Supplementary Data

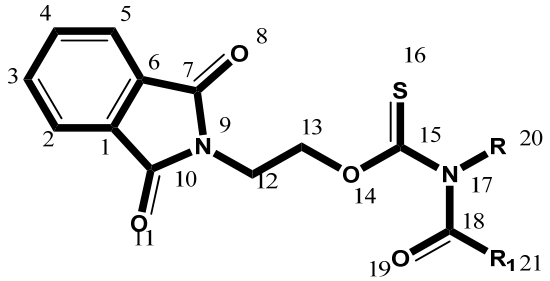
In silico de novo design of NNRTIs of HIV-1: Functional group based computational molecular modelling approach

U Raghuvanshi & N S Sapre*

Department of Applied Chemistry, SGSITS, Indore, India

Email: sukusap@yahoo.com

Table S1 — Substituent (R and R_1) on the parent structure, anti-HIV-1 activity (Effective concentration pEC_{50} in μM terms), and electrotopological indices (Ring A, E_{O8}, E_{N9} O14, S16, N17, O19, R, R1) of ATC analogues (Training, Test and Outliers)



Template structure of ATC

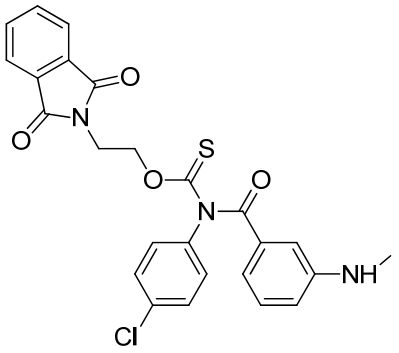
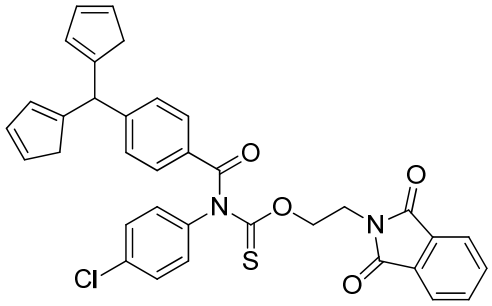
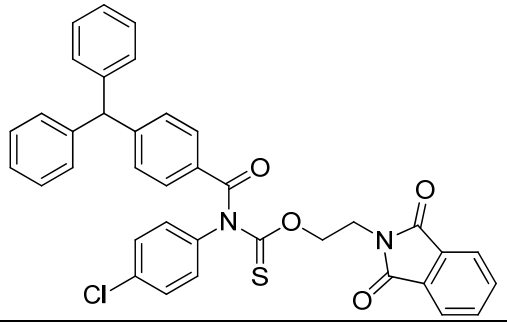
COMP	PEC_{50}	R	R1	Ring A	E_{O8}	E_{N9}	O14	S16	N17	O19	R	R1
1[a]	6.40	$-C_6H_5$	$-C_6H_5$	7.8160	12.4836	1.1242	5.6491	5.3965	1.2987	13.0988	9.4805	9.1972
2	4.96	-2- CH_3-C_6H_4	$-C_4H_3S$	7.8554	12.5001	1.1314	5.6853	5.4347	1.3461	13.1225	10.7220	7.1624
3	4.96	-2- C_2H_5-C_6H_4	$-C_4H_3S$	7.8822	12.5389	1.1376	5.7298	5.4826	1.3774	13.2385	11.8116	7.2029
4	4.96	-2- Cl-C_6H_4	$-C_4H_3S$	7.6978	12.4826	1.1054	5.6266	5.3760	1.2142	13.0639	13.7890	6.9517
5	4.96	-2- $OCH_3C_6H_4$	$-C_4H_3S$	7.7329	12.5224	1.1129	5.6742	5.4270	1.2524	13.1829	14.7173	7.0033
6[a]	5.00	-3- CH_3-C_6H_4	$-C_4H_3S$	7.8517	12.4929	1.1303	5.6701	5.4183	1.3299	13.0816	10.8441	7.1535
7[a]	4.96	-3- CF_3-C_6H_4	$-C_4H_3S$	6.9173	12.4803	0.9891	5.4745	5.2305	0.8599	13.0500	38.1373	6.1032
8[a]	5.22	-3-CO CH_3-C_6H_4	$-C_4H_3S$	7.5901	12.5248	1.0923	5.6457	5.3991	1.2043	13.1665	20.4212	6.8628
9[a]	4.96	-3- F-C_6H_4	$-C_4H_3S$	7.4769	12.4512	1.0703	5.5477	5.2959	1.0899	12.9592	18.8688	6.6954
10[a]	5.22	-3- Cl-C_6H_4	$-C_4H_3S$	7.7198	12.4782	1.1092	5.6270	5.3752	1.2455	13.0385	13.6676	6.9924
11[a]	5.92	-3- Br-C_6H_4	$-C_4H_3S$	7.8050	12.4877	1.1228	5.6548	5.4030	1.2999	13.0663	11.8472	7.0963
12	6.42	-3- NO_2-C_6H_4	$-C_4H_3S$	7.3039	12.4929	1.0477	5.5608	5.3142	1.0521	13.0816	27.1822	6.5358
13[a]	5.46	-3-O CH_3-C_6H_4	$-C_4H_3S$	7.7495	12.5121	1.1153	5.6630	5.4138	1.2699	13.1291	14.7040	7.0318
14	7.52	-4- C_2H_5-C_6H_4	$-C_4H_3O$	7.7443	12.5020	1.1131	5.6405	5.3900	1.2321	13.0206	11.8792	9.8942
15	7.00	-4- F-C_6H_4	$-C_4H_3S$	7.5310	12.4518	1.0800	5.5668	5.3142	1.1557	12.9615	18.4986	6.7860
16	7.00	-4- Cl-C_6H_4	- C_6H_5	7.7164	12.5030	1.1100	5.6463	5.3963	1.2533	13.1507	13.5666	9.0761
17	7.60	-4- Cl-C_6H_4	-4- Cl-C_6H_4	7.6303	12.5197	1.0981	5.6440	5.3960	1.2195	13.2189	13.4455	13.1222
18[b]	8.15	-4- Cl-C_6H_4	$-C_4H_3O$	7.6129	12.4611	1.0915	5.5812	5.3286	1.1596	12.9181	13.3129	9.7258
19	8.10	-4- Cl-C_6H_4	$-C_4H_3S$	7.7373	12.4748	1.1121	5.6275	5.3749	1.2638	13.0222	13.5869	7.0206
20[b]	8.30	-4- Cl-C_6H_4	-2-Cl-C_5H_3N	7.5369	12.5189	1.0831	5.6227	5.3761	1.1497	13.2471	13.2491	14.7105
21	8.22	-4- Cl-C_6H_4	-2-Cl-C_5H_3N	7.5677	12.5128	1.0881	5.6236	5.3756	1.1795	13.1789	13.3234	14.5598

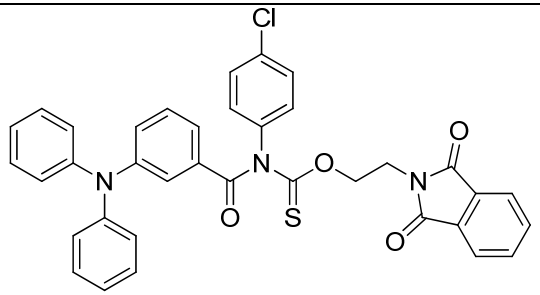
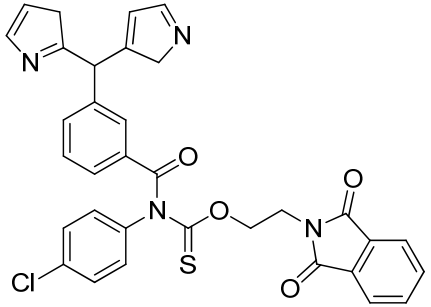
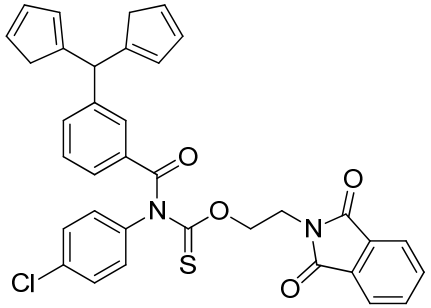
22	7.46	-4- Br-C_6H_4	- C_6H_5	7.7885	12.5111	1.1213	5.6676	5.4176	1.2912	13.1720	11.8365	9.1695
23	8.00	-4- I-C_6H_4	-C_4H_3S	7.8430	12.4866	1.1286	5.6586	5.4060	1.3191	13.0534	11.0720	7.1407
24	8.10	-4- NO_2-C_6H_4	-C_4H_3O	7.2535	12.4735	1.0395	5.5278	5.2793	1.0142	12.9511	26.4920	9.3861
25	8.00	-4- NO_2-C_6H_4	-C_4H_3S	7.3777	12.4873	1.0601	5.5741	5.3256	1.1183	13.0552	26.7994	6.6475
26 [b]	6.30	-4-(OC_2H_5)-C_6H_4	-2-Cl-C_5H_3N	7.6093	12.5667	1.0967	5.6728	5.4286	1.2067	13.3064	15.6744	14.6635
27[b]	5.07	-4- F-C_6H_4	-4-$O(CH_2)_4$ CH_3-C_6H_4	7.4931	12.5438	1.0732	5.6232	5.3783	1.1318	13.3183	18.5481	15.1642
28	7.40	-4- Cl-C_6H_4	-2- CH_3-C_6H_4	7.7310	12.5405	1.1142	5.6862	5.4396	1.2742	13.3716	13.6292	10.2475
29	7.15	-4- Cl-C_6H_4	-2- NO_2-C_6H_4	7.1829	12.5405	1.0315	5.5768	5.3354	0.9964	13.3716	12.6980	27.1084
30	8.00	-4- Cl-C_6H_4	-3- CH_3-C_6H_4	7.7284	12.5349	1.1134	5.6766	5.4292	1.2667	13.3029	13.6149	10.3819
31	8.82	-4- Cl-C_6H_4	-3- NO_2-C_6H_4	7.2569	12.5349	1.0439	5.5901	5.3469	1.0626	13.3029	12.8729	26.6319
32	7.30	-4- Cl-C_6H_4	-4- CH_3-C_6H_4	7.7266	12.5305	1.1128	5.6700	5.4221	1.2626	13.2620	13.6050	10.4652
33	8.30	-4- Cl-C_6H_4	-4 -CF_3-C_6H_4	7.0263	12.5210	1.0115	5.5443	5.3014	0.9967	13.2305	12.5687	36.8874
34 [a]	7.22	-4- Cl-C_6H_4	-4- NO_2-C_6H_4	7.3164	12.5305	1.0536	5.6000	5.3554	1.1064	13.2620	12.9998	26.2901
35	7.46	-4- Br-C_6H_4	-4 -Br-C_6H_4	7.7648	12.5347	1.1188	5.6821	5.4341	1.2851	13.2680	11.8087	11.5119
36	6.70	-4- Br-C_6H_4	-2-Cl-C_5H_3N	7.6089	12.5270	1.0943	5.6440	5.3974	1.1875	13.2684	11.5053	14.8270
37	6.70	-4- CH_3-C_6H_4	-2- CH_3-C_6H_4	7.8428	12.5530	1.1316	5.7192	5.4726	1.3328	13.4045	10.9245	10.4134
38 [a]	6.40	-4- CH_3-C_6H_4	-2- NO_2-C_6H_4	7.2952	12.5530	1.0490	5.6098	5.3684	1.0550	13.4045	9.9578	27.3092
39	7.40	-4- CH_3-C_6H_4	-3- CH_3-C_6H_4	7.8406	12.5474	1.1308	5.7095	5.4622	1.3253	13.3359	10.9135	10.5442
40	7.70	-4- CH_3-C_6H_4	-3- NO_2-C_6H_4	7.3690	12.5474	1.0614	5.6231	5.3799	1.1212	13.3359	10.1422	26.8235
41	7.70	-4- CH_3-C_6H_4	-3- CH_3-C_6H_4	7.8384	12.5430	1.1302	5.7030	5.4551	1.3212	13.2950	10.9065	10.6247
42	7.52	-4- CH_3-C_6H_4	-4- NO_2-C_6H_4	7.4285	12.5430	1.0711	5.6330	5.3884	1.1650	13.2950	8.3843	26.4746
43	8.10	-4- Cl-C_6H_4	-2,3-di-F-C_6H_3	7.0501	12.4952	1.0079	5.5002	5.2562	0.8808	13.1171	12.3765	28.1533
44	8.00	-4- Cl-C_6H_4	-2,4-di-F-C_6H_3	7.0925	12.4957	1.0153	5.5134	5.2689	0.9210	13.1204	12.4813	27.8135
45	7.70	-4- Cl-C_6H_4	-2,5-di-F-C_6H_3	7.0501	12.4952	1.0079	5.5002	5.2562	0.8808	13.1171	12.3765	28.1534
46	7.70	-4- Cl-C_6H_4	-2,6-di-F-C_6H_3	6.9963	12.4947	0.9983	5.4812	5.2380	0.8150	13.1124	12.2256	28.6697
47	8.10	-4- Cl-C_6H_4	-3,4-di-F-C_6H_3	7.1463	12.4963	1.0249	5.5325	5.2871	0.9868	13.1251	12.6322	27.2972
48	8.15	-4- Cl-C_6H_4	-3,5- di-F -C_6H_3	7.1039	12.4958	1.0175	5.5193	5.2745	0.9467	13.1218	12.5277	27.6368
49	7.00	-4- Cl-C_6H_4	-2,3- di-Cl -C_6H_3	7.4993	12.5453	1.0789	5.6404	5.3964	1.1444	13.3807	13.2195	17.3940
50	7.52	-4- Cl-C_6H_4	-2,4- di-Cl -C_6H_3	7.5131	12.5426	1.0811	5.6408	5.3962	1.1559	13.3553	13.2501	17.3286
51 [a]	5.25	-4- Cl-C_6H_4	-2,6- di-Cl -C_6H_3	7.4821	12.5487	1.0761	5.6399	5.3967	1.1261	13.4235	13.1760	17.4887
52	8.05	-4- Cl-C_6H_4	-3,4- di-Cl -C_6H_3	7.5302	12.5392	1.0840	5.6412	5.3959	1.1742	13.3125	13.2939	17.2338
53	8.10	-4- Cl-C_6H_4	-3,5- di-Cl -C_6H_3	7.5168	12.5418	1.0818	5.6408	5.3960	1.1627	13.3379	13.2634	17.2993
54	6.40	-4- Cl-C_6H_4	-2,6-di-(O CH_3) -C_6H_3	7.5411	12.6163	1.0883	5.7117	5.4737	1.1750	13.7502	13.3398	19.3016
55	8.10	-4- Cl-C_6H_4	-2-Cl,4-NO_2-C_6H_3	7.1994	12.5533	1.0366	5.5968	5.3556	1.0427	13.3984	12.8043	30.4963
56 [a]	7.15	-4- Cl-C_6H_4	-2-Cl,5-NO_2-C_6H_3	7.1399	12.5577	1.0269	5.5869	5.3471	0.9990	13.4393	12.6778	30.8382
57	7.70	-4- Cl-C_6H_4	-4-Cl,5-NO_2-C_6H_3	7.1708	12.5516	1.0320	5.5878	5.3466	1.0288	13.3712	12.7516	30.6780
58	8.15	-4- Cl-C_6H_4	-4-Br,5- CH_3-C_6H_3	7.7045	12.5586	1.1109	5.6910	5.4457	1.2606	13.3989	13.5964	12.7150
59[b]	7.70	-4- Cl-C_6H_4	-2,5 CH_3-C4HO	7.6581	12.5324	1.1025	5.6624	5.4158	1.2275	13.3249	13.4867	11.9246
60	7.52	-4- CH_3-C_6H_4	-2,4-di-F-C_6H_3	7.2047	12.5082	1.0327	5.5464	5.3018	0.9796	13.1534	9.7616	27.9943
61	7.40	-4- CH_3-C_6H_4	-2,5-di-F-C_6H_3	7.1620	12.5077	1.0254	5.5332	5.2892	0.9395	13.1501	9.6542	28.3368
62 [a]	6.00	-4- CH_3-C_6H_4	-2,6-di-F-C_6H_3	7.1081	12.5071	1.0157	5.5142	5.2710	0.8737	13.1454	9.4995	28.8572
63	7.30	-4- CH_3-C_6H_4	-3,5-di-F-C_6H_3	7.2156	12.5083	1.0350	5.5523	5.3075	1.0053	13.1548	9.8088	27.8164
64	6.70	-4- CH_3-C_6H_4	-2,3-di-Cl-C_6H_3	7.6114	12.5577	1.0964	5.6733	5.4294	1.2031	13.4137	10.4973	17.5773

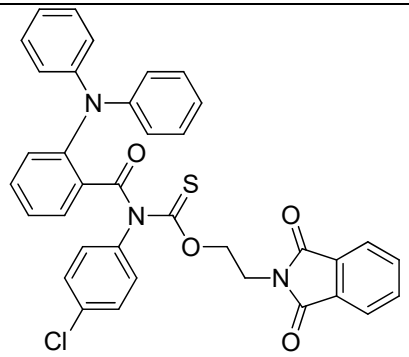
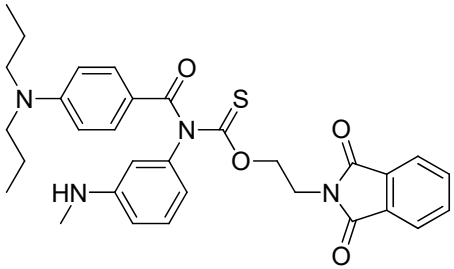
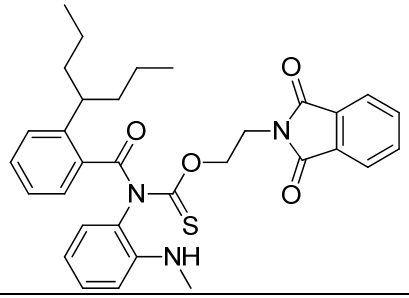
65 [a]	6.52	-4- CH_3-C_6H_4	-2,4-di-Cl-C_6H_3	7.6250	12.5551	1.0986	5.6738	5.4292	1.2145	13.3883	10.5305	17.5089
66 [a]	5.05	-4- CH_3-C_6H_4	-2,6-di-Cl-C_6H_3	7.5941	12.5612	1.0935	5.6729	5.4297	1.1848	13.4565	10.4497	17.6755
67	7.30	-4- CH_3-C_6H_4	-3,4-di-Cl-C_6H_3	7.6423	12.5516	1.1015	5.6742	5.4289	1.2329	13.3455	10.5779	17.4107
68[b]	4.66	-4- CH_3-C_6H_4	-2,6-di-(OCH_3) -C_6H_3	7.6533	12.6288	1.1057	5.7447	5.5067	1.2337	13.7832	10.5789	19.5236
69 [a]	7.05	-4- CH_3-C_6H_4	-2-Cl,4-NO_2-C_6H_3	7.3113	12.5658	1.0541	5.6298	5.3886	1.1014	13.4314	10.0597	30.7021
70 [a]	7.22	-4- CH_3-C_6H_4	-4-Cl,5-NO_2-C_6H_3	7.2829	12.5641	1.0495	5.6208	5.3796	1.0874	13.4041	10.0065	30.8843
71	7.30	-4- CH_3-C_6H_4	-4-Br,5- CH_3-C_6H_3	7.8168	12.5711	1.1284	5.7240	5.4787	1.3193	13.4319	10.8803	12.8918
72 [a]	6.30	- C_6H_5	-3,4-diCl-C_6H_3	7.6303	12.5197	1.0981	5.6440	5.3960	1.2195	13.2606	9.2074	17.3550
73 [a]	6.46	- C_6H_5	-4-Cl,5-NO_2-C_6H_3	7.2705	12.5322	1.0461	5.5906	5.3467	1.0741	13.3192	8.6934	30.7709
74	7.52	-4- Cl-C_6H_4	-1($C_{10}H_7$)	7.7271	12.6275	1.1175	5.7584	5.5191	1.2875	13.7647	13.6913	15.1702
75	7.70	-4- Cl-C_6H_4	-2($C_{10}H_7$)	7.7251	12.6275	1.1175	5.7584	5.5191	1.2875	13.7647	13.6913	15.1702
76	5.59	-2,4- di-Cl-C_6H_3	-C_4H_3O	7.4735	12.4883	1.0707	5.5776	5.3296	1.0647	13.0117	17.5557	9.5818
77	4.77	-2,6- di-Cl-C_6H_3	-3-NO_2-C_6H_4	7.0779	12.5699	1.0163	5.5856	5.3490	0.9181	13.4382	17.3229	26.4139
78	6.16	-3,4- di-Cl-C_6H_3	-C_4H_3O	7.4959	12.4839	1.0745	5.5780	5.3288	1.0960	12.9863	17.4420	9.6145

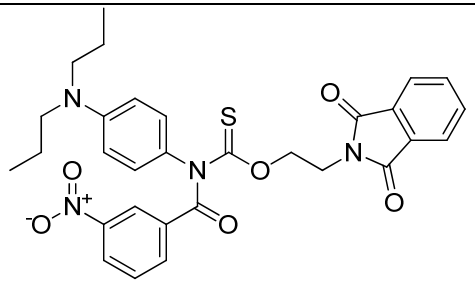
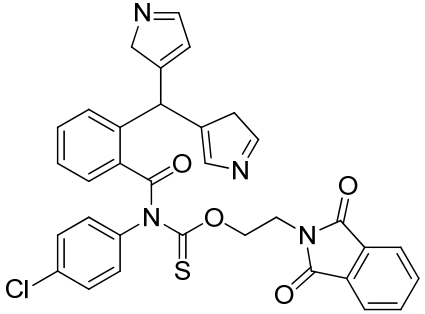
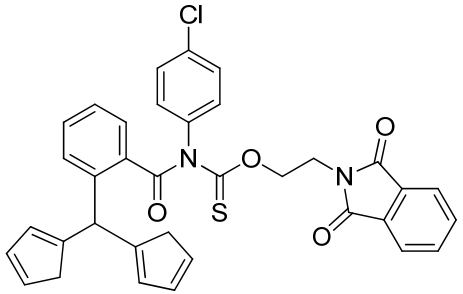
pEC_{50} = -logEC_{50} (it is the compound dose required to achieve 50 % protection of MT-4 cells from HIV-1 induced cytopathogenecity, as determined by MTT method expressed in terms of EC50 and refers to the concentration of a drug which induces a response halfway between the baseline and maximum). [a]Data points present in Test Set. [b]Data points not included in deriving the equation(Outliers)

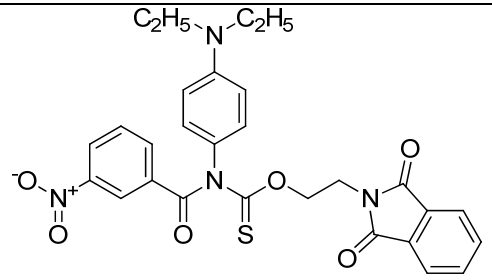
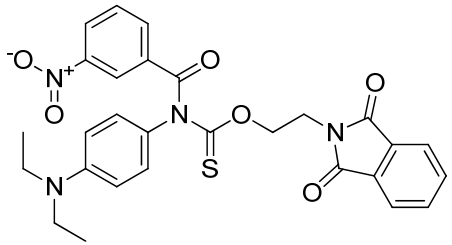
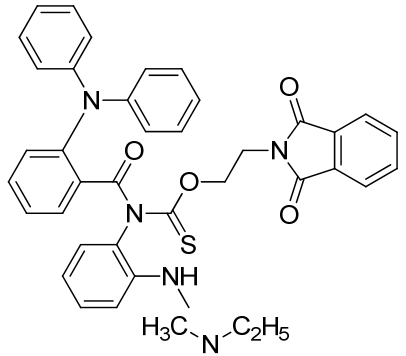
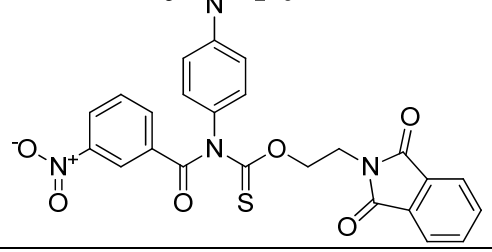
Table S2 — Virtual dataset (26 compounds those are having pEC_{50} (in μM terms) value higher than 8)

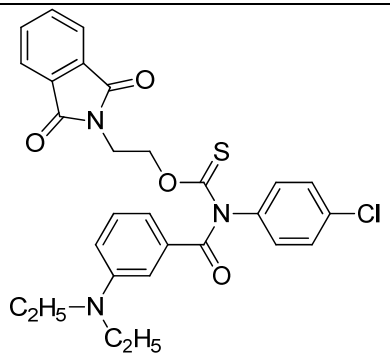
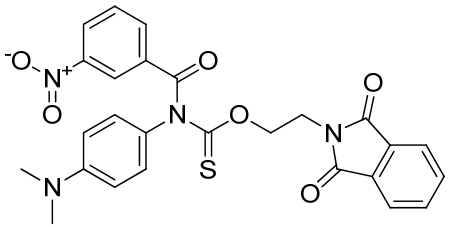
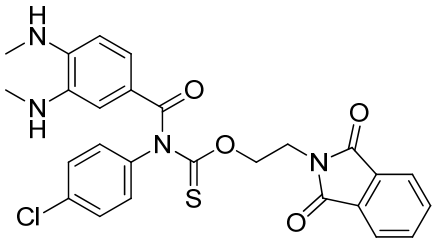
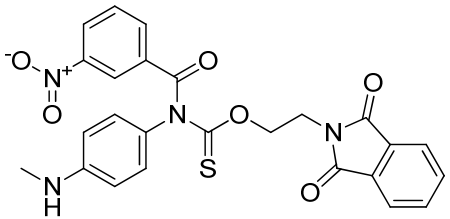
COMP. NO.	R	Ring A	O8	N9	O14	S16	N17	O19	R	R1	ANN (10-2-1)#4
VDS01		7.6948	12.5575	1.1092	1.2528	5.4419	1.2528	13.3911	13.5743	12.8785	8.55535
VDS02		7.7639	12.7326	1.1265	5.8145	5.58	1.299	13.9016	13.8017	32.6646	8.55416
VDS03		7.7204	12.7595	1.1214	5.824	5.5915	1.2866	13.9719	13.7593	31.5628	8.55285

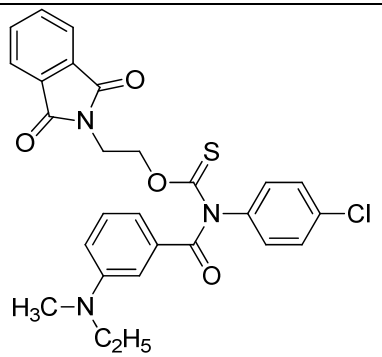
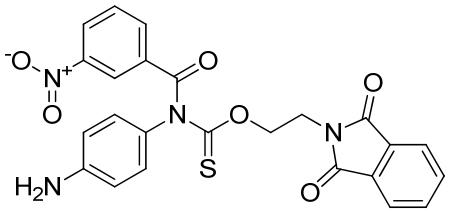
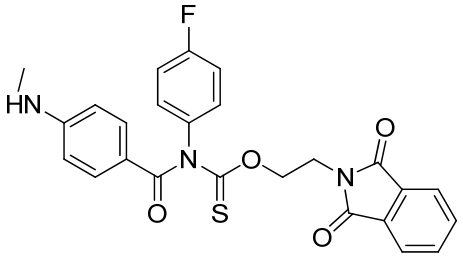
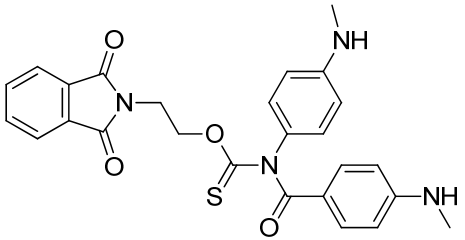
VDS04		7.6868	12.791	1.118	5.8523	5.6235	1.2803	14.18	13.7346	32.0292	8.5472
VDS05		7.697	12.7564	1.1184	5.8336	5.6025	1.2865	14.0862	13.7363	28.5057	8.5332
VDS06		7.7716	12.7648	1.1294	5.8518	5.6207	1.3144	14.1142	13.8496	26.1221	8.53115

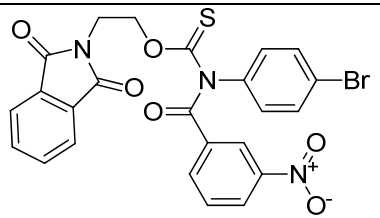
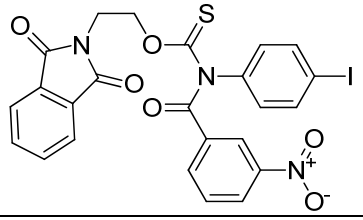
VDS07		7.6833	12.8336	1.1197	5.9024	5.6786	1.2927	14.512	13.7691	31.4944	8.50529
VDS08		7.8425	12.7012	1.1371	5.8198	5.5837	1.3367	13.797	13.4706	19.5463	8.45465
VDS09		7.8727	12.7046	1.1417	5.828	5.5919	1.3503	13.8106	13.5266	18.7017	8.45217

VDS10		7.4438	12.6946	1.0784	5.7444	5.5114	1.1684	13.6755	18.8629	26.4009	8.44673
VDS11		7.702	12.7954	1.1214	5.8835	5.657	1.3085	14.4043	13.7916	27.9203	8.43115
VDS12		7.7817	12.8043	1.1333	5.9035	5.677	1.3398	14.4356	13.9151	25.5014	8.42123

VDS13		7.3873	12.648	1.0683	5.7025	5.4667	1.1416	13.5717	16.2516	27.0935	8.33123
VDS14		7.4153	12.6511	1.0725	5.7096	5.4738	1.1528	13.5829	16.3092	26.2645	8.3267
VDS15		7.7536	12.8895	1.1328	5.9768	5.7585	1.3438	14.7004	13.3066	31.7164	8.27568
VDS16		7.3709	12.6232	1.0649	5.6813	5.4438	1.1314	13.5156	15.0198	27.0109	8.26416

VDS17		7.7469	12.6355	1.1203	5.7559	5.516	1.2871	13.6837	13.7158	16.5167	8.24625
VDS18		7.3542	12.5984	1.0614	5.6602	5.421	1.1212	13.4594	13.7877	26.9282	8.16748
VDS19		7.3822	12.6016	1.0656	5.6673	5.4281	1.1324	13.4705	13.839	26.1053	8.16141
VDS20		7.3351	12.57	1.0573	5.6338	5.3926	1.1073	13.3898	12.6117	26.8304	8.13244

VDS21		7.0504	12.5119	1.0118	5.5294	5.2861	0.9546	13.2422	17.903	26.2791	8.13123
VDS22		7.6757	12.6046	1.1085	5.7198	5.4782	1.2518	13.5703	13.5809	16.7614	8.12126
VDS23		7.363	12.5731	1.0614	5.641	5.3998	1.1185	13.401	12.6592	26.0114	8.11567
VDS24		7.1565	12.5238	1.0283	5.5606	5.3174	1.0101	13.2734	15.3162	26.4606	8.1131

VDS25		7.3294	12.543	1.0552	5.6114	5.3681	1.1004	13.3242	11.1126	26.7554	8.08671
VDS26		7.3628	12.5467	1.0604	5.6213	5.378	1.1179	13.3341	10.2975	26.8128	8.07679