Application of Quantitative Structure-Activity Relationship Models of 5-HT$_{1A}$ Receptor Binding to Virtual Screening Identifies Novel and Potent 5-HT$_{1A}$ Ligands

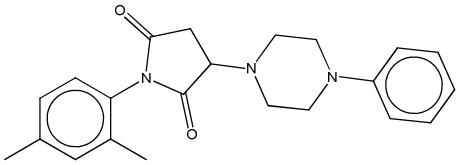
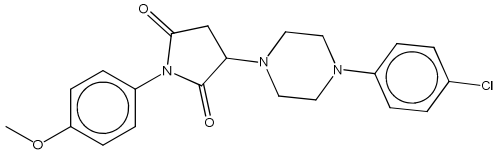
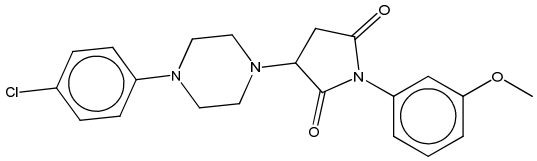
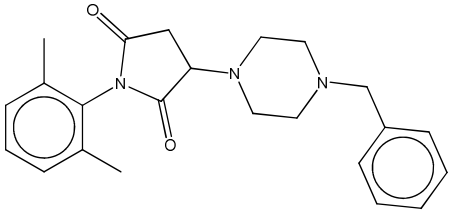
Man Luo[‖], Xiang Simon Wang[†*], Bryan L. Roth[≠‖], Alexander Golbraikh[‖], and Alexander Tropsha[‖*]

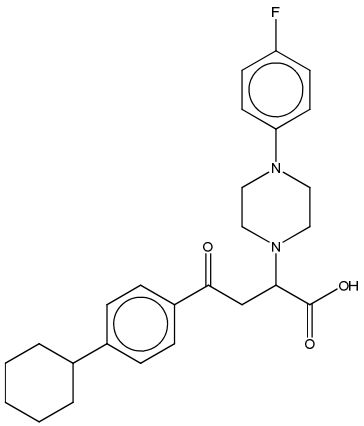
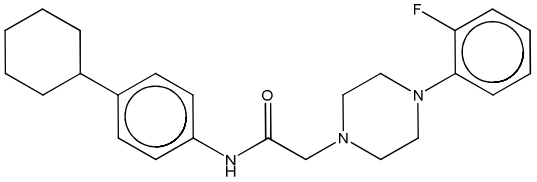
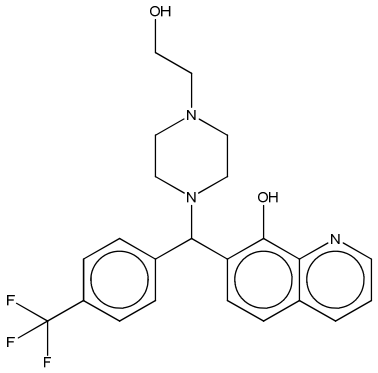
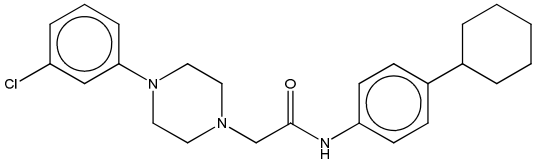
Supplementary Materials

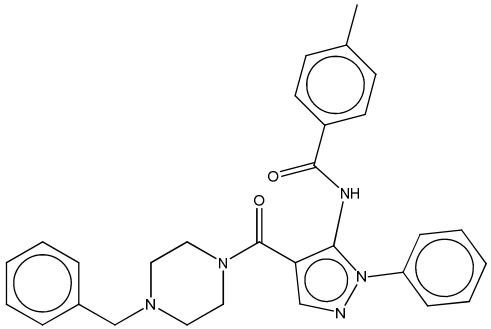
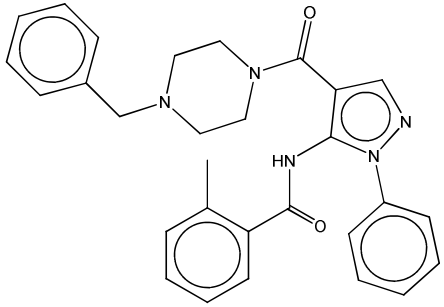
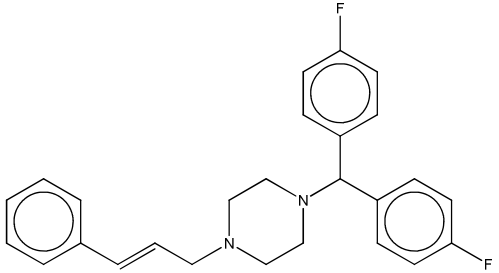
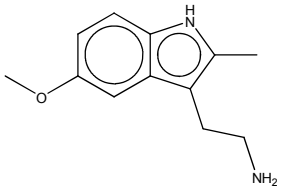
Table S1. The number of screening compounds within different n_{cutoff} values for 5-HT_{1A} binders virtual screening against 6 different types of chemical libraries.

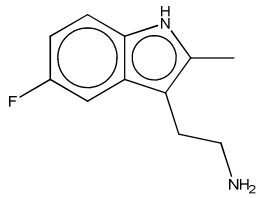
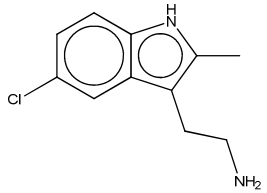
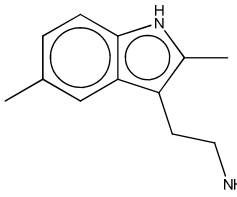
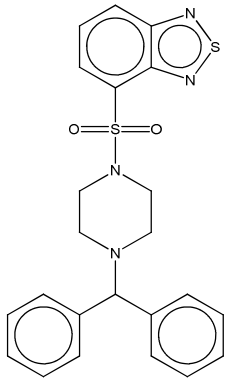
Screening libraries	Number of cmpds	n_{cutoff}	Cmpds in AD	Cmpds in AD (%)
Prestwick	1,552	-0.5	209	13.5%
		0	304	19.6%
		0.5	458	29.5%
World Drug Index	53,382	-0.5	1334	2.5%
		0	3295	6.2%
		0.5	7371	13.8%
TimTec GPCRTargeted Library	2,300	-0.5	31	1.3%
		0	151	6.6%
		0.5	542	23.6%
Asinex GPCR Targeted Library	3,233	-0.5	144	4.5%
		0	890	27.5%
		0.5	2279	70.5%
TimTec Diversity Library	10,000	-0.5	46	0.5%
		0	222	2.2%
		0.5	803	8.0%
Asinex Diversity Library	5,072	-0.5	39	0.8%
		0	267	5.3%
		0.5	811	16.0%

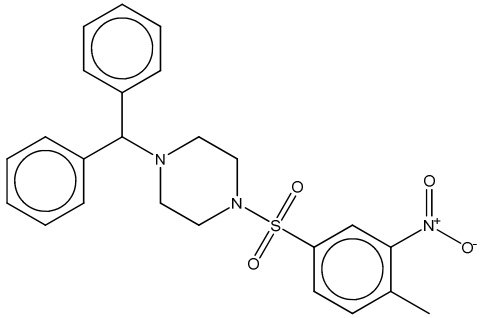
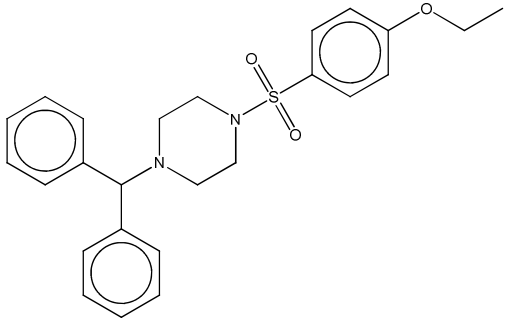
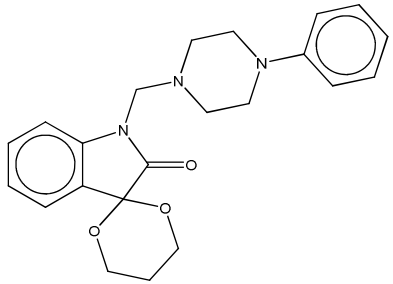
Table S2. The 21 screening hits predicted as 5-HT_{1A} binders as a result of mining the TimTec GPCR targeted library.

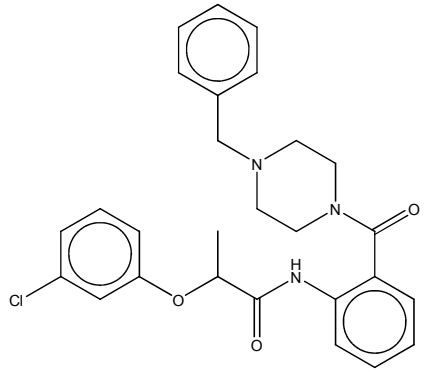
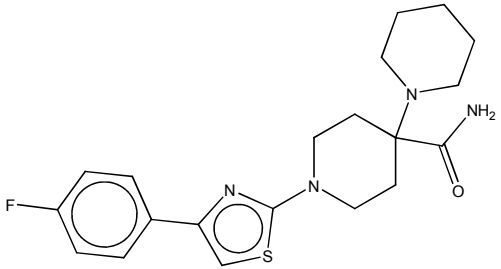
Structure	Serial No.	TimTec ID	*k*-Nearest Neighbor prediction score	Random Forest prediction score	Support Vector Machines prediction score	Percent of inhibition	Exp. IC_{50} (nM)[a]
	1	ST007335	1.08	1.30	1.00	ND[c]	ND
	2	ST007839	1.08	1.17	1.06	ND	ND
	3	ST007188	1.10	1.27	1.00	ND	ND
	4	ST039362	1.08	1.33	1.00	ND	ND

	5	ST007484	1.04	1.12	1.00	ND	ND
	6	ST016950	1.07	1.10	1.00	5.5%[a]	ND
	7	ST014829	1.06	1.27	1.00	39.20%	ND
	8	ST007472	1.03	1.09	1.00	21.80%	ND

	9	ST015321	1.07	1.17	1.00	ND	ND
	10	ST016028	1.08	1.15	1.00	ND	ND
	11	ST024768	1.05	1.20	1.00	ND	ND
	12	ST038835	1.07	1.12	1.00	ND	ND

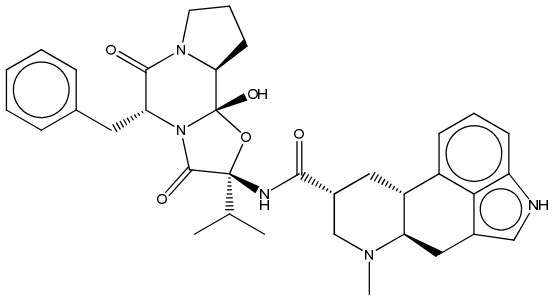
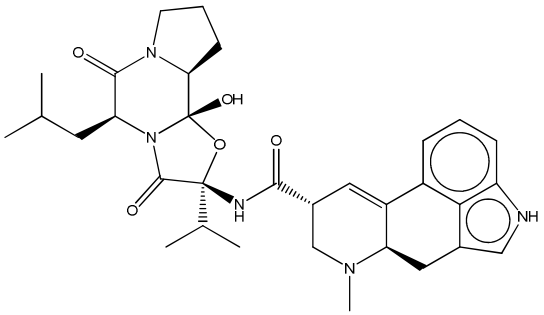
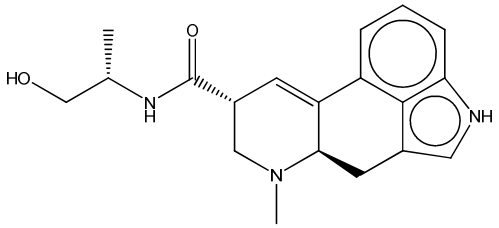
	13	ST038838	1.03	1.10	1.00	ND	ND
	14	ST037879	1.09	1.15	1.00	ND	ND
	15	ST039394	1.09	1.07	1.00	ND	ND
	16	ST044960	1.05	1.27	1.00	ND	ND

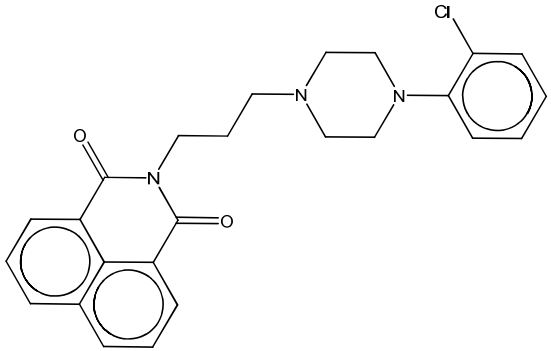
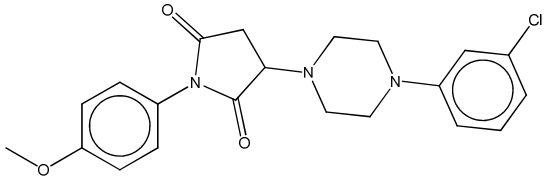
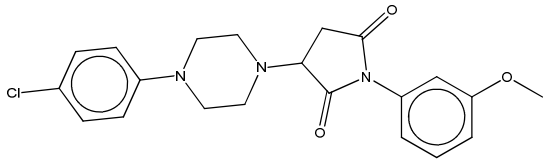
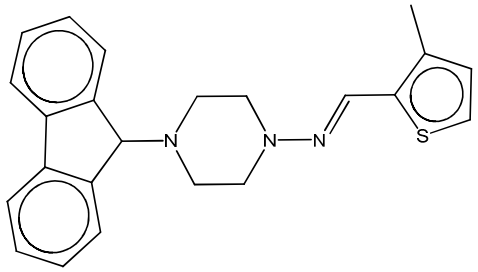
	17	ST024676	1.09	1.29	1.00	ND	ND
	18	ST055405	1.06	1.18	1.00	ND	ND
	19	ST030580	1.02	1.09	1.00	95.50%	306.0[b]

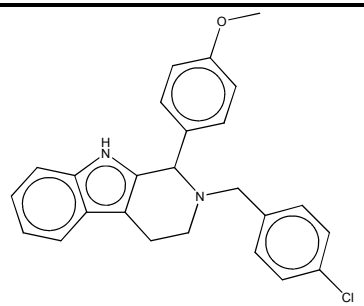
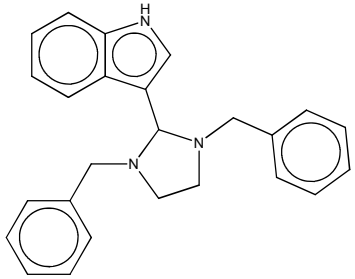
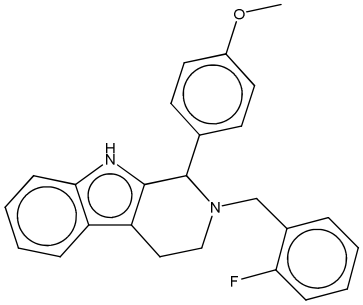
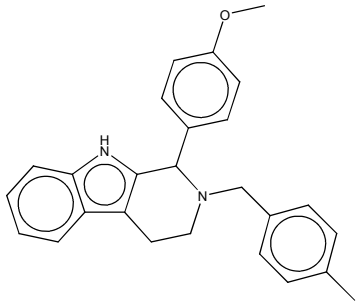
	20	ST015856	1.06	1.16	1.00	ND	ND
	21	ST041900	1.07	1.09	1.00	4.70%	ND

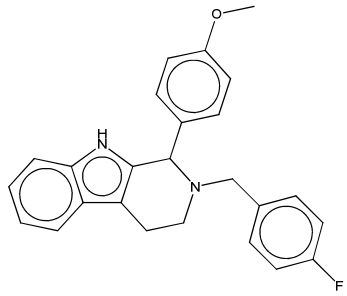
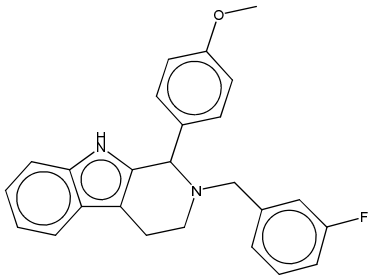
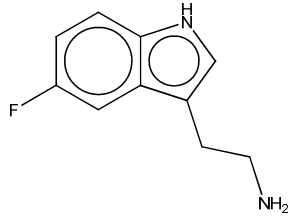
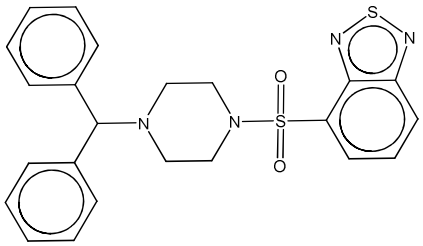
[a]: Only five hit compounds were sent to experimental validation upon their structural similarity and the commercial availability.
[b]: The full IC_{50} curve was generated in further experiment and the Ki value was determined.
[c]: Not determined.

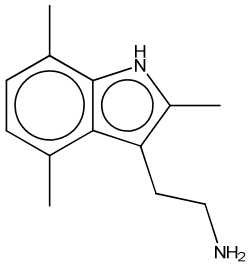
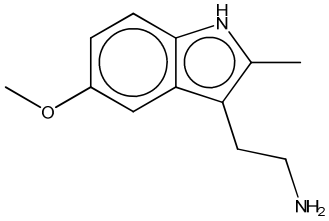
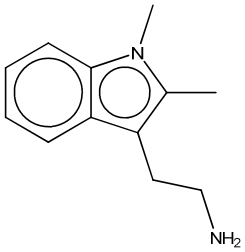
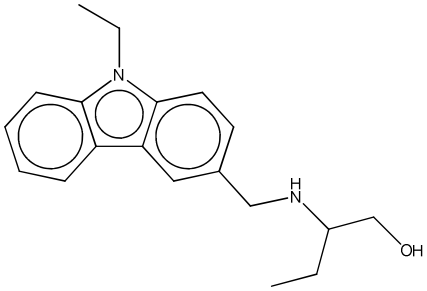
Table S3. The 22 screening hits predicted as 5-HT_{1A} binders as a result of mining the TimTec Diversity library.

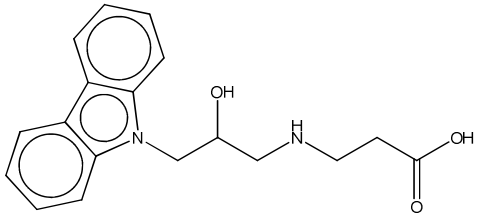
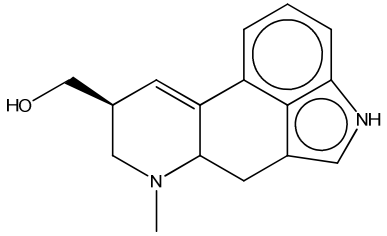
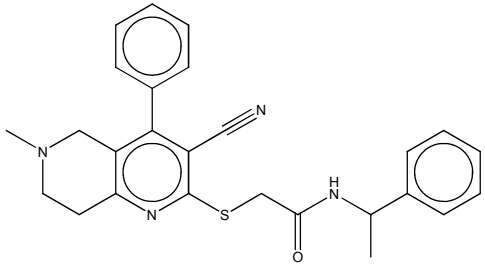
Structure	Serial No.	TimTec ID	*k*-Nearest Neighbor prediction score	Random Forest prediction score	Support Vector Machines prediction score	Percent of inhibition	Exp. IC_{50} (nM)[a]
	22	ST056345	1.01	1.14	1.00	ND[c]	ND
	23	ST057581	1.01	1.14	1.00	ND	ND
	24	ST057158	1.01	1.03	1.00	ND	ND

	25	ST023860	1.02	1.12	1.00	95.60%[a]	159.0[b]
	26	ST007110	1.08	1.14	1.06	11.10%	ND
	27	ST007188	1.10	1.27	1.06	ND	ND
	28	ST041338	1.03	1.22	1.00	ND	ND

	29	ST074312	1.06	1.20	1.00	ND	ND
	30	ST055328	1.08	1.43	1.00	ND	ND
	31	ST074310	1.09	1.18	1.00	ND	ND
	32	ST074317	1.09	1.14	1.00	ND	ND

	33	ST074308	1.10	1.15	1.00	ND	ND
	34	ST074311	1.10	1.14	1.00	70.80%	8194.0
	35	ST073803	1.03	1.22	1.00	ND	ND
	36	ST044960	1.05	1.28	1.00	ND	ND

	37	ST038832	1.06	1.12	1.00	ND	ND
	38	ST038835	1.07	1.12	1.00	ND	ND
	39	ST037883	1.08	1.16	1.00	ND	ND
	40	ST072504	1.09	1.20	1.00	ND	ND

	41	ST072476	1.09	1.33	1.00	ND	ND
	42	ST057540	1.05	1.14	1.00	98.20%	2.3
	43	ST066677	1.06	1.13	1.00	50.0%	ND

[a]: Only five hit compounds were sent to experimental validation upon their structural similarity and the commercial availability.
[b]: The full IC_{50} curve was generated in further experiment and the Ki value was determined.
[c]: Not determined.

Table S4. The 105 5-HT_{1A} actives from the PDSP database employed in the QSAR modeling studies.

Cmpds ID	PDSP ID	SMILES
1	3333	CC(CC1=CC=CC=C1)N
2	3334	C=CCN1CCC2=C(C(=C(C=C2C(C1)C3=CC=CC=C3)O)O)Cl
3	3336	CC(C)CC1C(=O)N2CCCC2C3(N1C(=O)C(O3)(C(C)C)NC(=O)C4CN(C5CC6=C(NC7=CC=CC(=C67)C5=C4)Br)C)O
4	3338	CN1CCN(CC1)C2=NC3=C(C=CC(=C3)C(F)(F)F)N4C2=CC=C4.C(=CC(=O)O)C(=O)O
5	3342	CN1CCC(=C2C3=CC=CC=C3C=CC4=CC=CC=C42)CC1
6	3345	CC1(C(=O)N2C(C(=O)N3CCCC3C2(O1)O)CC4=CC=CC=C4)NC(=O)C5CC6C(CC7=CNC8=CC=CC6=C78)N(C5)C
7	3347	C1=CC(=C(C=C1CCN)O)O
8	3349	CCC1=CC(=C(C(=C1O)C(=O)NCC2CCCN2CC)OC)Cl
9	3350	C1CN(CCN1CCC=C2C3=CC=CC=C3SC4=C2C=C(C=C4)C(F)(F)F)CCO
10	3351	CCCCCCC(=O)OCCN1CCN(CC1)CCCN2C3=CC=CC=C3SC4=C2C=C(C=C4)C(F)(F)F
11	3352	C1CN(CCN1CCCC2=CC=CC=C2)CCOC(C3=CC=C(C=C3)F)C4=CC=C(C=C4)F
12	3354	C1CN(CCC1(C2=CC=C(C=C2)Cl)O)CCCC(=O)C3=CC=C(C=C3)F
13	3357	C1CN=C(N1)C2COC3=CC=CC=C3O2
14	3364	CC1=C(C2=C(N1)C=CC(=C2)O)CCN
15	3367	CN1CCN2C(C1)C3=CC=CC=C3CC4=CC=CC=C42
16	3369	CN1CC(C2=C(C1)C(=CC=C2)N)C3=CC=CC=C3
17	3371	CC1=CC=C(C=C1)N(CC2=NCCN2)C3=CC(=CC=C3)O.CS(=O)(=O)O
18	3372	C1CN(CCC1N2C3=CC=CC=C3NC2=O)CCCC(C4=CC=C(C=C4)F)C5=CC=C(C=C5)F
19	3373	CCCN1CCCC(C1)C2=CC(=CC=C2)O
20	3374	CC(C)NCC(COC1=CC=CC2=CC=CC=C21)O

21	3375	CCCN1CCCC2C1CC3=C(C2)NN=C3
22	3377	COC1C(CC2CN3CCC4=C(C3CC2C1C(=O)OC)NC5=C4C=CC(=C5)OC)OC(=O)C6=CC(=C(C(=C6)OC)OC)OC
23	3378	CC1=C(C(=O)N2C=CSC2=N1)CCN3CCC(=C(C4=CC=C(C=C4)F)C5=CC=C(C=C5)F)CC3
24	3379	CN1CCC2=CC(=C(C=C2C(C1)C3=CC=CC=C3)O)Cl
25	3381	C1CNCC(C2=CC(=C(C=C21)O)O)C3=CC=CC=C3
26	3382	C=CCN1CCC2=CC(=C(C=C2C(C1)C3=CC=CC=C3)O)O
27	3383	C1CNCC(C2=CC(=C(C(=C21)Cl)O)O)C3=CC=CC=C3
28	3386	C1C(C1N)C2=CC=CC=C2
29	3387	CC1C(CCN1CC2=CC=CC=C2)NC(=O)C3=CC(=C(C=C3OC)NC)Cl
30	6981	CCCN(CCC)C1CCC2=C(C1)C=C(C=C2)O
31	6984	CC(C)(C)NCC(COC1=CC=CC2=C1SCCC2)O
32	11304	CC1=C(C(=O)N2CCCC(C2=N1)O)CCN3CCC(CC3)C4=NOC5=C4C=CC(=C5)F
33	11306	C1CCN(CC1)C2(CCN(CC2)CCCC(=O)C3=CC=C(C=C3)F)C(=O)N
34	11307	C1CN(CCC12C(=O)NCN2C3=CC=CC=C3)CCCC(C4=CC=C(C=C4)F)C5=CC=C(C=C5)F
35	11310	CC1=CC2=C(NC3=CC=CC=C3N=C2S1)N4CCN(CC4)C
36	11311	C1CN(CCN1CCOCCO)C2=NC3=CC=CC=C3SC4=CC=CC=C42
37	11312	CN1CC2C(C1)C3=C(C=CC(=C3)Cl)OC4=CC=CC=C24
38	11313	C1CN(CCC1C2=CN(C3=C2C=C(C=C3)Cl)C4=CC=C(C=C4)F)CCN5CCNC5=O
39	11314	C1CN(CCN1CCC2=C(C=C3C(=C2)CC(=O)N3)Cl)C4=NSC5=CC=CC=C54
40	11490	C1CN(CCC12C(=O)NCN2C3=CC=CC=C3)CC4COC5=CC=CC=C5O4
41	11861	CCCN(CCC)C1CCC2=C(C1)C(=CC=C2)O
42	11862	COC(=O)C1C(CCC2C1CC3C4=C(CCN3C2)C5=CC=CC=C5N4)O

43	11863	C1C2C(CN1)OC3=C(O2)C=CC=C3F
44	11912	CC1=C(C=CC(=C1)C2=NOC(=N2)C)C3=CC=C(C=C3)C(=O)NC4=CC(=C(C=C4)OC)N5CCN(CC5)C
45	11913	CC1=C(C=CC(=C1)C2=NOC(=N2)C)C3=CC=C(C=C3)C(=O)NC4=CC(=C(C=C4)OC)OCCN(C)C
46	14602	C1CC(CCC1CCN2CCC3=C(C2)C=CC(=C3)C#N)NC(=O)C4=CC=NC5=CC=CC=C45
47	15233	C1CN(CCN1CCCCN2C(=O)C3=CC=CC=C3S2(=O)=O)C4=NC=CC=N4.Cl
48	15237	CN1CC(CC2C1CC3=CN(C4=CC=CC2=C34)C)NS(=O)(=O)N(C)C
49	15238	C1CN(CCC1C(=O)C2=CC=C(C=C2)F)CCN3C(=O)C4=CC=CC=C4NC3=O
50	15245	C1=CC2=C(C=C1C(=O)N)C(=CN2)CCN
51	15246	COC1=CC=CC=C1N2CCN(CC2)CCCCN3C(=O)C4=CC=CC=C4C3=O.Br
52	15251	C1CN(CCN1CCNC(=O)C2=CC=C(C=C2)F)C3=C4C(=CC=C3)OC(CO4)CO
53	15254	CC(C)NCC(COC1=CC=CC2=C1C=CN2)O
54	15256	CNS(=O)(=O)CC1=CC2=C(C=C1)NC=C2CCN(C)C
55	15975	C1CCC2(C1)CC(=O)N(C(=O)C2)CCCCN3CCN(CC3)C4=NC=CC=N4
56	15982	CC(C)(C)C1(CCN2CC3C4=CC=CC=C4CCC5=C3C(=CC=C5)C2C1)O
57	15983	CN1CCN(CC1)C2CC3=CC=CC=C3SC4=C2C=C(C=C4)SC
58	15985	C1CN(CCC12C(=O)NCN2C3=CC=CC=C3)CCCC(=O)C4=CC=C(C=C4)F
59	16351	C1CN(CCN1)C2=NC3=CC=CC=C3C=C2
60	16353	CC1=CC(=C(C=C1OC)CC(C)N)OC
61	16356	CC1=CC=CN2C1=NC(=C(C2=O)CCN3CCC(CC3)C4=NOC5=C4C=CC(=C5)F)C
62	16358	C1CC2=C(C=CC(=C2)F)OC1C(CNCC(C3CCC4=C(O3)C=CC(=C4)F)O)O
63	16359	CN1C=C(N=C1C2=CC=C(C=C2)OCC(CNCCOC3=CC(=C(C=C3)O)C(=O)N)O)C(F)(F)F
64	16482	CC(CF)NCC(COC1=CC=CC2=C1C3=CC=CC=C3N2)O

65	16571	CN(C)CCC1=CNC2=C1C(=CC=C2)O
66	17031	CC(=O)C1=CC(=C(C=C1)OCCCN2CCC(CC2)C3=NOC4=C3C=CC(=C4)F)OC
67	21889	CCN(CC)C(=O)C1CN(C2CC3=CNC4=CC=CC(=C34)C2=C1)C
68	23627	CC(CN1C2=CC=CC=C2SC3=C1C=C(C=C3)C#N)CN(C)C
69	24816	CN(C)CCCN1C2=CC=CC=C2SC3=C1C=C(C=C3)Cl
70	25006	CN1CCN(CC1)C2=NC3=CC=CC=C3OC4=C2C=C(C=C4)Cl
71	25141	CC1=C(C(=O)N2CCCCC2=N1)CCN3CCC(CC3)C4=NOC5=C4C=CC(=C5)F
72	25270	CN1CCN(CC1)CCC=C2C3=CC=CC=C3SC4=C2C=C(C=C4)S(=O)(=O)N(C)C
73	25326	CN(C)CCOC1=CC2=CC=CC=C2SC3=C1C=C(C=C3)Cl
74	25378	C1CC(=O)NC2=C1C=CC(=C2)OCCCCN3CCN(CC3)C4=C(C(=CC=C4)Cl)Cl
75	26890	CCCCCCOC1=NSN=C1C2=CCCN(C2)C
76	27135	C1=CC2=C(C=C1O)C(=CN2)CCN
77	27136	CN1C=C(C2=C1C=CC3=C2CCCO3)CCN
78	27142	CC(CC1=CC(=C(C=C1OC)I)OC)N
79	27405	CN1CCC2=CC=CC3=C2C1CC4=C3C(=C(C=C4)O)O
80	27407	CCNC(=O)N(CCCN(C)C)C(=O)C1CC2C(CC3=CNC4=CC=CC2=C34)N(C1)CC=C
81	27408	CCN(CC)C(=O)NC1CN(C2CC3=CNC4=CC=CC(=C34)C2=C1)C
82	27409	CCCN1CC(CC2C1CC3=CNC4=CC=CC2=C34)CSC.CS(=O)(=O)O
83	27410	C1CN(CCN1CC2=CC3=C(C=C2)OCO3)C4=NC=CC=N4
84	27411	CCCNC1CCC2=C(C1)SC(=N2)N
85	27414	CCCN(CCC)CCC1=C2CC(=O)NC2=CC=C1
86	27415	C1CN(CC=C1C2=CC=CC=C2)CCCCC3=CNC4=C3C=C(C=C4)O

87	27416	C=CCN1CCC2=C(CC1)SC(=N2)N.Cl.Cl
88	27417	CCN(CC)C(=O)NC1CC2C(CC3=CNC4=CC=CC2=C34)N(C1)C
89	27820	CCCN(CCC)C1CCC2=C(C1)C=CC(=C2)CS(=O)(=O)C3=CC=C(C=C3)OC
90	28099	CCNC(C)CC1=CC(=CC=C1)C(F)(F)F
91	28153	CCC(CO)NC(=O)C1CN(C2CC3=CN(C4=CC=CC(=C34)C2=C1)C)C
92	28315	CC(CC1=CC(=CC=C1)C(F)(F)F)N
93	28369	CC1(C(=O)N2C(C(=O)N3CCCC3C2(O1)O)CC4=CC=CC=C4)NC(=O)C5CN(C6CC7=CNC8=CC=CC(=C78)C6=C5)C
94	28618	CCC1=C(NC2=C1C(=O)C(CC2)CN3CCOCC3)C
95	28620	C1CN(CCN1CCCN2C3=CC=CC=C3SC4=C2C=C(C=C4)Cl)CCO
96	28625	CN1CCCCC1CCN2C3=CC=CC=C3SC4=C2C=C(C=C4)SC
97	28627	CN1CCN(CC1)CCCN2C3=CC=CC=C3SC4=C2C=C(C=C4)C(F)(F)F
98	32488	CCC1=CC2=C(O1)C=CC3=C2N(N=C3)CC(C)N
99	36772	CN1CCN(CC1)C2=C3C=CC=CC3=NC4=C(N2)C=C(C=C4)Cl
100	38854	C1CN(CCC1C(C2=CC=CC=C2)O)CCC3=CC=CC=C3
101	38855	CN(C)CCONC(=CC=C1C=CC(=O)C=C1)C2=CC=CC=C2F.CN(C)CCONC(=CC=C1C=CC(=O)C=C1)C2=CC=CC=C2F.C(=CC(=O)O)C(=O)O
102	39352	CCCN(CCC)CCC1=CNC2=CC=CC=C21
103	41610	C1CN(CCN1)C2=C3C=CC=CC3=NC4=C(N2)C=C(C=C4)Cl
104	47317	COC1=CC=CC=C1N2CCN(CC2)CCN(C3=CC=CC=N3)C(=O)C4CCCCC4.Cl.Cl.Cl
105	48236	COC1=C(C=C2C3CC4=C(CN3CCC2=C1)C(=C(C=C4)O)OC)O

Table S5. The 61 5-HT_{1A} non-actives from the PDSP database employed in the QSAR modeling studies.

Cmpds ID	PDSP ID	SMILES
106	3337	O=C(N)N1c2c(C=Cc3c1cccc3)cccc2
107	3340	Clc1cc2N(c3c(CCc2cc1)cccc3)CCCN(C)C
108	3343	N(C(Cc1ccccc1)C)(CC#C)C
109	3344	N(CCCN1c2c(CCc3c1cccc3)cccc2)C
110	3353	O=C1c2c(n(c3c2cccc3)C)CCC1Cn1ccnc1C
111	3355	O(C)c1cc2c3CCN4C5C(CC(CC5CC)C4)c3[nH]c2cc1
112	3356	O(C(=O)c1c2c([nH]c1)cccc2)C1CC2N(C(C1)CC2)C
113	3358	N(CCCN1c2c(CCc3c1cccc3)cccc2)(C)C
114	3359	Fc1ccc(cc1)C(=O)C1CCN(CC1)CCN1C(=O)c2c(NC1=O)cccc2
115	3360	Clc1cc(cc(Cl)c1)C(OC1CC2N(C(C1)CC2)C)=O
116	3361	O1c2cc(ccc2OC1)CC(N)C
117	3363	N(C(Cc1ccccc1)C)C
118	3365	O(C(=O)C(C1NCCCC1)c1ccccc1)C
119	3366	Clc1cc(C(=O)NCCN(CC)CC)c(OC)cc1N
120	3370	Oc1cc(ccc1O)C(O)CN
121	3376	n1c2c(ccc1N1CCNCC1)cccc2
122	3385	S(=O)(=O)(N)c1cc(C(=O)NCC2N(CCC2)CC)c(OC)cc1
123	3389	Brc1ccc(cc1)\C(=C\CN(C)C)\c1cccnc1
124	16360	o1cccc1C(=O)N1CCN(CC1)c1nc(N)c2cc(OC)c(OC)cc2n1
125	16362	O(C)c1ccc(cc1)CN(CCN(C)C)c1ncccc1

126	16797	O(C)c1ccc(cc1)C(CN(C)C)C1(O)CCCCC1
127	16799	Fc1ccc(cc1)C1(OCc2c1ccc(c2)C#N)CCCN(C)C
128	16800	O1CCNCC1C(Oc1ccccc1OCC)c1ccccc1
129	22224	O1C(CC2(C3C(CCC2C1=O)(C)C(CC(OC(=O)C)C3=O)C(OC)=O)C)c1ccoc1
130	23143	SCCC(N)C(O)=O
131	23144	SCC(N)C(O)=O
132	23146	S1CCC(N)C1=O
133	23149	S(O)(=O)CC(N)C(O)=O
134	23150	S(O)(=O)(=O)CCC(N)C(O)=O
135	23153	S(O)(=O)(=O)CC(N)C(O)=O
136	23154	S(SCC(N)C(O)=O)CC(N)C(O)=O
137	23157	S(CCC(N)C(O)=O)C
138	25643	O(C(=O)c1ccccc1)C1CC2N(C(CC2)C1C(OC)=O)C
139	25815	OC(C(NC)C)c1ccccc1
140	25970	OC(C(N)C)c1ccccc1
141	26038	O=C(C(N)C)c1ccccc1
142	26069	O=C(C(NC)C)c1ccccc1
143	26728	Oc1c2c3c(c4c5c(C(=O)c6c7c(c8c(c(C2=O)c(O)cc8O)c3c57)c(O)cc6O)c(O)cc4C)c(c1)CO
144	26730	OC=1C2(CC(C\C=C(\C)/C)C(CC\C=C(\C)/C)(C)C(C(=O)C(C)C)(C2=O)C(=O)C=1C\C=C(\C)/C)C\C=C(\C)/C
145	26731	O1C(CO)C(O)C(O)C(O)C1OC1=C(OC=2C(=C1O)C(O)=CC(=O)C=2)c1cc(O)c(O)cc1
146	26732	O1C(CO)C(O)C(O)C(O)C1OC1=C(OC=2C(=C1O)C(O)=CC(=O)C=2)c1cc(O)c(O)cc1
147	26733	O1C(COC2OC(C)C(O)C(O)C2O)C(O)C(O)C(O)C1OC1=C(Oc2c(C1=O)c(O)cc(O)c2)c1cc(O)c(O)cc1

148	26734	O1c2c(C(=O)C(O)=C1c1cc(O)c(O)cc1)c(O)cc(O)c2
149	26735	O1C(C)C(O)C(O)C(O)C1OC1=C(OC=2C(=C1O)C(O)=CC(=O)C=2)c1cc(O)c(O)cc1
150	26736	O1c2c(C(=O)C=C1c1ccc(O)cc1)c(O)cc(O)c2-c1cc(ccc1O)C=1Oc2c(C(=O)C=1)c(O)cc(O)c2
151	27412	n1c2CC3C(N(CCC3)CCC)Cc2cnc1N
152	27666	O1c2cc(ccc2OC1)CC(NC)C
153	31276	Fc1ccc(cc1)CCN1CCC(CC1)C(O)c1cccc(OC)c1OC
154	33211	Clc1cccc(Cl)c1Nc1ccccc1CC(=O)[O-]
155	33247	S(=O)(=O)(Nc1ccc([N+](=O)[O-])cc1Oc1ccccc1)C
156	33287	O(C(=O)C)c1ccccc1C(O)=O
157	33323	S1(=O)(=O)N(C)\C(=C(/O)\Nc2ncccc2)\C(=O)c2c1cccc2
158	33400	OC(=O)C(C)c1ccc(cc1)CC(C)C
159	33516	O(C)c1cc2c(cc(cc2)C(C(O)=O)C)cc1
160	33556	S(=O)(=O)(N)c1ccc(-n2nc(cc2-c2ccc(cc2)C)C(F)(F)F)cc1
161	33595	Clc1cc(c(nc1)-c1ccc(nc1)C)-c1ccc(S(=O)(=O)C)cc1
162	33631	S(=O)(=O)(NC(=O)CC)c1ccc(cc1)-c1c(noc1C)-c1ccccc1
163	33671	S(=O)(=O)(C)c1ccc(cc1)C=1COC(=O)C=1c1ccccc1
164	33708	S(=O)(=O)(N)c1ccc(cc1)-c1c(noc1C)-c1ccccc1
165	33745	Clc1ccc(cc1)C(=O)n1c2c(cc(OC)cc2)c(CC(=O)NCCc2ccc(NC(=O)C)cc2)c1C
166	48248	O(c1ccccc1N(Cc1ccccc1OC)C(=O)C)c1ccccc1

Table S6. The 66 5-HT_{1A} actives used as the independent validation set (extracted from WOMBAT).

Cmpds ID	WOMBAT ID	SMILES
1	552	O1c2c(OCC1)cccc2OCCN1CCC(=CC1)c1c2c([nH]c1)cccc2
2	553	Fc1cc2c([nH]cc2C=2CCN(CC=2)CCOc2c3OCCOc3ccc2)cc1
3	554	Fc1cc2[nH]cc(c2cc1)C=1CCN(CC=1)CCOc1c2OCCOc2ccc1
4	555	Fc1cc2c([nH]cc2CC2CCN(CC2)CCOc2c3OCCOc3ccc2)cc1
5	556	Fc1cc2[nH]cc(c2cc1)C1CCN(CC1)CCOc1c2OCCOc2ccc1
6	557	O1c2c(OCC1)cccc2OCCN1CCC(=CC1)c1c2cccnc2[nH]c1
7	558	O(CCN1CCC(=CC1)c1c2c([nH]c1)cccc2)c1c2c([nH]cc2)ccc1
8	559	O(CCN1CCC(=CC1)c1c2cccnc2[nH]c1)c1c2c([nH]cc2)ccc1
9	560	O(CCN1CCC(=CC1)c1c2c([nH]c1)cccc2)c1c2c(nccc2)ccc1
10	561	O(CCN1CCC(=CC1)c1c2cccnc2[nH]c1)c1c2c(nccc2)ccc1
11	2800	Clc1cc2c(sc(S(=O)(=O)Nc3cc(N4CCNCC4)c(OC)cc3)c2C)cc1
12	2801	Clc1cc2c(sc(S(=O)(=O)Nc3cc(N4CCN(CC4)C)c(OC)cc3)c2C)cc1
13	2802	Clc1cc2c(sc(S(=O)(=O)Nc3cc(N4CC5N(CCCC5)CC4)c(OC)cc3)c2C)cc1
14	2804	Clc1cc2c(sc(S(=O)(=O)Nc3cc(N4CCN5[C@H](CCC5)C4)c(OC)cc3)c2C)cc1
15	28425	O(C)c1cc2c([nH]cc2CCN)cc1
16	28426	O(CCC)c1cc2c([nH]cc2CCN)cc1
17	28427	O(CCCCC)c1cc2c([nH]cc2CCN)cc1
18	28428	O(CCCCCCC)c1cc2c([nH]cc2CCN)cc1
19	28429	O(CCCCCCCC)c1cc2c([nH]cc2CCN)cc1
20	28430	O(CCCCCCCCC)c1cc2c([nH]cc2CCN)cc1

21	28431	O(CCCCCCCCCC)c1cc2c([nH]cc2CCN)cc1
22	28432	O(CCCCCCCCCCC)c1cc2c([nH]cc2CCN)cc1
23	28434	O(CCCCCCC(C)C)c1cc2c([nH]cc2CCN)cc1
24	28435	O(CCCCCCCC(C)(C)C)c1cc2c([nH]cc2CCN)cc1
25	28436	O(CCCCCCC(C)(C)C)c1cc2c([nH]cc2CCN)cc1
26	28437	O(CCCCC1CCCCC1)c1cc2c([nH]cc2CCN)cc1
27	28438	O(CCCCCCCCC)c1cc2c([nH]cc2CCN(C)C)cc1
28	28439	O(CCCCCCCCC)c1cc2c([nH]cc2CCNCCCC)cc1
29	28443	[nH]1cc(c2cc(ccc12)CC)CCN
30	28444	[nH]1cc(c2cc(ccc12)CC)CCN(C)C
31	28445	[nH]1cc(c2cc(ccc12)CCCCCCCCCC)CCN
32	28446	O=C(C)c1cc2c([nH]cc2CCN)cc1
33	28448	O(C(=O)CC)c1cc2c([nH]cc2CCN(C)C)cc1
34	28451	O(C(=O)CCCCCCCC)c1cc2c([nH]cc2CCN(C)C)cc1
35	34398	S(Oc1cc2c([nH]cc2CCN(CCC)CCC)cc1)(=O)(=O)C(F)(F)F
36	34399	S(Oc1cc2c([nH]cc2CCN(CC)CC)cc1)(=O)(=O)C(F)(F)F
37	34400	S(Oc1cc2c([nH]cc2CCN(C)C)cc1)(=O)(=O)C(F)(F)F
38	34401	S(Oc1cc2c([nH]cc2CCN)cc1)(=O)(=O)C(F)(F)F
39	34402	O(C)c1ccc(cc1)CNC(=O)\C=C\c1cc2c([nH]cc2CCN(C)C)cc1
40	44145	O1CCc2c(cccc2)C1CCN1CCN(CC1)c1ccc(OC)cc1
41	44146	O1CCc2cc(ccc2[C@@H]1CCN1CCN(CC1)c1ccc(OC)cc1)C(=O)N
42	44148	O1CCc2cc(ccc2[C@@H]1CCN1CCN(CC1)c1ccc(OC)cc1)C(=O)NC

43	49513	Brc1ccc(S(=O)(=O)Nc2cc(N3CCN(CC3)C)c(OC)cc2)cc1
44	49956	O=C1Nc2c3c(CCCC13CCCCN1CCC(=CC1)c1ccccc1)ccc2
45	52802	O1c2c(OCC1CNCCOc1c(OC)cccc1OC)cccc2
46	52803	O1c2c(cccc2)C(=O)[C@@H](CNCCOc2c(OC)cccc2OC)[C@@H]1c1ccccc1
47	52804	S1c2c(cccc2)C(=O)[C@@H](CNCCOc2c(OC)cccc2OC)[C@@H]1c1ccccc1
48	52805	O(CCNC[C@H]1[C@@H](Cc2c(cccc2)C1=O)c1ccccc1)c1c(OC)cccc1OC
49	52807	O1c2c(C[C@H]([C@@H]1CNCCOc1c(OC)cccc1OC)c1ccccc1)cccc2
50	52808	O1c2c(C=C(CNCCOc3c(OC)cccc3OC)C1c1ccccc1)cccc2
51	52809	O(CCNCCOc1ccccc1OCc1ccccc1)c1c(OC)cccc1OC
52	52810	O(C)c1ccccc1N1CCN(CC1)CCN1C(=O)CC2(CC1=O)CCCC2
53	54082	O1c2c(OCC1CN[C@@H]1CC[C@@H](Oc3ccccc3OC)[C@@H]1O)cccc2
54	54084	O1c2c(OCC1CN[C@@H]1CC[C@@H](Oc3c(OC)cccc3OC)[C@@H]1O)cccc2
55	56453	S(=O)(=O)(N1CCCC[C@@H]1CCN1CCC(CC1)C)c1c2c(ccc1)cccc2
56	56455	S(=O)(=O)(N1CCC[C@@H]1CCN1CCC(CC1)C)c1c2c(ccc1)cccc2
57	56457	Clc1cc(S(=O)(=O)N2CCC[C@@H]2CCN2CCC(CC2)C)ccc1Cl
58	56458	Brc1cc(S(=O)(=O)N2CCC[C@@H]2CCN2CCC(CC2)C)ccc1
59	56459	S(=O)(=O)(N1CCC[C@@H]1CCN1CCC(CC1)C)c1cc(ccc1)C
60	57927	N1([C@@H]2Cc3c4-c5c(Cc4ccc3)ccc(CC1)c25)C
61	57933	O=C(N[C@@H]1c2c3-c4c1ccc1CCN([C@H](Cc3ccc2)c14)C)C
62	57936	N[C@H]1c2c3-c4c1ccc1CCN([C@H](Cc3ccc2)c14)C
63	57937	N1([C@@H]2Cc3c-4c(ccc3)[C@H](c3c-4c2c(CC1)cc3)C)C

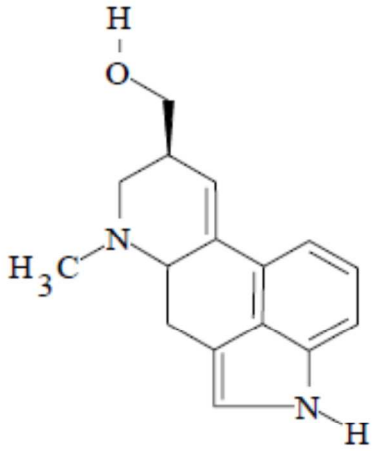
64	57940	O(C)[C@]1(c2c3-c4c1ccc1CCN([C@H](Cc3ccc2)c14)C)C
65	57941	O(CC)[C@@]1(c2c3-c4c1ccc1CCN([C@H](Cc3ccc2)c14)C)CC
66	57947	Fc1c2c(C[C@@H](N(CCC)CCC)CC2)c(O)cc1

Certificate of Analysis for compound Lysergol

TimTec Corporation
Harmony Business Park, Bldg 301-A,
Newark DE 19711, USA
Phone: 1-302-292-8500 Fax: 1-302-292-8520
http://www.timtec.net e-mail info@timtec.net

Certificate of Analysis

Product Name	Lysergol
IDNUMBER	ST057540



CAS Registry Number	602-85-7
Molecular Formula	$C_{16}H_{18}N_2O$
Molecular Weight	254.36
Original Storage Container	Screw top vial
Storage temperature	Ambient
Transportation Temperature	Ambient
PURITY	>96%