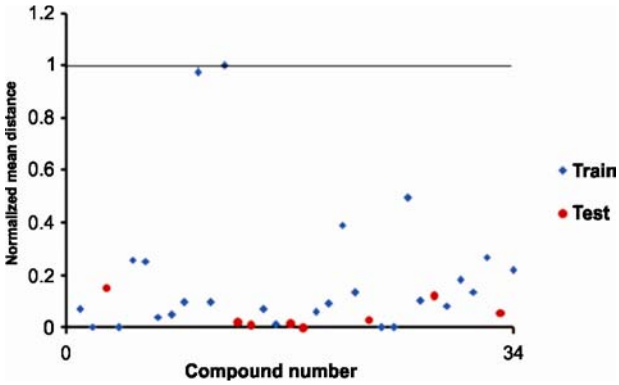
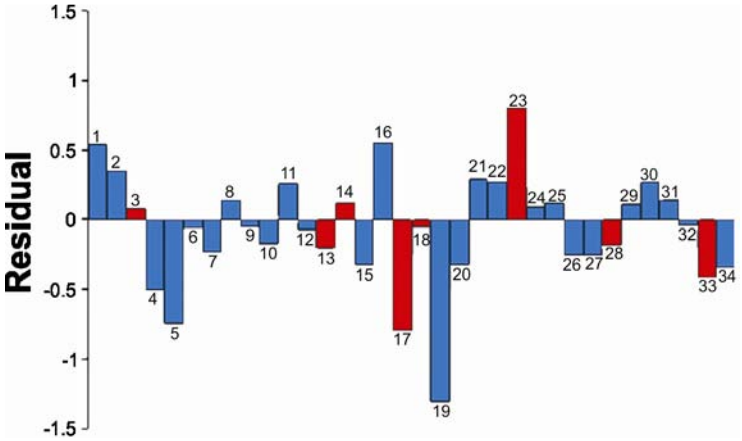
Supplementary Figures



Supplementary Fig 1—Test of the applicability domain by the Euclidean distance approach for the model (1)



Supplementary Fig 2—Residual plot between predicted and experimental values of train set (light blue box) and test set (red box) compounds for the model (1)

Supplementary Tables

Table S1—Correlation matrix among biological activity and selected parameters of model (1)

	T_N_O_7	***GATS2c***	***ZMIC3***	***nHsNH2***	***Quadrupole1***	***pEC_{50}***
T_N_O_7	1.00	-0.40	0.10	0.19	-0.62	-0.56
GATS2c		1.00	-0.02	-0.13	-0.07	-0.22
ZMIC3			1.00	-0.31	-0.08	-0.49
nHsNH2				1.00	-0.09	-0.21
Quadrupole1					1.00	0.32
pEC_{50}						1.00

Table S2—VIF, ***t***-Value and ***p***-Value for model (1)

Parameter	***t***-Value	***p***-Value	VIF
Intercept	13.79568	0.000000	-
T_N_O_7	-8.43747	0.000000	2.41866
GATS2c	-7.65670	0.000000	1.48171
ZMIC3	-6.83043	0.000001	1.13627
nHsNH2	-4.01176	0.000684	1.16793
Quadrupole1	-3.79572	0.001134	1.99387

Table S3—Calculation of the Euclidean based applicability domain for model (1)

Compound No	Distance Score	Mean Distance	Normalized Mean Distance
1	429.5563	16.5214	0.071425
2	382.9112	14.72735	0.003017
3	484.9993	18.65382	0.152734
4	382.4617	14.71007	0.002358
5	557.2156	21.43137	0.258643
6	554.0536	21.30975	0.254006
7	408.459	15.70996	0.040485
8	414.9476	15.95952	0.05
9	446.5726	17.17587	0.09638
10	1045.537	40.21297	0.97479
11	447.8998	17.22691	0.098326
12	1062.727	40.87413	1.000
13	396.4538	15.24822	0.022878
14	388.2509	14.93273	0.010848
15	427.7671	16.45258	0.068801
16	388.3947	14.93826	0.011059
17	392.2101	15.085	0.016655
18	382.0019	14.69238	0.001684
19	421.0173	16.19297	0.058902
20	441.9855	16.99944	0.089653
21	648.6415	24.94775	0.392723
22	471.1188	18.11995	0.132378
23	400.9712	15.42197	0.029503
24	380.8537	14.64822	0.000
25	381.5256	14.67406	0.000985
26	719.7892	27.6842	0.497065
27	449.7418	17.29776	0.101028
28	463.7109	17.83504	0.121514
29	437.1747	16.81441	0.082597
30	504.9787	19.42226	0.182035
31	473.5374	18.21298	0.135925
32	565.0006	21.73079	0.27006
33	419.0944	16.11901	0.056082
34	530.9782	20.42224	0.220165

Table S4—Observed, LOO-predicted and Residual activities of training and test set compounds for model (1)

Compound Number	Observed	LOO-Predicted	Residual
1	4.6	4.057	0.542833
2	6.4	6.052	0.34777
3	4.5	4.419954	0.080046
4	3.9	4.403007	-0.50301
5	3.8	4.544291	-0.74429
6	4.7	4.761493	-0.06149
7	4.3	4.534136	-0.23414
8	4.1	3.960966	0.139034
9	3.8	3.849624	-0.04962
10	4	4.17154	-0.17154
11	4	3.734579	0.265421
12	3.7	3.77314	-0.07314
13	3.9	4.100774	-0.20077
14	5.7	5.583816	0.116184
15	6.1	6.425905	-0.32591
16	6.3	5.742119	0.557881
17	4.4	5.191409	-0.79141
18	3.7	3.749056	-0.04906
19	3.7	5.006716	-1.30672
20	4.5	4.825935	-0.32594
21	3.6	3.308884	0.291116
22	3.9	3.625019	0.274981
23	4	3.193811	0.806189
24	4.7	4.605927	0.094073
25	4.5	4.381638	0.118362
26	4.4	4.655493	-0.25549
27	5.1	5.350799	-0.2508
28	5.1	5.277049	-0.17705
29	5.4	5.286454	0.113546
30	5.3	5.024066	0.275934
31	4.5	4.358512	0.141488
32	4.4	4.445156	-0.04516
33	4.2	4.609664	-0.40966
34	3.9	4.243503	-0.3435