CC=C2)=O)OCCCCCC.[Br-]	-23.3191981	-12.35175544
Pyri_43_46	CCOC([C@H](CC1=CC=C(O)C=C1)NC(C[N+]2=CC(OC)=CC=C2)=O)=O.[Br-]	-21.04493861	-15.02975578
Pyri_55_46	O=C(N[C@@H](CC1=CC=CC=C1)C(N[C@@H](CC2=CC=CC=C2)C(OCCCC)=O)=O)C[N+]3=CC=CC=C3.[Br-]	-27.69501948	-12.85429455
NI_14_45	OCCN(C)CC(N[C@@H](CC1=CC=CC=C1)C(OCC)=O)=O	3.480790876	13.37404125
Imid_5_44	CN1C=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCCCCCCCCCC)=O)=O)=C1.[Br-]	-18.60819771	-9.415035003
Imid_9_44	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CN(C)C=C3)=O)=O)=O)C[N+]4=CN(C)C=C4.[Br-].[Br-]	-39.03674402	-20.34936884
Pyri_17_44	[n+]1(cc(O)ccc1)COCCCCCCC.[Cl-]	-30.75216881	0.027377054
Pyri_10_43	[n+]1(cc(O)ccc1)COCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-24.28708633	2.954290266
Chol_27_42	C[N+](C)(C)CCO.[O-]C(=O)COc1ccc2ccccc2c1	6.931767822	33.02598095
Pyri_7_41	[n+]1(cc(O)ccc1)COCCCCCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-25.88017927	-0.931274499
Pyri_34_41	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-42.70510734	-21.00352546
Imid_8_40	CCOC([C@H](CC1=CC=CC=C1)NC(CN2C=C[N+](CC(N[C@@H](CC3=CC=CC=C3)C(OCC)=O)=O)=C2)=O)=O.[Br-]	-28.68863346	-13.5999403
Imid_10_40	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CN(C)C=C3)=O)=O)=O)C[N+]4=CN(C)C=C4.[Br-].[Br-]	-39.23469619	-20.2459143
Imid_78_40	CCOC([C@H](CC1=CC=C(O)C=C1)NC(C[N+]2=CN(C)C=C2)=O)=O.[Br-]	-16.4252457	-12.1592158
Pyri_16_40	[n+]1(cc(O)ccc1)COCCC.[Cl-]	-31.62386281	0.295497523
Quat_21_40	[O-]C(=O)[C@@H]1C[C@@H](O)CN1.C[N+](C)(C)C	6.836001743	44.28854017
NI_43_40	CC(CCO)OC	39.57576649	1.588378871

Pyri_6_39	[n+]1(cc(O)ccc1)COCCCCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-26.26669914	-0.687223315
Imid_54_37	C[N+]1=CN(CCCC)C=C1.CC2=CC=C(S(=O)([O-])=O)C=C2	-4.095410008	-35.32126446
Pyrr_2_37	SCC(N)C([O-])=O.C[N+]1(CC)CCCC1	-1.863569173	41.91264498
Pyri_28_37	OC(C1=CC=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCC)=O)=O)=C1)=O.[Br-]	-22.81041669	-15.51797437
Pyri_24_36	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCC)=O)C[N+]2=CC=CC=C2.[Br-]	-23.43356326	-10.65127558
Chol_37_36	OCC[N+](C)(C)CC(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCC)=O)=O.[Br-]	-7.976210191	13.90056549
Pyri_52_35	CC[N+]1=CC=CC=C1.[Br-]	-39.12213142	3.1033291
NI_10_35	O=C(N[C@H](C(OCC)=O)CC1=CC=CC=C1)CN2C=NC=C2	4.072615644	12.76585949
Imid_29_34	N#C\N=C=[N-].O=C(OCCC)Cn1cc[n+](C)c1	9.38175245	-25.92303897
Imid_33_33	[Br-].O=C(OCCCCC)Cn1cc[n+](C)c1C	6.826574592	-25.4769798
Imid_66_33	CCCCn1cc[n+](C)c1.[O-]C(=O)CN	0.373180727	-32.19765325
Pyri_30_33	CN(C)C1=CC=[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCC)=O)=O)C=C1.[Br-]	-23.74233692	-14.81449945
Imid_64_32	CCCCn1cc[n+](C)c1.[O-]C(=O)C1CCCN1	-0.95912671	-29.79150857
Pyri_9_32	[n+]1(cc(O)ccc1)COCCCCCCCCCCCCCCCCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-20.17415067	-0.09681773
Pyri_13_32	[n+]1(cc(O)ccc1)COCCCCCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-23.29665397	2.12478414
Pyri_27_32	CCOC(C1=C[N+](CC(N[C@H](C(OCC)=O)CC2=CC=CC=C2)=O)=CC=C1)=O.[Br-]	-22.56753043	-16.73452815
Pyri_12_31	[n+]1(cc(O)ccc1)COCCCCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-23.68902102	2.30685113
Pyri_44_31	CCOC([C@H](CC1=CC=C(O)C=C1)NC(C[N+]2=CC=CC(C(OCC)=O)=C2)=O)=O.[Br-]	-21.93045312	-16.9579751
Pyri_50_31	CCOC(C1=CC=C[N+](C(C(OCCCC)=O)C2=CC=CC=C2)=C1)=O.[Br-]	-16.77039386	-22.00275238
Imid_69_30	CCCCn1cc[n+](C)c1.[O-]C(=O)C(N)Cc1ccccc1	-3.611413409	-30.68789222
Quat_19_30	[O-]C(=O)[C@@H]1CCCN1.CCC[N+](CCC)(CCC)CCC	-22.05205164	29.81904446
Imid_32_29	[Br-].O=C(OCCCCC)Cn1cc[n+](C)c1	6.763455456	-26.67303912
Imid_73_29	N#CCn1cc[n+](CCOCC)c1.N#C[N-]C#N	13.4841017	-27.18036423
Chol_36_29	O=C([C@H](CC1=CC=CC=C1)NC(C[N+](C)(CCO)C)=O)OCCCCCC.[Br-]	-7.651334015	14.82744608
Piperi_1_28	C[N+]1(CC)CCCCC1.[O-]C(=O)C(N)CS	-3.227221466	42.1706276
Chol_42_28	OCC[N+](C)(C)CC(N[C@H](C(OCCCCCCCCOC([C@H](CC1=CC=CC=C1)NC(C[N+](C)(C)CCO)=O)=O)=O)CC2=CC=CC=C2)=O.[Br-].[Br-]	-35.53840283	-16.82827561
Quat_18_28	[O-]C(=O)[C@@H]1CCCN1.CC[N+](CC)(CC)CC	-4.770752343	40.54262008
Imid_21_27	[Br-].O=C(OCCCCCCCC)Cn1cc[n+](C)c1	5.219057246	-26.30973151
Imid_81_27	CN1C=C[N+](C(C(OCCCC)=O)C2=CC=CC=C2)=C1.[Br-]	-11.25420203	-25.16620156
Chol_38_27	OCC[N+](C)(CC(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCC)=O)=O)C.[Br-]	-8.183564962	13.1823941
NI_13_27	CCOC([C@@H]1CCCN1CC(N[C@H](C(OCC)=O)CC2=CC=CC=C2)=O)=O	2.70639416	14.27879592
Imid_19_26	[Br-].O=C(OCCCC)Cn1cc[n+](C)c1	7.3808758	-26.89103285
Imid_40_26	CCCCCCCCCCN1C=C[N+](=C1)C.[Br-]	2.26724639	-28.57170362
Imid_13_25	CCCCCCCCOS(=O)(=O)[O-].CCCCN1C=C[N+](=C1)C	-5.050833412	-10.63959349

Imid_30_25	FC(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F.O=C(OCCC)Cn1cc[n+](C)c1	18.92473298	-23.91726635
Imid_68_25	CCCCn1cc[n+](C)c1.[O-]C(=O)C(N)Cc1c[NH]cn1	-3.098499826	-29.61180525
Pyri_19_25	[Cl-].Oc1ccc[n+](COCCCCCCCCCCCCCCCCCC)c1	-19.50277223	-0.710364989
Pyri_67_25	[Br-].O\N=C\c1cc[n+](CC(=O)NC(Cc2ccccc2)C(=O)NCCCCCCCC)cc1	-25.96267011	-9.3615767
Imid_20_24	[Br-].O=C(OCCCCCC)Cn1cc[n+](C)c1	6.111804665	-26.4103544
Pyri_4_24	[n+]1(cc(O)ccc1)COCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-27.1276826	-0.616881688
Pyri_65_24	[Br-].O\N=C\c1cccc[n+]1CC(=O)NC(Cc1ccccc1)C(=O)NCCCCCCCC	-25.74966581	-9.99263294
Chol_34_24	OCC[N+](C)(C)CC(N[C@@H](CC1=CC=CC=C1)C(OCC)=O)=O.[Br-]	-6.633957165	16.75225511
Chol_40_24	OCC[N+](C)(C)CC(N[C@H](C(OCCCCCCCCCCCOC([C@H](CC1=CC=CC=C1)NC(C[N+](C)(C)CCO)=O)=O)=O)CC2=CC=CC=C2)=O.[Br-].[Br-]	-35.70049085	-16.99681356
Imid_31_23	[Br-].O=C(OCCC)Cn1cc[n+](C)c1C	7.689550567	-25.71839355
Imid_74_23	N#CCn1cc[n+](CCOCC)c1.FC(F)(F)C([O-])=O	13.52647003	-27.1659266
Prol_1_23	O=C(C[N@+]1(C)CCC[C@@]1([H])C(OCC)=O)N[C@@H](CC2=CC=CC=C2)C(OCC)=O.[Br-]	-8.851479876	17.66663813
Pyri_5_22	[n+]1(cc(O)ccc1)COCCCC.[N-]1C(=O)C=C(C)OS1(=O)=O	-26.8716079	-0.810481612
Pyri_49_22	COC1=CC=C[N+](C(C(OCCCC)=O)C2=CC=CC=C2)=C1.[Br-]	-16.10014585	-22.44053131
Quat_22_22	[O-]C(=O)[C@H]1C[C@H](O)CN1.CC[N+](CC)(CC)CC	-4.891199558	40.48030188
NI_12_22	CCOC(=O)[C@H](CC1=CC=CC=C1)NC(=O)CN1CCOCC1	3.286583239	13.58462283
Imid_18_21	[Br-].O=C(OCC)Cn1cc[n+](C)c1	8.389375289	-27.43641929
Imid_25_21	[Br-].O=C(OCCC)Cn1cc[n+](C)c1	8.065362064	-26.9693913
Imid_83_21	[B-](F)(F)(F)F.CCCCCCCCCCN1C=C[N+](=C1)C	1.851733144	-27.52155556
Pyri_47_21	O=C(OCCCC)C(C1=CC=CC=C1)[N+]2=CC=CC=C2.[Br-]	-15.51920769	-22.32087062
Chol_35_21	O=C([C@H](CC1=CC=CC=C1)NC(C[N+](C)(CCO)C)=O)OCCCC.[Br-]	-7.208918961	15.75423992
Quat_5_21	[N+](CCCC)(CCCC)(CCCC)CCCC.[O-]C(=O)[C@@H]1[C@@H](O)[C@H](O)[C@@H](O)C(O)O1	-23.34562278	24.77984258
Imid_80_20	CN1C=C[N+](C(C2=CC=CC=C2)C(OCC)=O)=C1.[Br-]	-11.34737908	-25.77789546
Pyri_15_20	[n+]1(cc(O)ccc1)COCCCCCCCCCCCCCCCCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-21.02736081	0.382189572
Pyri_66_20	[Br-].O\N=C\c1ccc[n+](CC(=O)NC(Cc2ccccc2)C(=O)NCCCCCCCC)c1	-25.69225247	-9.546902094
Chol_39_20	OCC[N+](C)(CC(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCCCC)=O)=O)C.[Br-]	-8.305392165	12.74220176
Chol_43_20	CCOC([C@H](CC1=CC=C(O)C=C1)NC(C[N+](C)(C)CCO)=O)=O.[Br-]	-6.41603923	16.08653388
Imid_67_19	CCCCn1cc[n+](C)c1.[O-]C(=O)C(N)CS	-1.171277995	-31.57027008
Chol_41_19	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+](C)(CCO)C)=O)=O)=O)C[N+](C)(CCO)C.[Br-].[Br-]	-35.92530682	-17.04818525
Quat_14_19	[O-]C(=O)C(N)CS.CCCC[N+](CCCC)(CCCC)CCCC	-24.5421769	27.39599978
Quat_23_19	[O-]C(=O)[C@@H]1C[C@@H](O)CN1.CCC[N+](CCC)(CCC)CCC	-22.02053121	29.78907384
Prol_2_19	OC([C@]1([H])[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCC)=O)=O)(C)CCC1)=O.[Br-]	-8.031611944	17.75488417
NI_33_19	C(CCO)CO	39.32086504	1.918876973
Imid_63_18	[Br-].COC(=O)Cn1cc[n+](C)c1	8.453229295	-24.2223948

Pyri_35_18	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCCCCOC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-42.36048912	-21.5566495
Quat_13_18	CC[N+](CC)(CC)CC.[O-]C(=O)C(N)CS	-2.888488135	42.00208458
Quat_20_18	[O-]C(=O)[C@@H]1CCCN1.CCCC[N+](CCCC)(CCCC)CCCC	-23.46003535	27.47023114
Imid_7_17	OC(CN1C=C[N+](CC(N[C@H](C(OCC)=O)CC2=CC=CC=C2)=O)=C1)=O.[Br-]	-15.70174478	-11.36439787
Imid_72_17	[O-]C(=O)C(N)CS.C[n+]1ccn(CCCCCC)c1	-1.780875498	-31.29969502
Morph_2_17	C[N+]1(CC(N[C@@H](CC2=CC=CC=C2)C(OCC)=O)=O)CCOCC1.[Br-]	-8.061957896	16.34886423
Quat_24_17	[O-]C(=O)[C@@H]1C[C@@H](O)CN1.CCCC[N+](CCCC)(CCCC)CCCC	-23.04855581	27.25669603
NI_21_17	CCCCCCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	1.87442254	6.287644503
Imid_27_16	F[B-](F)(F)F.O=C(OCCC)Cn1cc[n+](C)c1	5.161745748	-39.94489
Imid_28_16	F[P-](F)(F)(F)(F)F.O=C(OCCC)Cn1cc[n+](C)c1	0.915121093	-44.64956317
Imid_71_16	SCC(N)C([O-])=O.C[n+]1ccn(CC)c1	-0.822332435	-32.44186149
Pyri_46_16	O=C(OCC)C(C1=CC=CC=C1)[N+]2=CC=CC=C2.[Br-]	-14.76951928	-22.74853011
Quat_6_16	[O-]C(=O)[C@H](C)O.CCCC[N+](CCCC)(CCCC)CCCC	-25.20079773	27.84350506
Quat_7_16	[O-]C(=O)[C@@H](O)[C@H](O)C(O)=O.CCCC[N+](CCCC)(CCCC)CCCC	-24.45771342	25.28140203
Quat_29_16	CCCCCCCCCCCCCC[N+](C)(C)CC1=CC=CC=C1.[Cl-]	-5.22046841	6.449501759
NI_32_16	C(CN)N	31.80844271	5.056992436
Imid_6_15	CN1C=C[N+](CC(N[C@@H](CC2=CC=CC=C2)C(OCCCCCCCCCCCC)=O)=O)=C1.[Br-]	-19.01202302	-9.145665749
Prol_3_15	CCOC([C@]1([H])[N+](CC(N[C@@H](CC2=CC=CC=C2)C(O)=O)=O)(C)CCC1)=O.[Br-]	-8.212547527	18.02840648
Imid_53_14	C[N+]1=CN(CCCC)C=C1.O=S(OC)([O-])=O	-0.878756119	-35.76442664
Pyri_73_14	C1=CC=[N+](C=C1)CC(=O)NN.[Cl-]	-38.18002562	1.73793403
Quat_8_14	[O-]C(=O)[C@@H](O)CC(=O)O.CCCC[N+](CCCC)(CCCC)CCCC	-24.66259543	26.05360691
Quat_28_14	CN(C)C(=O)Oc1cccc(c1)[N+](C)(C)C.COS(=O)(=O)[O-]	7.611167453	38.20272429
Pyri_11_13	[n+]1(cc(O)ccc1)COCCCC.[N-]1C(=O)c2ccccc2S1(=O)=O	-23.92854118	2.898001218
Quat_11_12	CC(=O)C([O-])=O.CCCC[N+](CCCC)(CCCC)CCCC	-25.88588899	27.78355438
NI_22_12	CCCCCCCCCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	0.635182561	5.922142868
Phos_2_11	CCCCCCCCCCCCCC[P+](CCCCCC)(CCCCCC)CCCCCC.C(=[N-])=NC#N	-18.04457474	0.62520453
NI_20_11	CCCCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	3.159369023	6.832539934
Quat_10_10	CCCC[N+](CCCC)(CCCC)CCCC.[O-]C(=O)CC(O)=O	-25.50807784	27.00677688
NI_66_10	OC(CN1C=CN=C1)=O	27.8093555	5.412186
Imid_14_9	C[N+]1=CN(CCCC)C=C1.F[B-](F)(F)F	4.105153655	-39.43945656
Imid_44_9	[I-].Cc1c(C)[n+](C)cn1CCCCCC	7.742157349	-34.55209787
Imid_48_9	CCN1C=C[N+](=C1)C.COS(=O)(=O)[O-]	-0.592001963	-36.28707034
Imid_55_9	O=C([O-])C.C[n+]1ccn(CC)c1	2.190320409	-34.62361106
Pyri_64_9	[Br-].O\N=C\c1cc[n+](CC(=O)NC(C)C(=O)NCCCCCCCC)cc1	-26.6376448	-7.976537308

NI_23_9	CCCCCCCCCCCCCCOC(=O)[C@@H](N)CC1=CC=CC=C1	-0.52044435	5.88865173
Pyri_45_8	O=C(OC)C(C1=CC=CC=C1)[N+]2=CC=CC=C2.[Br-]	-14.50632622	-22.87300101
Quat_4_8	[Br-].CCCC[N+](CCCC)(CCCC)CCCC	-26.38855837	28.93680273
Imid_15_7	C[N+]1=CN(CCCC)C=C1.F[P-](F)(F)(F)(F)F	0.430368311	-44.65814006
Imid_41_7	CCN1C(=C([N+](=C1C(C)C)C)C)C.[I-]	7.954015654	-35.82334259
Pyri_3_7	CN(C1=CC=[N+](CC(O)=O)C=C1)C.[Br-]	-37.35228075	-0.18765785
Pyri_41_7	O=C(N[C@@H](CC1=CC=CC=C1)C(NCCOCCOCCNC([C@@H](NC(C[N+]2=CC=CC=C2)=O)CC3=CC=CC=C3)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-45.42708847	-20.58478244
Imid_43_6	[I-].Cc1c(C)n(CC)c[n+]1C	7.289575961	-36.03880536
Imid_57_6	CN1C=CN(CC=C)[CH+]1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	25.6157687	-26.32774776
Quat_45_6	C[N+]1([C@@H]2C[C@H](C[C@H]1[C@H]3[C@@H]2O3)OC(=O)[C@H](CO)c4ccccc4)C.[Br-]	-5.896156388	18.44504459
Imid_12_5	C[N+]1=CN(CCCC)C=C1.[Br-]	2.65464434	-33.22758187
Imid_17_5	CCCCN1C=C[N+](=C1)C.C(F)(F)(F)S(=O)(=O)[N-]S(=O)(=O)C(F)(F)F	18.93972802	-23.68362767
Imid_22_5	[Br-].O=C(ONCCCC)Cn1cc[n+](C)c1	7.40977509	-28.83942332
Imid_23_5	[Br-].CCCCN(C)OC(=O)Cn1cc[n+](C)c1	8.470529195	-29.75035788
Imid_24_5	[Br-].CCN(OC(=O)Cn1cc[n+](C)c1)CC	8.565056296	-29.90217676
Imid_38_5	[N+]1(C)=CN(C=C1)CCCCCC.[Br-]	2.470957238	-30.92184435
Imid_79_5	CN1C=C[N+](C(C(OC)=O)C2=CC=CC=C2)=C1.[Br-]	-11.41154853	-25.98320699
Pyri_2_5	COC1=CC=C[N+](CC(O)=O)=C1.[Br-]	-38.8336287	0.05104749
Pyri_40_5	O=C(N[C@@H](CC1=CC=CC=C1)C(NCCCCNC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-45.67316044	-21.7547582
Pyri_72_5	CC[N+]1=CC=CC=C1C.[Br-]	-39.44051482	3.920061398
Chol_33_5	OCC[N+](C)(C)CC(N[C@@H](CC1=CC=CC=C1)C(O)=O)=O.[Br-]	-5.87296979	16.9520367
Pyri_75_5	C[n+]1cccc(c1)[C@@H]2CCCN2C.[I-]	-37.29868824	5.60233308
NI_74_5	CCCCN(C=N1)C2=C1C=CC=C2	26.84228748	4.953543004
Imid_39_4	CCN1C=C[N+](=C1)C.[Br-]	3.054699757	-34.65737554
Pyri_25_4	O=C(N[C@@H](CC1=CC=CC=C1)C(OCCCCCCCCCCCCC)=O)C[N+]2=CC=CC=C2.[Br-]	-23.198392	-10.19619441
Pyri_37_4	O=C(N[C@@H](CC1=CC=CC=C1)C(NCCCCCCCNC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-45.14805128	-21.61590389
Pyri_63_4	[Br-].O\N=C\c1cc[n+](CC(=O)NCCCCCCCC)cc1	-26.86593888	-7.578806929
Quat_3_4	[OH-].CCCC[N+](CCCC)(CCCC)CCCC	-26.0973948	29.16134386
Imid_16_3	C[N+]1=CN(CCCC)C=C1.N#CN=C=[N-]	5.098085378	-32.25081259
Imid_59_3	CN1C=CN(CCCCCC)[CH+]1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	25.77987848	-25.4908823
Imid_87_3	CCCCCCCCN1C=C[N+](=C1)C.[Br-]	2.420117886	-29.48328171
Pyri_51_3	CC[N+]1=CC=C(C=C1)C.[Br-]	-38.69696827	3.550510133
Pyri_54_3	CCCC[N+]1=CC=C(C)C=C1.[Cl-]	-37.50907857	4.148742377
Pyri_68_3	C1=CC=[N+](C=C1)CC(=O)N.[Cl-]	-38.63850308	1.450906346

Quat_1_3	[Br-].C[N+](C)(C)C	31.26396589	1.964185586
NI_35_3	C(CO)NCCO	44.91312804	7.155989225
Imid_42_2	[I-].C[n+]1c(C)c(C)n(CC)c1C	7.645882901	-36.14898815
Imid_60_2	CN1C=CN(CCO)[CH+]1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	25.38458063	-25.81991598
Imid_88_2	CCCCCCCCN1C=C[N+](=C1)C.[Cl-]	2.393562255	-29.4686076
Pyrr_3_2	CCCC[N+]1(CCCC1)C.[Br-]	0.633045288	46.61144737
Pyri_53_2	CCCC[N+]1=CC=CC=C1.[Br-]	-37.49149407	3.558347648
Thia_2_2	Cc1c(sc[n+]1Cc2cnc(nc2N)C)CCO.Cl.[Cl-]	-11.322466	-35.07132645
Imid_37_1	CCN1C=C[N+](=C1)C.[Cl-]	3.064627483	-34.66972543
Pyri_38_1	O=C(N[C@@H](CC1=CC=CC=C1)C(NCCCCCCCCCNC([C@H](CC2=CC=CC=C2)NC(C[N+]3=CC=CC=C3)=O)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-44.78566771	-22.01736131
Pyri_48_1	O=C(NCCCC)C(C1=CC=CC=C1)[N+]2=CC=CC=C2.[Br-]	-14.7165468	-21.85081744
Pyri_57_1	C[n+]1cccc(c1)OC(=O)N(C)C.[Br-]	-40.3121591	-0.195310218
QUAT_44_1	[Cl-].O[C@H](c1ccc(O)cc1)[C@H](C)[NH2+]CCc1ccc(O)cc1	8.676929266	10.00511236
NI_1_1	c1cc(ccc1C(=O)O)NC(=N)N.Cl	19.71055185	2.611205353
NI_2_1	C1CN2CCN1CC2	39.17510128	14.49261952
NI_75_1	C1CNC(=O)CN1	44.83336269	4.900489387
QUAT_39_1	CCCCCCCCCCCCCC[N+](C)(C)CC1=CC=CC=C1.[Cl-]	-5.228526507	6.455050801
Imid_61_0	C[N+]1=CN(CC)C=C1.O=S(C(F)(F)F)([O-])=O	0.138577619	-36.59875877
Imid_45_0	[B-](F)(F)(F)F.CCN1C=C[N+](=C1)C	4.016217932	-39.3697372
Imid_46_0	CCN1C=C[N+](=C1)C.COP(=O)([O-])OC	12.38051721	-34.30187343
Imid_47_0	CCN1C=C[N+](=C1)C.CCOP(=O)([O-])OCC	12.39375976	-34.28922361
Imid_49_0	CCN1C=C[N+](=C1)C.CCOS(=O)(=O)[O-]	-1.191787602	-36.48579383
Imid_50_0	CCN1C=C[N+](=C1)C.C(=[N-])=NC#N	5.102385036	-32.4300142
Imid_52_0	CCN1C=C[N+](=C1)C.F[P-](F)(F)(F)(F)F	0.338384676	-44.82632905
Imid_56_0	OC(=O)C[n+]1ccn(C)c1.[Br-]	2.557453682	-36.1927116
Imid_62_0	CCC[N+]1=CN(C)C=C1.[I-]	3.310197582	-35.83920519
Imid_82_0	CN1C=C[N+](C(C(NCCCC)=O)C2=CC=CC=C2)=C1.[Br-]	-10.67128868	-25.05366756
Imid_86_0	C[N+]1=CN(C=C1)CCO.[Cl-]	3.778800971	-34.67628574
Morph_1_0	C[N+]1(CCCC)CCOCC1.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	26.49163796	-25.42760241
Pyri_39_0	O=C(N[C@@H](CC1=CC=CC=C1)C(NCCCCCCCCCCNC([C@@H](NC(C[N+]2=CC=CC=C2)=O)CC3=CC=CC=C3)=O)=O)C[N+]4=CC=CC=C4.[Br-].[Br-]	-44.40229853	-22.33068803
Pyri_58_0	C[n+]1ccccc1/C=N/O.[Cl-]	-40.94743395	3.679844658
Pyri_71_0	CCCC[N+]1=CC=CC(=C1)C.[Br-]	-36.89739245	4.079049046
Quat_2_0	[Br-].CC[N+](CC)(CC)CC	0.862061895	46.35217058
Quat_9_0	[O-]C(=O)CCC(=O)O.CCCC[N+](CCCC)(CCCC)CCCC	-25.67603807	26.18594175

Quat_12_0	CCCC[N+](CCCC)(CCCC)CCCC.OS(=O)(=O)[O-]	-25.25378522	28.83682995
Thia_1_0	CC[N+]1=CSC(=C1C)CCO.[Br-]	-11.32579914	-35.08336367
NI_25_0	CN(CCO)CCO	40.60592914	15.94572495
NI_26_0	CCN(CCO)CCO	40.28220312	16.22604739
NI_27_0	C(CO)N(CCO)CCO	40.35008266	16.86662315
Pipera_1_0	O=C(O)C1C[NH2+]CC[NH2+]1.[Cl-].[Cl-]	47.07318770	6.333133763
Pyri_76_0	C[n+]1cc(C2CCC[N+](C)2C)ccc1.[I].[I]	-35.68529422	4.388572054
NI_46_0	C1=CC=C2C(=C1)C(=O)NS2(=O)=O	20.61541965	3.762897433
Anion_1_0	CC1=CC(=O)[N-]S(=O)(=O)O1.[K+]	36.90049655	5.18098124
QUAT_42_0	[Cl-].CC[NH2+]CC	42.83775872	6.519943859
NI_67_0	C1=CN=CC=C1C(=O)NN	22.56985305	1.021392412
NI_71_0	C1CNCCN1	46.6378824	6.658716502
NI_72_0	O=C1NCCOC1	44.17138445	4.119381808
Piperi_3_0	[Cl-].[Cl-].[NH3+]C1CCC[NH2+]C1	41.49728957	8.716321519
NI_76_0	C1CC(CNC1)CN	43.80140011	9.443238955
NI_77_0	C1CCNC(C1)CN	43.83108976	10.25770877
NI_80_0	C(=O)(C(F)(F)F)O	30.03701668	5.427030798
Piperi_4_0	[Cl-].OC1CC[NH2+]CC1	44.09238591	7.28685604
Piperi_5_0	[Cl-].OC1(O)CC[NH2+]CC1	43.98849926	7.088442595
Imid_58_0	CN1C=CN(CCCC)[C+]1C.FC(F)(F)C(C(F)(F)F)O[Al-](OC(C(F)(F)F)C(F)(F)F)(OC(C(F)(F)F)C(F)(F)F)OC(C(F)(F)F)C(F)(F)F	26.21137907	-26.19305015

References

1 D. Rogers and M. Hahn, *J. Chem. Inf. Model.*, 2010, **50**, 742–754.

2 S. Chakravarti, *Chem. Res. Toxicol.*, 2023, **36**, 959–970.

3 S. K. Chakravarti, *ACS omega*, 2018, **3**, 2825–2836.

4 J. Friedman, T. Hastie and R. Tibshirani, *J. Stat. Soft.*, 2010, **33**.

5 R. Tibshirani, J. Bien, J. Friedman, T. Hastie, N. Simon, J. Taylor and R. J. Tibshirani, *J. R. Statist. Soc. B*, 2012, **74**, 245–266.

6 OECD, *Guidance Document on the Validation of (Quantitative)Structure-Activity Relationships [(Q)SAR] Models*, 2007.

7 P. Gramatica, *QSAR Comb. Sci.*, 2007, **26**, 694–701.

8 P. Gramatica, in *Computational Toxicology. Methods in Molecular Biology*, ed. B. Reisfeld and A. Mayeno, Humana Press, Totowa, NJ, vol 930, pp. 499–526.