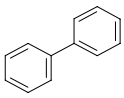
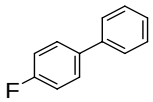
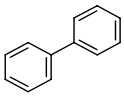
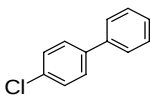
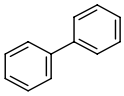
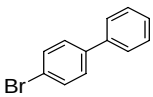
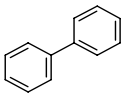
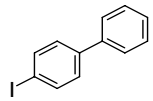
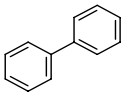
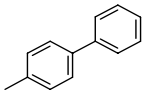
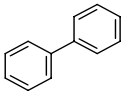
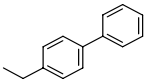
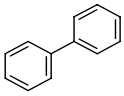
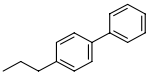
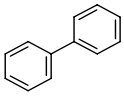
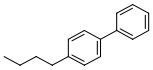
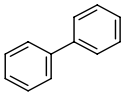
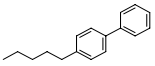
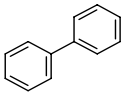
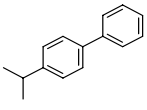
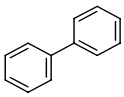
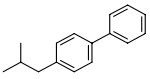
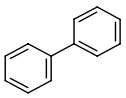
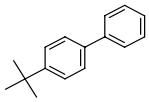
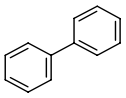
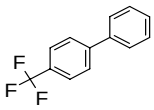
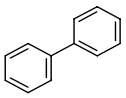
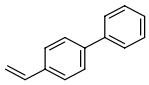
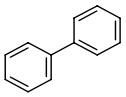
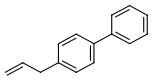
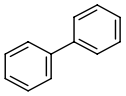
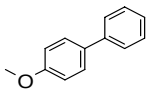
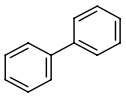
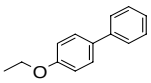
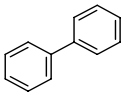
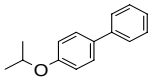
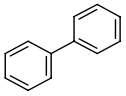
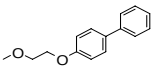
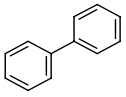
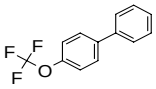
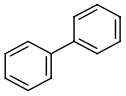
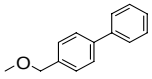
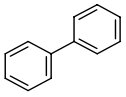
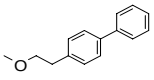
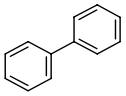
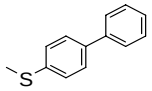
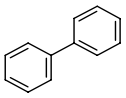
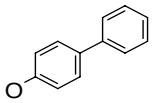
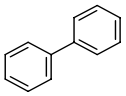
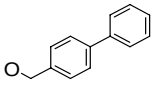
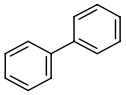
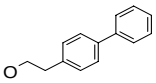
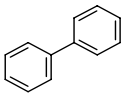
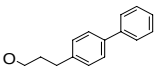
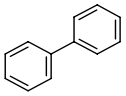
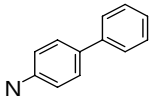
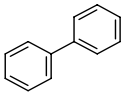
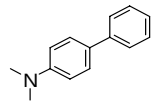
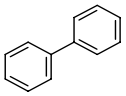
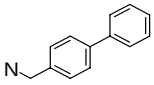
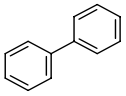
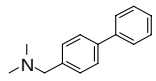
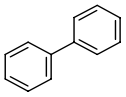
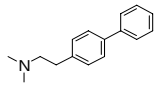
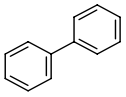
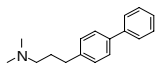
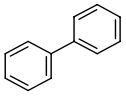
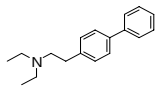
Supporting Information: A Statistical Analysis of the Effects of Common Chemical Substituents on Ligand Potency,

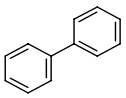
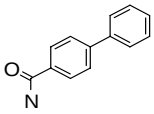
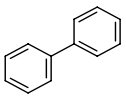
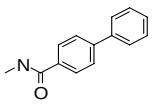
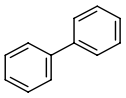
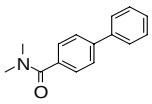
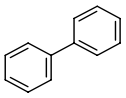
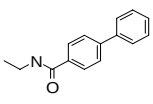
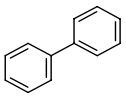
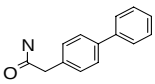
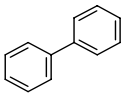
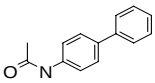
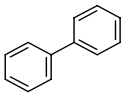
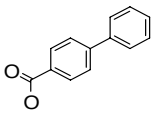
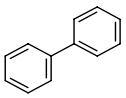
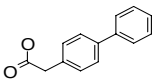
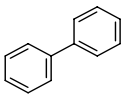
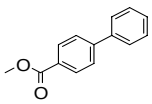
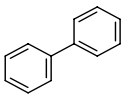
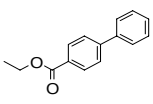
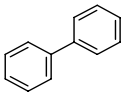
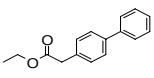
Philip J. Hajduk and Daryl R. Sauer

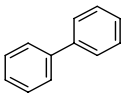
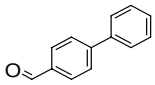
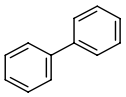
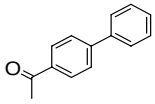
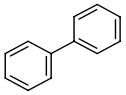
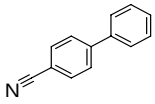
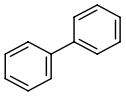
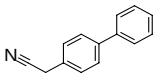
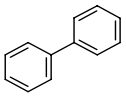
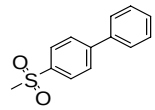
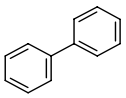
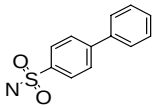
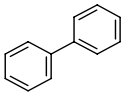
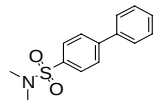
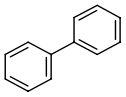
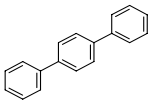
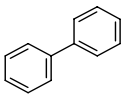
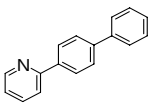
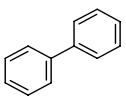
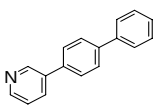
Table S2 Examples of compound modifications described in the text

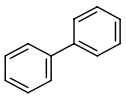
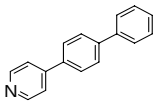
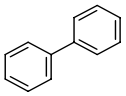
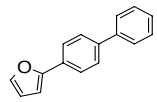
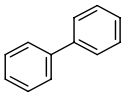
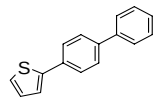
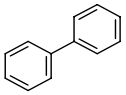
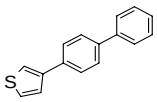
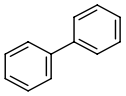
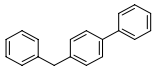
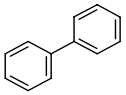
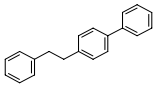
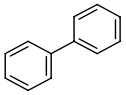
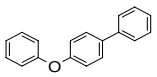
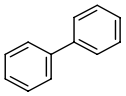
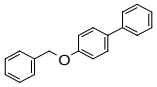
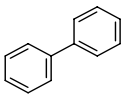
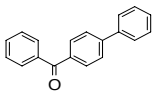
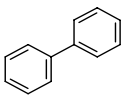
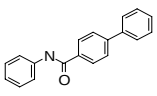
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
1	F			2587	30	0.031	0.066	18.0	0.1	0.0
2	Cl			3885	30	0.063	0.078	34.4	0.7	0.0
3	Br			1048	29	0.115	0.093	78.9	0.9	0.0
4	I			95	21	0.135	0.125	125.9	1.1	0.0
5	C			9867	30	0.053	0.101	14.0	0.5	0.0
6	CC			1425	29	0.077	0.115	28.1	1.0	0.0
7	CCC			503	28	0.074	0.153	42.1	1.6	0.0
8	CCCC			233	26	0.094	0.162	56.1	2.1	0.0
9	CCCCC			73	18	0.081	0.189	70.2	2.6	0.0
10	C(C)C			528	29	0.101	0.144	42.1	1.4	0.0
11	CC(C)C			172	25	0.129	0.208	56.1	2.0	0.0

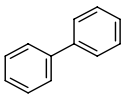
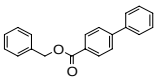
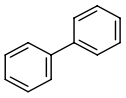
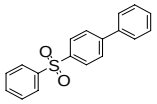
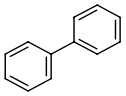
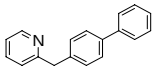
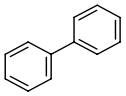
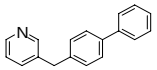
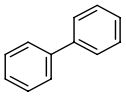
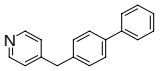
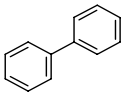
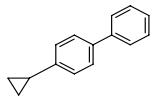
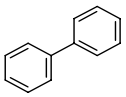
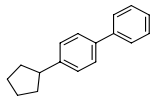
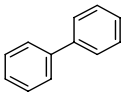
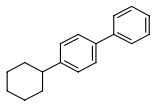
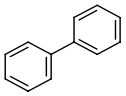
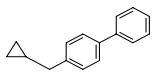
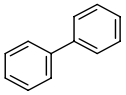
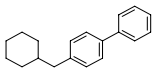
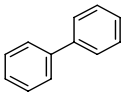
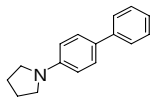
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
12	C(C)(C)C			251	27	0.069	0.234	56.1	1.8	0.0
13	C(F)(F)F			1141	29	0.113	0.134	68.0	0.9	0.0
14	C=C			161	20	0.124	0.124	26.0	0.7	0.0
15	CC=C			84	19	0.107	0.214	40.1	1.1	0.0
16	OC			2941	30	0.040	0.095	30.0	-0.1	9.2
17	OCC			195	27	0.062	0.077	44.1	0.4	9.2
18	OC(C)C			57	16	0.140	0.105	58.1	0.8	9.2
19	OCCOC			50	17	0.060	0.120	74.1	-0.2	18.5
20	OC(F)(F)F			245	27	0.121	0.137	84.0	1.0	9.2
21	COC			221	26	0.044	0.093	44.1	-0.2	9.2
22	CCOC			130	21	0.038	0.135	58.1	-0.1	9.2
23	SC			128	23	0.154	0.069	46.1	0.6	0.0

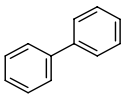
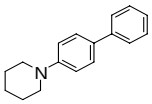
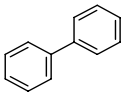
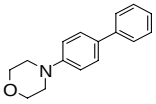
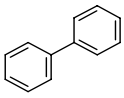
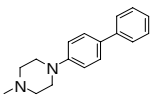
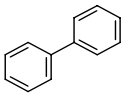
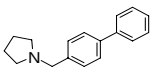
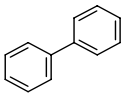
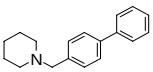
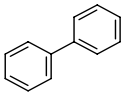
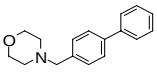
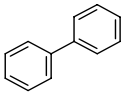
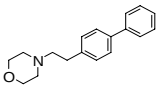
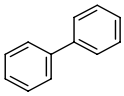
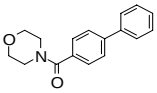
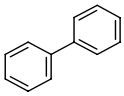
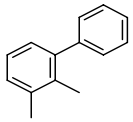
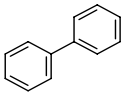
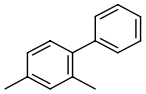
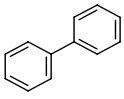
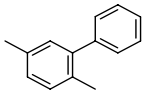
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
24	O			1447	30	0.053	0.144	16.0	-0.7	20.2
25	CO			490	27	0.028	0.124	30.0	-1.0	20.2
26	CCO			211	24	0.042	0.075	44.1	-0.8	20.2
27	CCCO			65	16	0.046	0.062	58.1	-0.4	20.2
28	N			652	27	0.056	0.093	15.0	-1.2	26.0
29	N(C)C			324	29	0.046	0.098	43.1	0.2	3.2
30	CN			77	22	0.051	0.115	29.1	-1.0	26.0
31	CN(C)C			243	25	0.069	0.162	57.1	-0.2	3.2
32	CCN(C)C			215	20	0.129	0.069	71.1	0.0	3.2
33	CCCN(C)C			66	15	0.119	0.060	85.2	0.4	3.2
34	CCN(CC)CC			58	15	0.034	0.103	99.2	1.0	3.2

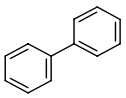
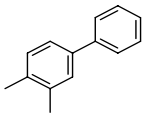
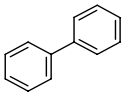
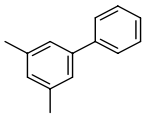
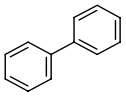
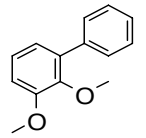
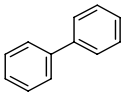
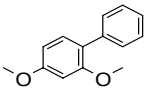
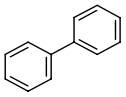
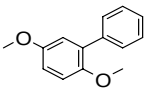
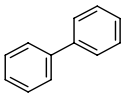
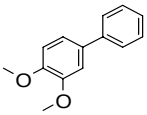
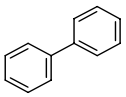
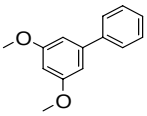
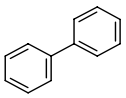
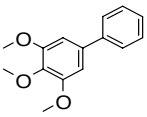
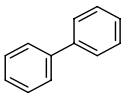
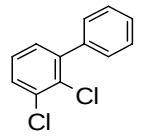
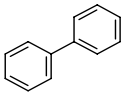
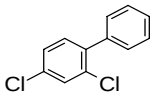
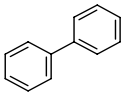
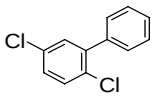
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
35	C(=O)N			305	25	0.069	0.163	43.0	-1.5	43.1
36	C(=O)NC			53	15	0.018	0.273	57.1	-1.3	29.1
37	C(=O)N(C)C			94	19	0.105	0.126	71.1	-1.5	20.3
38	C(=O)NCC			88	17	0.080	0.125	71.1	-0.8	29.1
39	CC(=O)N			75	17	0.026	0.105	57.1	-1.7	43.1
40	NC(=O)C			172	25	0.058	0.179	57.1	-1.0	29.1
41	C(=O)O			498	26	0.056	0.256	44.0	-0.3	37.3
42	CC(=O)O			133	21	0.022	0.172	58.0	-0.7	37.3
43	C(=O)OC			333	27	0.050	0.158	58.0	0.0	26.3
44	C(=O)OCC			193	27	0.026	0.170	72.1	0.5	26.3
45	CC(=O)OCC			55	15	0.073	0.182	86.1	0.2	26.3

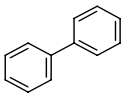
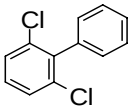
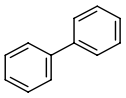
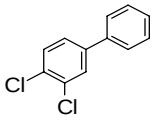
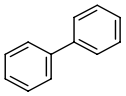
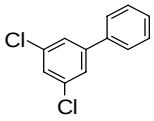
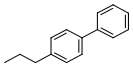
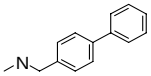
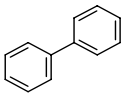
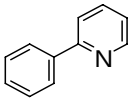
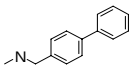
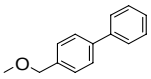
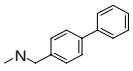
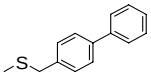
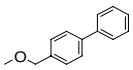
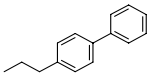
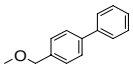
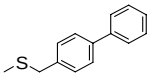
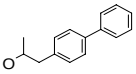
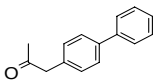
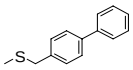
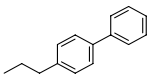
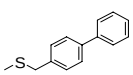
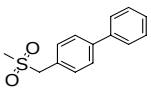
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
46	C=O			58	20	0.068	0.119	28.0	-0.6	17.1
47	C(=O)C			467	29	0.070	0.113	42.0	-0.6	17.1
48	C#N			679	30	0.077	0.104	25.0	-0.6	23.8
49	CC#N			75	20	0.040	0.187	39.0	-0.6	23.8
50	S(=O)(=O)C			277	26	0.130	0.083	78.1	-1.6	34.1
51	S(=O)(=O)N			51	16	0.137	0.039	79.1	-1.8	60.2
52	S(=O)(=O)N(C)C			65	15	0.000	0.308	107.1	-0.8	37.4
53	c1ccccc1			1395	30	0.108	0.169	76.1	1.9	0.0
54	c1ccccn1			107	26	0.083	0.074	77.1	0.6	12.9
55	c1cccnc1			186	23	0.145	0.129	77.1	0.4	12.9

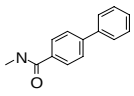
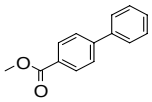
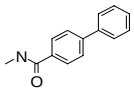
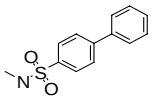
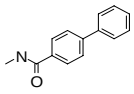
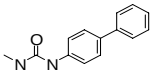
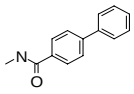
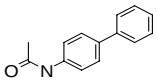
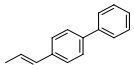
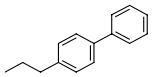
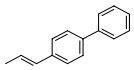
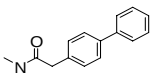
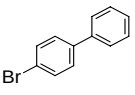
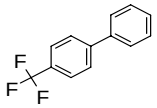
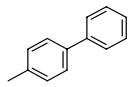
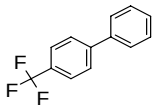
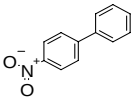
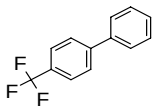
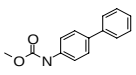
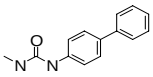
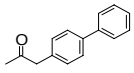
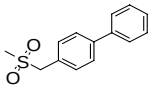
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
56	c1ccncc1			152	27	0.118	0.079	77.1	0.4	12.9
57	c1ccco1			53	16	0.226	0.057	66.1	1.3	13.1
58	c1cccs1			102	23	0.176	0.127	82.1	1.7	0.0
59	c1ccsc1			62	18	0.290	0.194	82.1	1.5	0.0
60	Cc1ccccc1			593	30	0.151	0.143	90.1	2.1	0.0
61	CCc1ccccc1			90	25	0.143	0.110	104.2	2.4	0.0
62	Oc1ccccc1			219	26	0.021	0.187	92.1	2.1	9.2
63	OCc1ccccc1			180	26	0.055	0.257	106.1	1.7	9.2
64	C(=O)c1ccccc1			110	23	0.117	0.126	104.1	1.0	17.1
65	C(=O)Nc1ccccc1			72	23	0.222	0.139	119.1	0.5	29.1

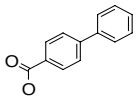
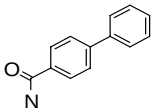
No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
66	C(=O)OCc1ccccc1			93	20	0.043	0.419	134.1	1.8	26.3
67	S(=O)(=O)c1ccccc1			65	18	0.212	0.106	140.2	0.3	34.1
68	Cc1ccccn1			65	20	0.091	0.076	91.1	0.6	12.9
69	Cc1cccnc1			74	18	0.081	0.149	91.1	0.6	12.9
70	Cc1ccncc1			65	20	0.121	0.106	91.1	0.6	12.9
71	C1CC1			121	23	0.107	0.066	40.1	0.9	0.0
72	C1CCCC1			91	23	0.154	0.154	68.1	2.1	0.0
73	C1CCCCC1			139	24	0.135	0.113	82.2	2.6	0.0
74	CC1CC1			83	16	0.083	0.143	54.1	1.5	0.0
75	CC1CCCCC1			56	18	0.339	0.179	96.2	3.1	0.0
76	N1CCCC1			87	21	0.080	0.125	69.1	0.3	3.2

No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
77	N1CCCCC1			64	20	0.068	0.219	83.2	0.8	3.2
78	N1CCOCC1			182	23	0.070	0.151	85.1	-0.5	12.5
79	N1CCN(C)CC1			61	19	0.065	0.145	98.2	0.0	6.5
80	CN1CCCC1			70	18	0.100	0.100	83.2	0.5	3.2
81	CN1CCCCC1			66	16	0.015	0.104	97.2	1.0	3.2
82	CN1CCOCC1			183	26	0.049	0.141	99.2	-0.3	12.5
83	CCN1CCOCC1			152	20	0.051	0.135	113.2	0.0	12.5
84	C(=O)N1CCOCC1			65	16	0.030	0.106	113.1	-1.4	29.5
85	2,3-dimethyl			56	16	0.054	0.000	28.1	0.9	0.0
86	2,4-dimethyl			84	16	0.106	0.035	28.1	1.0	0.0
87	2,5-dimethyl			83	17	0.024	0.059	28.1	1.0	0.0

No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
88	3,4-dimethyl			87	18	0.102	0.034	28.1	0.9	0.0
89	3,5-dimethyl			74	18	0.027	0.040	28.1	1.0	0.0
90	2,3-dimethoxy			57	16	0.000	0.052	60.1	-0.3	18.5
91	2,4-dimethoxy			98	22	0.051	0.061	60.1	0.0	18.5
92	2,5-dimethoxy			83	21	0.036	0.119	60.1	0.0	18.5
93	3,4-dimethoxy			125	23	0.023	0.117	60.1	-0.3	18.5
94	3,5-dimethoxy			77	19	0.026	0.051	60.1	0.0	18.5
95	3,4,5-trimethoxy			72	20	0.028	0.153	90.1	-0.7	27.7
96	2,3-dichloro			111	21	0.126	0.072	68.9	1.3	0.0
97	2,4-dichloro			129	24	0.123	0.046	68.9	1.4	0.0
98	2,5-dichloro			73	17	0.081	0.027	68.9	1.4	0.0

No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
99	2,6-dichloro			57	17	0.086	0.121	68.9	1.4	0.0
100	3,4-dichloro			158	24	0.101	0.075	68.9	1.3	0.0
101	3,5-dichloro			87	20	0.080	0.161	68.9	1.4	0.0
102	n to c aliphatic			800	30	0.090	0.090	1.0	-2.2	12.0
103	n to c aromatic			2592	30	0.086	0.092	1.0	-1.5	12.9
104	nh to o			353	25	0.061	0.103	1.0	0.4	-2.8
105	nh to s			59	16	0.033	0.049	17.0	1.4	-12.0
106	ether to methylene			839	30	0.078	0.055	-2.0	1.8	-9.2
107	ether to thioether			261	29	0.026	0.067	16.1	1.0	-9.2
108	hydroxy to carbonyl			147	26	0.033	0.066	-2.0	-0.2	-3.2
109	sulfide to ethylene			70	19	0.014	0.070	-18.0	0.8	0.0
110	sulfide to sulfone			169	25	0.035	0.145	32.0	-2.6	34.1

No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
111	amide to ester			93	20	0.118	0.065	1.0	1.3	-2.8
112	amide to sulfonamide			296	28	0.017	0.100	36.1	0.1	17.1
113	amide to urea			269	26	0.077	0.095	1.0	0.4	12.0
114	amide to retroamide			390	23	0.020	0.060	0.0	0.3	0.0
115	olefin saturation			514	29	0.029	0.082	2.0	0.5	0.0
116	olefin to amide			77	17	0.025	0.152	17.0	-3.1	29.1
117	bromo to trifluoromethyl			357	27	0.011	0.063	-10.9	0.0	0.0
118	methyl to trifluoromethyl			602	29	0.033	0.098	54.0	0.4	0.0
119	nitro to trifluoromethyl			191	26	0.031	0.062	23.0	1.1	-43.1
120	carbamate to urea			76	17	0.118	0.092	-1.0	-0.5	2.8
121	carbonyl to sulfone			75	20	0.077	0.064	36.1	-1.2	17.1

No.	Modification	Parent Structure	Modified Structure	N	M	F(-1.0)	F(1.0)	ΔMW	ΔClogP	ΔPSA
122	carboxylate to amide			191	23	0.031	0.185	-1.0	-1.2	5.8
123	benzene to cyclohexane			410	29	0.036	0.128	6.1	1.2	0.0
124	benzene to thiophene			785	30	0.017	0.049	6.0	-0.4	0.0
125	benzene to thiazole			218	27	0.032	0.136	7.0	-1.7	12.9
126	pyridine to thiazole			56	16	0.036	0.411	6.0	-0.2	0.0
127	thiophene to thiazole			84	23	0.035	0.106	1.0	-1.3	12.9