SUPPLEMENTARY INFORMATION

Correlation of blood-brain penetration and human serum albumin binding with theoretical descriptors

Mati Karelson, [*a,b,c] Dimitar Dobchev,[a,c] Tarmo Tamm,[c,d] Indrek Tulp,[b,c] Jaak Jänes,[b,c] Kaido Tämm,[b,c] Andre Lomaka,[a,c] Deniss Savchenko,[a,c] Gunnar Karelson[a,c]

[a] Department of Chemistry, Tallinn University of Technology, Akadeemia tee 15, Tallinn 19086, Estonia

[b] Institute of Chemistry, University of Tartu, Jakobi Street 2, Tartu 51014, Estonia

[c] MolCode, Ltd.; Soola 8, Tartu 51013, Estonia

[d] Institute of Technology, University of Tartu, Nooruse 1, Tartu 50411, Estonia

E-mail: mati.karelson@ttu.ee

Table 1. SM1 – Descriptor values available for all models.

Table 2. SM2 – smile strings available for all compounds in Table 2.

Table 1. SM1 – Descriptor values available for all models.

CAS	Number of halogenide groups	SQRT(D0000368)	Max net atomic charge (Zefirov) for N atoms	Max electrophilic reactivity index (AM1) for H atoms
21672	0	1.64806	-0.0660815	0.000298049
51481-61-9	0	1.06444	-0.0431017	0.00138185
53-86-1	1	0.990184	-0.0912991	0.000599535
25614-03-3	1	1.20514	-0.0686689	0.000710489
69-72-7	0	0.859889	0	1.70365E-10
18559-94-9	0	1.12439	-0.119749	0.000969869
29122-68-7	0	1.08553	-0.101856	0.000358671
59-92-7	0	1.30951	-0.113841	0.000274397
52-53-9	0	0.826758	-0.0610633	9.06479E-05
154-93-8	2	1.0937	0.00833836	9.05455E-05
50-33-9	0	0.632498	-0.0723909	0.00221258
50-78-2	0	0.94083	0	7.20458E-05
56-54-2	0	0.927621	-0.0936874	0.000419356
487-54-7	0	1.13592	-0.0913271	4.67154E-05
5588-33-0	0	0.664966	-0.11013	6.70589E-05
36507-30-9	0	0.97822	-0.0863092	0.000504752
103-90-2	0	0.890426	-0.0979484	0.000306876
58-55-9	0	1.00562	-0.0741758	0.00022649
83-67-0	0	1.02093	-0.070262	0.000855648
15687-27-1	0	0.675953	0	0.000640915
57-27-2	0	1.02177	-0.121726	0.000460271
298-46-4	0	0.801996	-0.0883526	0.000104167
76-75-5	0	0.879817	-0.0831418	3.52063E-05
321-64-2	0	0.627366	-0.0932322	0.00046395
60-80-0	0	0.627302	-0.0794389	0.00180815
125-33-7	0	0.973415	-0.0939825	5.31485E-05
51-55-8	0	0.886746	-0.120971	0.000098448
21399	0	0.929704	-0.0745418	0.000850286

57-41-0	0	0.767769	-0.0721814	0.000414023
129618-40-2	0	0.892763	-0.0858496	6.95541E-05
58-15-1	0	0.659953	-0.0773455	0.00186094
28312	0	0.709058	-0.122641	0.000344293
28981-97-7	1	0.801851	-0.0554257	0.000257571
57-43-2	0	0.936793	-0.0714901	3.14691E-05
1622-62-4	1	0.832543	0.00915978	8.81322E-05
57-47-6	0	0.807845	-0.0901193	0.000280793
56-29-1	0	0.895461	-0.071686	0.000050732
76-74-4	0	0.933558	-0.0714425	0.000093008
14759-06-9	0	0.739127	-0.109843	8.83618E-05
22316-47-8	1	0.813297	-0.0892381	7.60053E-05
22316-55-8	1	0.771249	-0.0891474	0.000273822
68-88-2	1	1.04773	-0.118014	0.00015774
91-84-9	0	0.802615	-0.0906999	0.000169705
1088-11-5	1	0.771359	-0.08959	0.000424779
439-14-5	1	0.759283	-0.0896978	0.000478742
76-57-3	0	0.833238	-0.121732	0.000521057
604-75-1	1	0.971701	-0.0754405	0.000321058
525-66-6	0	0.827957	-0.119902	9.34373E-05
28399	0	0.256012	-0.121034	1.44458E-05
28911-01-5	2	0.802171	-0.0553543	0.000259217
50-48-6	0	0.291507	-0.125423	0.000831146
2095-95-6	0	0.665315	-0.113703	0.000241021
50-53-3	1	0.57912	-0.109574	1.00746E-05
50-49-7	0	0.347237	-0.113709	0.000582396
50-47-5	0	0.515445	-0.113706	0.000237639
58-40-2	0	0.509126	-0.111091	5.6723E-06
52-86-8	2	0.822929	-0.121425	0.000365981
10457-90-6	2	0.809988	-0.121432	0.000370275
117-89-5	3	0.782842	-0.109204	1.39253E-05
69-23-8	3	1.04185	-0.109203	0.000013949

Table 2. SM2 – smile strings available for all compounds in Table 2.

CAS number	BBB
76-74-4	C([C@@H](C)(C1(CC)C(=O)NC(NC1=O)=O))CC
59-92-7	C([C@H](C(=O)O)(N))c1cc(O)c(cc1)O
69-72-7	C(=O)(O)c1c(O)cccc1
50-78-2	C(=O)(O)c1c(OC(C)=O)cccc1
125-33-7	C(C)C1(C(=O)NCNC1=O)c1ccccc1
5588-33-0	C(C[C@H]1(N(C)CCCC1))N1c2c(Sc3c1cccc3)ccc(c2)[S@@](C)=O
57-43-2	C(CC(C)C)C1(CC)C(=O)NC(NC1=O)=O
50-49-7	C(CCN(C)C)N1c2c(CCc3c1cccc3)cccc2
50-53-3	C(CCN(C)C)N1c2c(Sc3c1cccc3)ccc(c2)Cl
58-40-2	C(CCN(C)C)N1c2c(Sc3c1cccc3)cccc2
10457-90-6	C(CCN1CCC(CC1)(c1ccc(cc1)Br)O)C(c1ccc(cc1)F)=O
56-29-1	C[C@@]1(C(=O)N(C)C(NC1=O)=O)C=1CCCCC1
57-47-6	C[C@]12c3c(N([C@H]1(N(CC2)C))C)ccc(c3)OC(NC)=O
50-48-6	C1(/c2c(CCc3c1cccc3)cccc2)=C\CCN(C)C
52-53-9	c1([C@](CCCN(CCc2cc(c(cc2)OC)OC)C)(C(C)C)C#N)cc(c(cc1)OC)OC
15687-27-1	c1([C@H](C(O)=O)(C))ccc(cc1)CC(C)C
51481-61-9	c1([nH]cnc1CSCCN\C(=N\C#N)NC)C
53-86-1	c1(c(C)n(c2c1cc(cc2)OC)C(c1ccc(cc1)Cl)=O)CC(O)=O
76-75-5	C1(C(NC(NC1=O)=S)=O)([C@H](CCC)(C))CC
487-54-7	c1(C(NCC(O)=O)=O)c(cccc1)O
77-10-1	C1(c2ccccc2)(N2CCCCC2)CCCCC1
18559-94-9	c1(cc(c(cc1)O)CO)[C@@H](CNC(C)(C)C)(O)
29122-68-7	c1(ccc(cc1)CC(N)=O)OC[C@H](CNC(C)C)(O)
77-07-6	c12[C@@]34[C@H]([C@@H](Cc1ccc(c2)O)(N(CC3)C))(CCCC4)
36507-30-9	c12[C@@H]3([C@@H](O3)(c3c(N(c1cccc2)C(N)=O)cccc3))
57-27-2	c12[C@]34[C@@H]5([C@@H](Cc2ccc(c1O[C@H]3([C@H](C=C5)(O)))O)(N(CC4)C))
76-57-3	c12[C@]34[C@@H]5([C@@H](Cc2ccc(c1O[C@H]3([C@H](C=C5)(O)))OC)(N(CC4)C))
56-54-2	c12c([C@@H]([C@@H]3(N4C[C@@H]([C@H](C3)(CC4))(C=C)))(O))ccnc1ccc(c2)OC
28911-01-5	c12c(C(c3c(cccc3)Cl)=NCc3n1c(nn3)C)cc(cc2)Cl
321-64-2	c12c(c(c3c(n1)CCCC3)N)cccc2

28981-97-7	c12c(C(c3ccccc3)=NCc3n1c(nn3)C)cc(cc2)Cl
1622-62-4	c12C(c3c(cccc3)F)=NCC(N(c1ccc(c2)[N+](=O)[O-])C)=O
525-66-6	c12c(OC[C@H](CNC(C)C)(O))cccc1cccc2
117-89-5	c12N(c3c(Sc1ccc(c2)C(F)(F)F)cccc3)CCCN1CCN(CC1)C
69-23-8	c12N(c3c(Sc1ccc(c2)C(F)(F)F)cccc3)CCCN1CCN(CC1)CCO
14759-06-9	c1c(ccc2Sc3ccccc3N(c12)CC[C@@H]1(N(CCCC1)C))S(=O)(=O)C
51-55-8	C1C[C@H]2(N([C@@H]1(CC(C2)OC([C@@H](c1ccccc1)(CO))=O))C)
604-75-1	Clc1cc2C(=N[C@H](C(Nc2cc1)=O)(O))c1ccccc1
1088-11-5	Clc1cc2C(=NCC(Nc2cc1)=O)c1ccccc1
439-14-5	Clc1ccc2c(c1)C(=NCC(N2C)=O)c1ccccc1
154-93-8	ClCCN(C(NCCCl)=O)N=O
83-67-0	Cn1c2c(c([nH]c1=O)=O)n(cn2)C
60-80-0	Cn1n(c(cc1C)=O)c1ccccc1
103-90-2	N(C(C)=O)c1ccc(cc1)O
91-84-9	N(Cc1ccc(cc1)OC)(CCN(C)C)c1ccccn1
68-88-2	N1([C@@H](c2ccc(cc2)Cl)(c2ccccc2))CCN(CC1)CCOCCO
22316-55-8	N1(C(CC(Nc2c1cc(cc2)Cl)=O)=O)c1ccccc1
129618-40-2	n1(c2c([nH]c(c3c1nccc3)=O)c(ccn2)C)C1CC1
2095-95-6	N1(c2c(CCc3c1cccc3)cccc2)CCCN
50-47-5	N1(c2c(CCc3c1cccc3)cccc2)CCCNC
298-46-4	N1(c2c(cccc2)C=Cc2c1cccc2)C(N)=O
25614-03-3	N12[C@]([C@H]3(N(C([C@@H]2(CC(C)C))=O)CCC3))(O[C@](C1=O)(NC([C@@H]1(C=C2c3c4c(C[C@H]2(N(C1)C))c([nH]c4ccc3)Br))=O)C(C)C)O
59-05-2	Nc1c2c(ncc(n2)CN(C)c2ccc(cc2)C(N[C@H](CCC(=O)O)(C(=O)O))=O)nc(n1)N
57-41-0	O=C1C(NC(N1)=O)(c1ccccc1)c1ccccc1
58-55-9	O=c1c2c(n(c(n1C)=O)C)[nH]cn2
58-08-2	O=c1c2c(n(c(n1C)=O)C)ncn2C
22316-47-8	O=C1N(c2c(N(C(C1)=O)C)ccc(c2)Cl)c1ccccc1
50-33-9	O=C1N(N(C(C1CCCC)=O)c1ccccc1)c1ccccc1
58-15-1	O=c1n(n(c(c1N(C)C)C)C)c1ccccc1
52-86-8	OC1(c2ccc(cc2)Cl)CCN(CC1)CCCC(=O)c1ccc(cc1)F