Electronic Supplementary Material (ESI) for RSC Advances.

Understanding Structural Requirements of Cyclic Sulfone Hydroxyethylamines as hBACE1 Inhibitors against Aβ Plaques in Alzheimer's Disease: A Predictive QSAR Approach

Pravin Ambure and Kunal Roy*

Drug Theoretics and Cheminformatics Laboratory,

Department of Pharmaceutical Technology,

Jadavpur University, Kolkata 700032, India

---

* Author for correspondence

E-mail: kunalroy_in@yahoo.com , kroy@pharma.jdvu.ac.in ;

URL: http://sites.google.com/site/kunalroyindia/

Supplementary Materials

A failed QAAR Model

As one can observe that the classical and/or R^2-based parameters (see the respective values mentioned in **Box** S1) show acceptable results for the reported QAAR model, irrespective of presence of high percentage of moderate prediction errors in both the training and test sets (see **Table** S1 for prediction error values). Further, the Golbraikh and Tropsha's criteria are also satisfied for the reported QAAR model. However, based on the MAE-based criteria, the prediction quality for both the training and test sets were categorized as '*Bad*' (see **Table** S2). Please note that the response range of the training set (i.e., *3.1832*) employed in QAAR study is lower than response range of the training set (i.e., 4.5185) employed in the 2D-QSAR and G-QSAR model, so the threshold for allowable prediction errors should be lower for QAAR compared to the 2D-QSAR and G-QSAR models. Further, for comparing the prediction quality, the distributions of prediction errors for all the three models (see **Table** S3). One can clearly observe that the percentage of compounds (*%NComp*) in the training as well as test sets with moderate-high errors (i.e., *%NComp*> 0.15×trainingSetRange) are considerably higher in the QAAR models as compared to the 2D-QSAR and G-QSAR models.

Box S1: QAAR Model (***Failed Model based on MAE-based criteria***)

$$pIC_{50}^{(hBACE1\text{-}hCatD)} = -0.7437(\pm 0.35012) + 0.33458(\pm 0.18781)B07[F-F] - 1.87395(\pm 0.4698)R1\_ETA\_dBetaP - 0.66285(\pm 0.2619)R2\_nArOR - 1.01635(\pm 0.31859)B07[O-S] + 1.63303(\pm 0.3309)F03[N-O] \quad \textbf{(S1)}$$

***Training Set*:** $N_{Training}$ =57; R^2 =0.7727, $R^2_{adj.}$ =0.7504, Q^2 =0.7354, $\overline{r^2_{m(training)}}$ =0.6376, $\Delta r^2_{m(training)}$ =0.1710

***Test Set*:** N_{Test}=25; R^2_{test} =0.7426, R^2_{pred} / Q^2_{F1} =0.7462 Q^2_{F2} =0.7290, *CCC*=0.880, $\overline{r^2_{m(test)}}$ =0.6555, $\Delta r^2_{m(test)}$ =0.0024

Golbraikh and Tropsha's Model Acceptability Criteria: *All criterions passed*

Summarizing the QAAR study, the QAAR model was reported only to demonstrate how the error-based metrics provides a true picture of the prediction quality of QSAR models in comparison to the misleading interpretation that can be deduced from classical and/or R^2-based metrics.

Table S1. The observed and predicted activity values [$\boldsymbol{pIC_{50}}$(hBACE1-hCatD)] along with the absolute prediction errors for all the compounds present in training (A) and test set (B) of QAAR model.

(A) TRAINING SET (QAAR Model)			
Compound	Observed	LOO* Predicted	Absolute

Number	$pIC_{50(\text{hBACE1-hCatD})}$	$pIC_{50(\text{hBACE1-hCatD})}$	Prediction Error
2	0.4809	0.2871	0.1938
3	-0.0607	-0.3637	0.3030
4	-0.0872	1.0439	1.1310
5	-0.0669	0.6368	0.7037
6	0.4771	0.3533	0.1238
7	0.0589	0.4734	0.4145
8	0.2430	0.7207	0.4776
9	0.2348	0.8307	0.5959
11	0.9765	0.7523	0.2243
13	0.6021	0.4091	0.1930
14	0.7901	0.7808	0.0092
15	1.6301	1.1613	0.4688
20	-0.3201	-0.1920	0.1280
22	0.0598	0.2397	0.1799
23	0.8573	0.4514	0.4059
24	0.7911	0.5006	0.2905
26	0.4824	0.5161	0.0338
27	1.7886	1.6368	0.1518
28	1.0291	0.8009	0.2281
30	0.4127	0.3197	0.0930
31	0.8698	0.6651	0.2046
32	0.3034	0.3292	0.0258
36	1.0969	1.4592	0.3622
40	2.4586	0.7431	1.7155
42	1.2674	1.0571	0.2103
43	1.3310	0.6420	0.6890
45	0.0000	0.6960	0.6960
46	0.8773	1.2855	0.4082
49	-0.3637	-0.0607	0.3030
50	2.0628	1.4118	0.6510
51	2.0842	1.6678	0.4164
52	1.5982	1.4976	0.1006
53	1.6326	1.4084	0.2242
54	1.5497	1.6485	0.0988
55	1.6628	1.7113	0.0486
56	0.6532	1.2533	0.6001
58	2.0544	1.9671	0.0872
59	2.0607	2.0591	0.0016
60	1.5051	1.7997	0.2946
62	1.5999	1.6437	0.0438
63	1.7129	1.4872	0.2257
64	2.2041	2.0968	0.1073
65	2.0607	2.2735	0.2128
66	2.0969	2.2402	0.1432
68	1.6546	2.3284	0.6738
69	2.8195	2.2470	0.5725
70	2.4771	2.2709	0.2062
71	2.4008	2.2763	0.1245

72	1.7597	1.5473	0.2123
73	1.8924	2.3118	0.4194
74	1.8625	2.3139	0.4514
77	2.1761	2.0501	0.1260
78	2.3118	2.2255	0.0863
80	0.5261	1.2716	0.7454
88	1.5081	1.5999	0.0918
89	2.6637	2.2014	0.4622
90	2.2739	2.2281	0.0459

* LOO = Leave-One-Out approach

(B) TEST SET (QAAR Model)

Compound Number	Observed $pIC_{50(\text{hBACE1-hCatD})}$	Predicted $pIC_{50(\text{hBACE1-hCatD})}$	Absolute Prediction Error
1	0.2374	0.2201	0.0173
10	1.1027	0.8503	0.2524
12	0.8120	0.6551	0.1569
19	0.8829	0.4208	0.4621
25	0.7112	0.4208	0.2904
33	1.3460	0.8224	0.5236
34	1.0940	0.9206	0.1734
35	1.0934	0.5770	0.5164
37	0.4873	1.1905	0.7031
39	0.3615	0.7190	0.3575
44	0.6587	0.6689	0.0102
47	1.3310	1.2838	0.0471
48	1.1648	1.3344	0.1696
57	1.6676	1.8766	0.2091
61	2.6532	2.2844	0.3688
67	1.7959	2.3537	0.5578
75	2.6385	2.2844	0.3541
76	1.5749	1.3356	0.2393
79	2.5441	2.1713	0.3728
82	2.3500	2.2310	0.1190
83	1.8751	2.2310	0.3559
84	1.4624	2.2310	0.7686
85	2.4510	2.2310	0.2200
86	2.7837	2.2310	0.5527
87	0.9890	1.5681	0.5791

Table S2. Error-based metrics and prediction quality for the training and test sets employed in QAAR model.

Set	MAE (100%)	σ_{AE} (100%)	MAE (95%)	σ_{AE} (95%)	MAE+3× σ_{AE} (95%)	Prediction Quality	Response Range	Response Mean
Training	0.3235	0.3018	0.2749	0.2040	0.8869	**Bad**	3.1832	1.2471

Test	0.3351	0.2079	0.3002	0.1765	0.8298	**Bad**	2.5463	1.4427

Table S3. The distribution of prediction error (in %) for all the three QSAR models (i.e., 2D-QSAR, G-QSAR and QAAR) for actual depiction of model prediction quality.

QSAR models		**Distribution of prediction error (in %)**				**Result**
Model	Set	%*NComp*[*]> (0.1×*TR*[#])	%*NComp*> (0.15×*TR*)	%*NComp*> (0.2×*TR*)	%*NComp*> (0.25×*TR*)	Prediction quality based on MAE-based criteria
2DQSAR	Training	32.692	13.462	5.769	0.000	*Moderate*
	Test	27.273	9.091	0.000	0.000	*Good*
	External	29.412	11.765	5.882	5.882	*Moderate*
GQSAR	Training	45.098	13.726	1.961	0.000	*Moderate*
	Test	36.364	13.636	9.091	0.000	*Moderate*
	External	29.412	11.765	0.000	0.000	*Good*
QAAR	Train	36.842	***21.053***	14.035	3.509	***Bad***
	Test	52.000	***28.000***	8.000	0.000	***Bad***

***** %*NComp* = percentage of compounds in the respective set.

[#] *TR*= Range of training set observed response values.