Indian Journal of Chemistry
Vol. 57B, October 2018, pp. 1322-1327

Simplified molecular-input line-entry system based quantitative structure–activity relationship (QSAR) models for serotonin 3 (5-HT_3) receptor

Sanija Begum[a], P Ganga Raju Achary*[a], Andrey A Toropov[b] & Alla P Toropova[b]

[a] Department of Chemistry, Institute of Technical Education and Research (ITER), Siksha 'O' Anusandhan University, Bhubaneswar 751 030, India

[b] IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Via La Masa 19, 20156 Milan, Italy

E-mail: pgrachary@soa.ac.in; pgrachary@gmail.com

Received 7 October 2017; accepted (revised) 3 July 2018

Anxiety, emesis and cognition are regulated significantly by the Serotonin 3 (5-HT_3) receptor, which are present in the central nervous system (CNS). Quantitative structure–activity relationship (QSAR) models have been built for the prediction of inhibitor constant (−log Ki) with the help 50 potent antagonists for the 5-hydroxytryptamine (5-HT_3) receptor. The –log (Ki) values have been modelled with simplified molecular input line entry system (SMILES) based molecular structures. The external validation characteristics such as, the randomization parameters like ${}^{c}r_p^2$, r_m^2 , $r^{*}_m{}^2$, average r_m^2 needs to be more than 0.5 and for Δr_m^2 if it is less than 0.2 then such models can be robust. All the five reported QSAR models have passed the external validation criteria test such as ${}^{c}r_p^2$, r_m^2, $r^{*}_m{}^2$, average r_m^2 test and have proved their robustness.

Keywords: Quantitative structure–activity relationship, 5-HT_3, SMILES, CNS, Monte Carlo method

The peripheral (PNS) and central (CNS) nervous system contain 5-HT3 receptors. Among the 5-HT (serotonin) receptors 5-HT3 receptors are ligand-gated ion channels (LGIC) which are mediated *via* G proteins, belongs to the Cys-loop family of ligand-gated ion channels. Around a central ion-conducting pore the five subunits of these 5-HT3 receptor are arranged symmetrically[1]. Activation of 5-HT_3 receptors can have significant effect on the number of sympathetic, parasympathetic and sensory functions[2,3]. The other functions where this 5-HT_3 play important role are: urinary tract, emesis, cognition and anxiety[4,5]. Also these receptors cab regulate the release of GABA, dopamine, cholecystokinin, *etc*.[6,7]

A good number of compounds are known which can act as antagonists for this receptor. These antagonists could be non selective compounds like morphine and cocaine[8] or selective compounds like bemesetron and tropisetron. The discovery of bemesetron and tropisetron led to the development of good number of potential compounds including ondansetron, granisetron and zacopride, which are known to have effect in the nanomolar concentrations (nM). The bioactivities of such potent molecules are available in the EBI database powered by ChEMBL with detailed bioassay report.

As we know the quantitative structure activity/property relationship (QSAR/QSPR) is a proven technique to predict the desired activities/ properties from the molecular properties and which works with the principle "*Endpoint=f* (*molecular property*)". Such models always require a sizable number of molecular properties, and we call them 'descriptors'. Although, the work of 5-HT3 receptors and the compounds which act as antagonists for this receptor is of prime importance as discussed above, there have been very little efforts in building the QSAR models. One such attempt is made CoMFA and CoMSIA techniques taking 20 thiazoles for binding affinities to the 5HT3 receptor[9].

In the present study QSAR models were built for the prediction of inhibitor constant (−log Ki) with the help 50 potent antagonists for the 5-Hydroxytryptamine (5-HT_3) receptor. The −log Ki values were modelled with simplified molecular input line entry system (SMILES) based molecular structures and five random splits of the data. These models were built with the usual paradigms such as "*Endpoint=f (SMILES*)[10,11]. The reported models validated by external validation techniques and with a set of eight inhibitors as external validation set to check the robustness of these models.

Methods

Data

The datasets of 50 inhibitors with Ki (nM) values for the Serotonin 3 (5-HT3) receptor was obtained from the EBI database (https://www.ebi.ac.uk/chembl/malaria/target/inspect/CHEMBL2094116). The potency of an inhibitor could be expressed in terms of Ki; which is defined as the minimum concentration required to have half of the maximum inhibition. More specifically it is the reflective of the binding affinity of the inhibitor for a drug. The –log (Ki) dataset of 50 antagonists with their molecular structure (SMILES) are given in the Table I.

Table I — Experimental and Predicted inhibition constant (−log (Ki) QSAR model-1, model-2, model-3, model-4 and model-5

ID	SMILES	Expt −log(Ki)	Pred. –log(Ki) M-I	Pred. −log(Ki) M-II	Pred. −log(Ki) M-III	Pred. −log(Ki) M-IV	Pred. −log(Ki) M-V
1	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3cc(Cl)cc(Cl)c3	7.27	7.55	7.51	7.45	7.53	7.51
2	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.61	8.28	8.24	8.27	8.28	8.30
3	#CNS(=O)(=O)Cc1ccc2[nH]cc(CCN(C)C)c2c1	6.00	5.00	5.00	5.50	5.00	5.00
4	+COc1cc(N)c(Cl)cc1C(=O)OCCN2CCCCC2	6.14	6.12	6.12	6.17	6.12	6.12
5	+C1CN(CCN1)c2ccc3ccccc3n2	8.52	8.70	8.72	8.73	8.70	8.72
6	+[O-][N+](=O)c1ccc2nc(ccc2c1)N3CCNCC3	7.24	7.24	7.24	7.24	7.24	7.24
7	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.12	8.28	8.24	8.27	8.28	8.30
8	+C1CN(CCN1)c2ccc3ccccc3n2	8.85	8.70	8.72	8.73	8.70	8.72
9	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.21	8.28	8.24	8.27	8.28	8.30
10	+C1CN(CCN1)c2ccc3ccccc3n2	8.92	8.70	8.72	8.73	8.70	8.72
11	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	9.00	8.28	8.24	8.27	8.28	8.30
12	#C1CN(CCN1)c2ccc3ccccc3n2	8.74	8.70	8.72	8.73	8.70	8.72
13	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.80	8.28	8.24	8.27	8.28	8.30
14	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.89	8.78	8.78	8.69	8.78	8.78
15	+CN1CCN2C(C1)c3ccccc3Cc4ccccc24	7.15	7.24	7.21	7.18	7.22	7.14
16	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.31	6.41	6.42	6.44	6.41	6.41
17	+COc1cc(N)c(Cl)cc1C(=O)NC2CCN(Cc3ccccc3)CC2	6.00	5.96	5.97	6.05	5.96	6.00
18	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.48	8.28	8.24	8.27	8.28	8.30
19	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.57	8.78	8.78	8.69	8.78	8.78
20	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.80	8.28	8.24	8.27	8.28	8.30
21	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	9.11	8.28	8.24	8.27	8.28	8.30
22	#CCN(CC)CCOC(=O)c1cc(Cl)c(N)cc1OC	6.69	6.05	6.12	6.28	5.94	6.14
23	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.00	6.41	6.42	6.44	6.41	6.41
24	+COc1cc(N)c(Cl)cc1C(=O)NC2CCN(Cc3ccccc3)CC2	6.00	5.96	5.97	6.05	5.96	6.00
25	+COc1cc(N)c(Cl)cc1C(=O)OCCN2CCCCC2	6.11	6.12	6.12	6.17	6.12	6.12
26	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3cc(Cl)cc(Cl)c3	7.61	7.55	7.51	7.45	7.53	7.51
27	#CN1C2CCC1CC(C2)NC(=O)c3cc(Cl)cc4[nH]cnc34	7.78	8.54	7.85	7.63	7.80	7.69
28	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.19	8.28	8.24	8.27	8.28	8.30
29	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.62	6.41	6.42	6.44	6.41	6.41
30	#COc1cc(N)c(Cl)cc1C(=O)N[C@H]2CCN3CCC[C@@H]2C3	8.28	8.66	8.40	8.17	8.51	8.35
31	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3cc(Cl)cc(Cl)c3	7.58	7.55	7.51	7.45	7.53	7.51
32	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.89	8.78	8.78	8.69	8.78	8.78
33	+CCOc1cc(N)c(Cl)cc1C(=O)NCC2CN(Cc3ccc(F)cc3)CCO2	5.92	5.92	5.92	5.92	5.92	5.92
34	#[O-][N+](=O)c1ccc2nc(ccc2c1)N3CCNCC3	7.24	7.24	7.24	7.24	7.24	7.24
35	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.68	6.41	6.42	6.44	6.41	6.41
36	#CN1CCN(CC1)c2ccc3ccccc3n2	8.52	8.66	8.61	8.56	8.68	8.58
37	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3cc(Cl)cc(Cl)c3	7.68	7.55	7.51	7.45	7.53	7.51
38	#C1CN(CCN1)c2ccc3ccccc3n2	8.74	8.70	8.72	8.73	8.70	8.72
39	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	7.79	8.28	8.24	8.27	8.28	8.30
40	+CNS(=O)(=O)Cc1ccc2[nH]cc(CCN(C)C)c2c1	5.00	5.00	5.00	5.50	5.00	5.00
41	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.35	6.41	6.42	6.44	6.41	6.41
42	#CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.43	8.78	8.78	8.69	8.78	8.78
43	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	5.47	8.28	8.24	8.27	8.28	8.30

(*Contd.*)

Table I — Experimental and Predicted inhibition constant (−log (Ki) QSAR model-1, model-2, model-3, model-4 and model-5

(*Contd*.)

ID	SMILES	Expt −log(Ki)	Pred. −log(Ki)	Pred. −log(Ki)	Pred. −log(Ki)	Pred. −log(Ki)	Pred. −log(Ki)
44	#Cl^CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc 34~	8.42	8.67	8.82	8.42	8.63	8.86
45	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.92	8.28	8.24	8.27	8.28	8.30
46	+CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.46	6.41	6.42	6.44	6.41	6.41
47	+CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.80	8.78	8.78	8.69	8.78	8.78
48	+CN1CCN(CC1)c2ccc3ccccc3n2	8.52	8.66	8.61	8.56	8.68	8.58
49	+CN1CCN(CC1)C2=Nc3cc(Cl)ccc3Nc4ccccc24	7.28	7.21	7.21	7.13	7.25	7.33
50	+Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.12	8.28	8.24	8.27	8.28	8.30

Results and Discussion

The CWs of different S_k for the five random splits are listed in Table II. The optimal descriptor is actually the summation of all CW (S_k) present in the SMILES string of the given molecule. Let us take an example 'Bemesetron' with IUPAC name '[(1S,5R)-8-methyl-8-azabicyclo[3.2.1]octan-3-yl] 3,5-dichlorobenzoate' which is our first inhibitor with ID-1,the SMILES code is: CN1[C@@H]2CC[C@H] 1C[C@H](C2)OC(=O)c3cc(Cl)cc(Cl)c3.

The illustration to obtain DCW is represented in the Table III for the molecule 'Bemesetron' and it was found to be -3.806, 0.153,-8.449, 3.919 and 2.567 respectively for the above five QSAR models to predict –log Ki for 5-HT_3 receptor.

QSAR models

A one descriptor QSAR model can take Equation (2) form with C_0, C_1, T and N being the intercept, slope, and threshold and is number of epochs respectively.

$$-\log (Ki) = C_0 + C_1 * DCW (T, N) \quad (2)$$

The five SMILES-based QSAR models for random split-1 to split-5 are given below:

I. −log(Ki) = 8.306 (± 0.02) + 0.198 (± 0.001) * DCW(1,30)

II. −log(Ki) = 7.480 (± 0.012) + 0.175 (± 0.001) * DCW(1,30)

III. −log(Ki) = 9.338 (± 0.027) + 0.224 (± 0.002) * DCW(1,30)

IV. −log(Ki) = 6.915 (± 0.008) + 0.156 (± 0.001) * DCW(1,30)

V. −log(Ki) = 8.096 (± 0.017) + 0.173 (± 0.001) * DCW(1,30)

It is observed in QSAR models I– V, the intercept (C_0) value lies between 6.9 and 9.3 and slope (C1) value between 0.156 and 0.224. The fitting data of the above five models for the 'T' between 1 and 5 is given in the Table IV, which was built by adopting classic scheme method present in the Monte Carlo optimization process.

Table II — CWs for SMILES attribute (Sk) for the five random splits

Sk	$CW(S_k)$ M-1	$CW(S_k)$ M-2	$CW(S_k)$ M-3	$CW(S_k)$ M-4	$CW(S_k)$ M-5
(	−1.060	−1.442	−0.998	−1.627	−1.565
+	−4.254	−4.058	−2.505	−2.818	−4.504
–	−3.130	−4.684	−1.691	−4.000	−3.124
1	2.061	4.372	−0.315	5.562	5.127
2	−0.809	−1.129	−0.873	−1.000	−1.250
3	3.622	2.310	2.254	2.750	2.378
4	0.005	−0.502	−0.058	−0.816	−0.380
=	−1.311	−2.059	−1.687	−1.378	−1.684
@	0.379	1.183	3.059	4.560	3.123
@@	2.816	5.003	2.942	4.373	4.059
C	−0.187	−0.628	−0.754	−0.127	−0.747
F	1.939	2.501	0.942	3.254	2.191
H	−0.250	−0.189	0.872	−0.753	0.066
Cl	−0.558	0.186	−0.433	−0.939	0.435
N	1.933	2.375	1.942	2.878	2.685
O	0.129	0.690	1.252	−0.126	1.254
S	−3.999	−5.317	−1.559	−5.192	−4.001
[	0.317	0.374	−0.812	−0.126	−0.186
c	−1.315	−0.814	−0.752	−1.066	−0.745
n	3.063	3.935	3.060	4.375	4.005

The experimental and predicted −log (Ki) for the four models are given in the Table I and the statistical characteristics of these five models are listed in Table IV for each training and validation set. All these reported models displayed good statistical characteristics such as for the training set correlation coefficients are around~0.80, and for validation set it is more than 0.92. Cross-validated correlation coefficients(*Q)* lies between 0.784 and 0.960 ; standard error of estimation(s) lies between 0.238 and 0.359; *MAE* is mean absolute error (MAE) lies between 0.238 and 0.359 and Fischer F-ratio (f) lies between 62 and 282. The lowest 'f' obtained for the

Table III — Calculation of optimal Descriptor for Bemesetron from the CW_S for the five random splits

Sk	CW(Sk) M-1	CW(Sk) M-2	CW(Sk) M-3	CW(Sk) M-4	CW(Sk) M-5
C	−0.1865	−0.628	−0.754	−0.127	−0.747
N	1.9325	2.375	1.9415	2.878	2.6845
1	2.0605	4.372	−0.3145	5.5615	5.127
[	0.3165	0.374	−0.8115	−0.126	−0.1855
C	−0.1865	−0.628	−0.754	−0.127	−0.747
@@	2.8155	5.003	2.9415	4.373	4.0585
H	−0.25	−0.1885	0.872	−0.753	0.0655
[	0.3165	0.374	−0.8115	−0.126	−0.1855
2	−0.8085	−1.129	−0.873	−1	−1.25
C	−0.1865	−0.628	−0.754	−0.127	−0.747
C	−0.1865	−0.628	−0.754	−0.127	−0.747
[	0.3165	0.374	−0.8115	−0.126	−0.1855
C	−0.1865	−0.628	−0.754	−0.127	−0.747
@	0.379	1.1825	3.0585	4.5595	3.123
H	−0.25	−0.1885	0.872	−0.753	0.0655
[	0.3165	0.374	−0.8115	−0.126	−0.1855
1	2.0605	4.372	−0.3145	5.5615	5.127
C	−0.1865	−0.628	−0.754	−0.127	−0.747
[	0.3165	0.374	−0.8115	−0.126	−0.1855
C	−0.1865	−0.628	−0.754	−0.127	−0.747
@	0.379	1.1825	3.0585	4.5595	3.123
H	−0.25	−0.1885	0.872	−0.753	0.0655
[	0.3165	0.374	−0.8115	−0.126	−0.1855
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
C	−0.1865	−0.628	−0.754	−0.127	−0.747
2	−0.8085	−1.129	−0.873	−1	−1.25
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
O	0.129	0.6895	1.252	−0.126	1.254
C	−0.1865	−0.628	−0.754	−0.127	−0.747
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
=	−1.3105	−2.0585	−1.6865	−1.378	−1.6835
O	0.129	0.6895	1.252	−0.126	1.254
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
3	3.622	2.3095	2.254	2.75	2.378
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
C	−0.5575	0.1855	−0.4325	−0.9385	0.4345
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
C	−0.5575	0.1855	−0.4325	−0.9385	0.4345
(	−1.0595	−1.4415	−0.998	−1.627	−1.5645
c	−1.3145	−0.8135	−0.752	−1.0655	−0.745
3	3.622	2.3095	2.254	2.75	2.378
∑	−3.806	0.153	−8.4495	3.919	2.567

validation set for the model-3, and for all models it 106. Though the predictability of these models appears to be good (Table I and Table IV) at this point, these models need to tested with some external validation characteristics to say that these models are robust for –log Ki for 5-HT$_3$ receptor. We know that, as rule, every model needs to be checked its prediction ability and reliability with dataset which is not used in building these models (to be treated as external validation set). So, a set of eight antagonists'

Table IV — QSAR model for −log (Ki) for 5-HT_3 receptor with the statistical parameters

Model no.	QSAR Models	Statistical Parameters						
		Set	*n*	r^2	Q^2	*s*	*MAE*	*f*
1	−log(Ki) = 8.306 (± 0.02) + 0.198 (± 0.001) * DCW(1,30)	Training	40	0.801	0.784	0.534	0.264	152
		Validation	10	0.929	0903	0.508	0.359	106
2	−log(Ki) = 7.480 (± 0.012) + 0.175 (± 0.001) * DCW(1,30)	Training	40	0.805	0.791	0.526	0.249	158
		Validation	10	0.934	0.909	0.467	0.322	114
3	−log(Ki) = 9.338 (± 0.027) + 0.224 (± 0.002) * DCW(1,30)	Training	40	0.804	0.787	0.527	0.263	157
		Validation	10	0.885	0.835	0.353	0.238	62
4	−log(Ki) = 6.915 (± 0.008) + 0.156 (± 0.001) * DCW(1,30)	Training	40	0.800	0.784	0.534	0.264	152
		Validation	10	0.972	0.960	0.448	0.285	282
5	−log(Ki) = 8.096 (± 0.017) + 0.173 (± 0.001) * DCW(1,30)	Training	40	0.802	0.787	0.522	0.246	154
		Validation	10	0.931	0.894	0.457	0.288	109

Table V — Set of eight antagonists dataset and its comparative data for the listed five models

IDs	External Validation set	Expt (−log(Ki))	Model-1 Pred. −log(Ki)	Model-2 Pred. −log(Ki)	Model-3 Pred. −log(Ki)	Model-4 Pred. −log(Ki)	Model-5 Pred. −log(Ki)
51	CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.35	6.412	6.415	6.317	6.41	6.407
52	Cc1[nH]c2ccc(O)cc2c1CCN	5.91	7.278	7.967	7.867	7.397	7.79
53	CCN(CC)CCNC(=O)c1cc(Cl)c(N)cc1OC	6.3	6.412	6.415	6.432	6.41	6.407
54	CN1CCN(CC1)c2oc3ccccc3n2	7.26	9.184	8.899	9.536	9.008	8.855
55	Cl^CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.68	8.666	8.815	8.775	8.629	8.862
56	CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.77	8.777	8.782	8.541	8.775	8.781
57	Cc1nccn1CC2CCc3c(C2=O)c4ccccc4n3C	8.21	8.282	8.237	8.372	8.281	8.295
58	CN1[C@@H]2CC[C@H]1C[C@H](C2)OC(=O)c3c[nH]c4ccccc34	8.49	8.777	8.782	8.699	8.775	8.781

Table VI — External validation characteristics for the SMILES based QSAR models for the prediction of inhibition constant (−log (Ki) in nM

QSAR models	r^2 *test set*	${}^c r_P^2$ S	${}^c r_P^2$ T	*k*	*kk*	r_m^2	$r^{*}{}_m^2$	r_m^2 *(Avg)*	Δr_m^2
1	0.9298	0.773	0.805	1.007	0.990	0.685	0.503	0.594	0.183
2	0.934	0.767	0.843	0.988	1.009	0.709	0.552	0.630	0.157
3	0.885	0.773	0.812	0.984	1.014	0.777	0.885	0.831	0.107
4	0.972	0.761	0.878	0.993	1.004	0.694	0.544	0.619	0.150
5	0.971	0.771	0.907	0.996	1.002	0.700	0.554	0.627	0.145

dataset from id 51 to 58 is chosen and the results obtained are summarized in Table V.

External validation characteristics

Y-scrambling technique could always be used to check the robustness of all the QSAR models. The randomization parameters like (${}^c r_p^2$, where ${}^c r_P^2 = r(r^2 - \bar{r}_r^2)^{1/2}$), r_m^2 , $r^{*}{}_m^2$, average r_m^2 needs to be more than 0.5 and for Δr_m^2 if it is less than 0.2 then such models can be robust[10–15]. If 'k' and 'kk' for a model is more than 0.85 but less than 1.15, then it can be another criteria to check robustness. All these external validation characteristics were explored for the five models for the –log(Ki) for the 5-HT_3, and listed in the Table VI. All the five reported QSAR models are passed this set of external validation criteria test and thus we can conclude that these models are robust for the 5HT_3.

Conclusion

Hence, it can be said in the concluding remark that the SMILES based QSAR model proved efficient to predict –log(Ki) for the inhibition of for Serotonin 3 (5-HT_3) receptor. All the reported five models exhibited good statistical characteristics and external validation characteristics. These models will help in designing and screening the drug candidates before they could be synthesized.

References

1 Thompson A J & Lummis S C R, *Curr Pharm Des*, 28 (2006) 3615.
2 Jackson M B & Yakel J L, *Annu Rev Physiol*, 57 (1995) 447.
3 Morales M & Wang S D, *J Neurosci*, 22 (2002) 6732.
4 Tecott L H, Maricq A V & Julius D, *Proc Natl Acad Sci* (*USA*), 90 (1993) 1430.
5 Bhattacharya A, Dang H, Zhu Q M, Schnegelsberg B, Rozengurt N, Cain G, Prantil R, Vorp D A, Guy N, Julius D, Ford A P D W, Lester H A & Cockayne D A, *J Neurosci*, 24 (2004) 5537.
6 Nayak S V, Ronde P, Spier A D, Lummis S C & Nichols R A, *Neurosci*, 91 (1999) 107.
7 Katsurabayashi S, Kubota H, Tokutomi N & Akaike N, *Neuropharmacology*, 44 (2003) 1022.
8 Gaddum J H & Picarelli Z P, *J Pharmacol*, 12 (1957) 323.
9 Zhu L-P, Ye D-Y & Tang Y, *J Mol Model*, 13 (2007) 121.
10 Achary P G R, *SAR QSAR Environ Res*, 25 (2014) 507.
11 Begum S & Achary P G R, *SAR QSAR Environ Res*, 26 (2015) 343.
12 Toropov A A, Achary P G R & Toropova A P, *Chem Phys Lett*, 660 (2016) 107.
13 Toropov A A & Toropova A P, *Chemosphere*, 104 (2014) 262.
14 Toropov A A & Toropova A P, *Chemosphere*, 124 (2015) 40.
15 Toropova A P & Toropov A A, *Mini Rev Med Chem*, 15 (2015) 608.