ORIGINAL ARTICLE

Physiochemical features for designing novel B-site APP-cleaving enzyme 1 inhibitor against Alzheimer's

Richa Arya[1]*, S. Paliwal[1], S. P. Gupta[2], Neetika tripathi[1]

[1]Department of Pharmacy, Banasthali Vidyapith, Vanasthali, Rajasthan, India, [2]Department of Pharmaceutical Technology, Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India.

Abstract

Background: B-site APP-cleaving enzyme 1 (BACE-1) is a target protein for therapeutic treatment of Alzheimer's disease. **Objective:** The present work focuses on designing of novel BACE-1 inhibitors by predicting physiochemical features relevant to binding with the target. **Methods:** In an attempt, the Piperazine as BACE-1 inhibitors with reported Ki values was considered for multiple linear regression method. **Results:** The descriptors, namely number of hydrogen bond acceptor (subs 1) and bond lipole (subs 2), are positively correlated, and the dipole moment Y component (subs 2) is negatively correlated with the dependent variable. A few compounds were predicted based on the regression equation. **Conclusion:** The predicted compounds were novel as determined by fingerprinting and have improved value of binding constant for receptor interaction.

Key words: Alzheimer, B-site APP-cleaving enzyme 1, fingerprint, multiple linear regression, QSAR, regression, secretase

INTRODUCTION

Alzheimer's disease is a fatal neurological condition characterized by the massive deposits of plaque affecting memory.[1,2] The pathology of plaque accumulation was explained by amyloidal hypothesis.[3] The hypothesis postulates that the major fractions of neurodegenerative plaque in brain are the fragments of mutated amyloidal precursor protein (APP). The cleavage at the β-site and γ-site of the APP is due to beta- and y-secretase, respectively.[4] B-site APP-cleaving enzyme 1 (BACE1) is a 501 amino acids containing protein functioning and is a rate-limiting step in the proteolysis of APP.[5] The survival of BACE-1 knockout mice also supports the preference of the current target, i.e., BACE-1 over other therapeutic targets.[6] A few chemical compounds such as AZD3293, GSK188909, and MK8931[7,8] reached clinical trials after struggling for years but unfortunately failed. Numbers of other small molecules have also been reported in the last decade as novel BACE-1 inhibitors, but none of them successfully reached market. Thus, an effort to identify the novel BACE-1 inhibitors in minimum time frame with improved selectivity was performed. The computational approach is still the most preferred alternative to laboratory experiments. To satisfy the long felt need of novel BACE-1 inhibitor, the regression analysis was performed using TSAR 3.3 (www.accelrys.com) to predict the binding of novel BACE-1 inhibitors.[9] The compounds were predicted based on regression equation and descriptors involved. The fingerprint and similarity indices were also calculated for the compounds using tanimoto index implemented in Discovery Studio version 2.0 (DS) (www.accelrys.com).

METHODS

QSAR and Molecular Modeling

A QSAR approach to establish a correlation between 2D descriptors and binding constant (Ki) was established by TSAR. The data set was prepared from 37 known inhibitors of

Address for correspondence:
Richa Arya, Department of Pharmacy, Banasthali Vidyapith, Vanasthali, Rajasthan, India.
Phone:+91-9970099156.
E-mail: richapharma2011@gmail.com

Received: 08-03-2018
Revised: 27-05-2018
Accepted: 18-06-2018

BACE-1 with reported Ki from the literature.[10] The chemical structures were then drawn using ChemDraw Ultra 8.0 (www.cambridgesoft.com) in.Mol file. The structures were ultimately imported in the TSAR sheet for energy minimization and deriving charge.[11,12] The dependent variable (Ki) was converted as log function for data standardization. The independent variables are the software enabled 2D chemical descriptors. The descriptor generation was a structurally guided protocol of the software. The chemical properties calculated on TSAR platform are molecular properties such as inertia, bulk, Verloop, charge, and log P. The connectivity, shape, topology, and the electrostatic properties were also calculated.[13] The mutually correlated descriptors were removed based on pairwise elimination method. The most relevant descriptors with highest correlation with binding constant were retained through data reduction. The division of test and training set was performed by prime consideration of variation of activity range in training set.[14] The applicability domain was established by calculating higher and lower leverage values of compounds in training set. The training set with best statistical value of regression standard deviation (s), Fisher value (f), and r and r^2 with set of descriptors was further validated by bootstrapping and Y randomization.[15,16] The model was cross-validated by leave-N-out (L-N-O) methodology.[17]

Predicted Compounds And Molecular Fingerprint

The low-molecular-weight compounds were designed from regression equation and relevance of descriptors involved in the study. Thus, the Piperazine scaffold N-((1S, 2S)-3-(3, 5-difluorophenyl)-1-hydroxy-1-((R)-4-(H-sulfonyl) piperazin-2-yl) propan-2-yl) carboxamide) was utilized for the structural optimization to identify virtual derivatives of interest. The prime considerations for structure optimization were the highly correlated descriptors for whole molecule as well as substituent. The binding is also calculated as per equation (1) for all predicted compounds. The novelty of compounds was checked by tanimoto index protocol enabled in DS. It is a molecular fingerprint and structural similarity indices technique. It determines novelty of the molecular structure in a set of ligands reported as BACE-1 inhibitors. The structural pool of 340 ligands was input reference ligand for similarity indices. The minimum value of tanimoto index verifies the novelty of the predicted ligand for a specific target.

RESULTS AND DISCUSSION

QSAR And Molecular Modeling

The regression analysis multiple linear regression (MLR) was performed on a dataset of 37 compounds with Ki values as dependent variable. The structures of all compounds in training set and test set are presented in Figures 1 and 2.

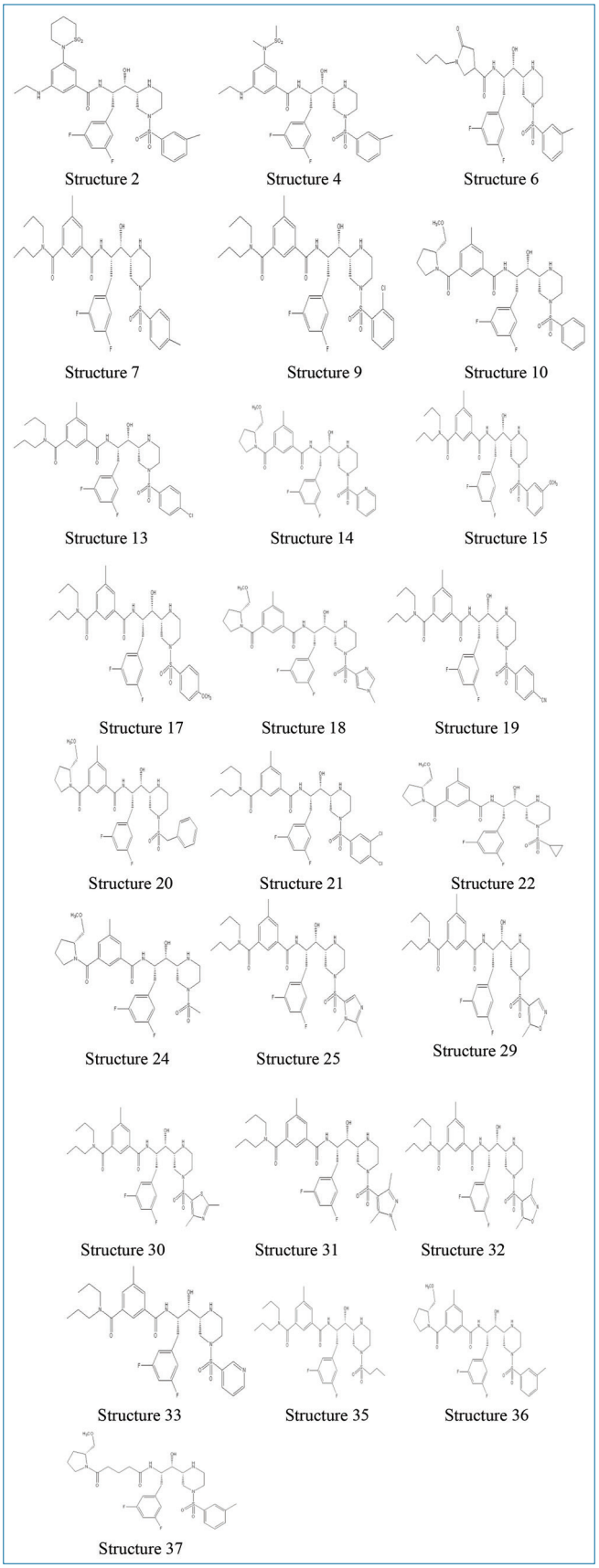


Figure 1: Structure of compounds in training set

The equation (1) was generated for 25 compounds in the training set presented below:

Log (1/Ki) = -0.528*X1 + 0.099*X2 + 1.247*X3 + 5.67 (1)

n= 25 r^2= 0.89 S value = 0.38 F value= 59.70

The regression coefficient (r) is 0.94. The most relevant descriptors entered for model development were dipole moment Y component (subs 2), bond lipole (subs 2), and number of hydrogen bond acceptor (subs 1) [Table 1: Confidence Table]. The validation was performed by bootstrapping, Y randomization, and cross-validation. The model was internally validated by the r^2 (0.78) obtained from test set. A graph representing value of r^2 for test and training set is presented in Graph 1. The model was cross-validated by L-N-O method, where N is 2. The value of r^2cv was 0.83.

A tabular representation of compounds for model development with values of descriptor for each molecule is presented in Table 2:

The descriptor with highest correlation with the dependent variable was found to be number of hydrogen bond acceptor (subs 1). A figure representing all the descriptors under study is presented in Figure 3.

The mutually correlated descriptors were not retained for model development. Thus, a correlation matrix of the descriptors is presented in Table 3.

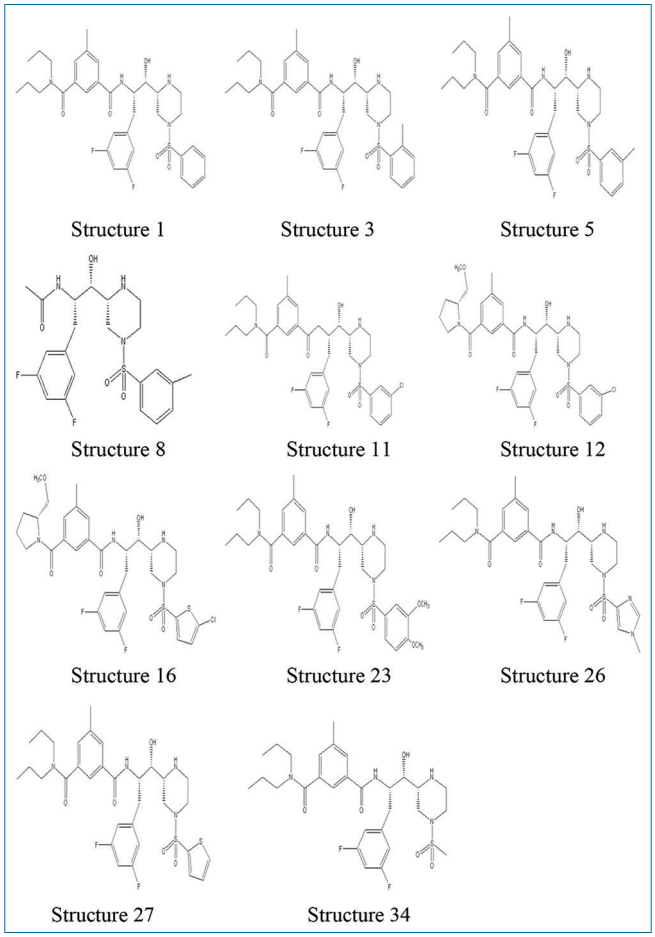


Figure 2: Structure of compounds in test set

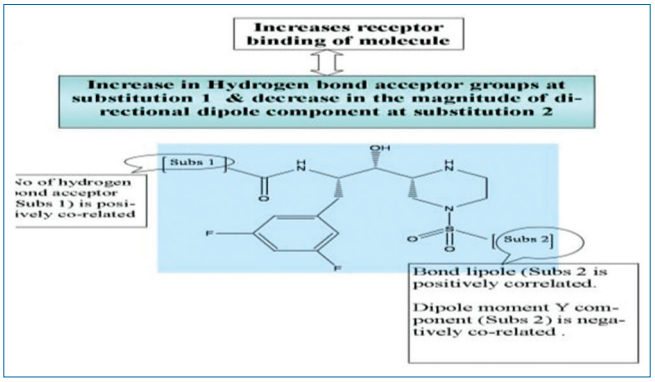
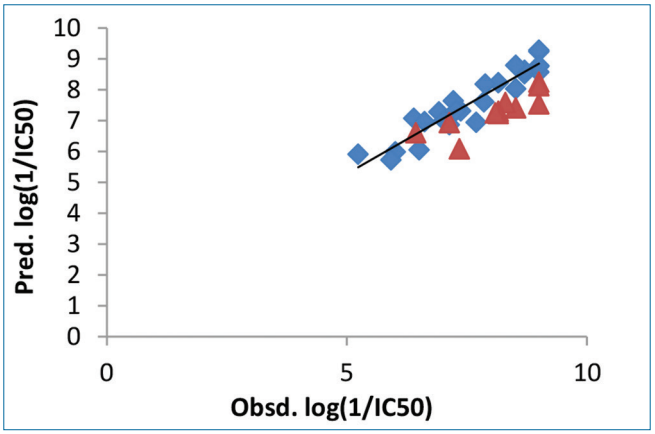


Figure 3: A figure representing all the descriptors under study is presented



Graph 1: Actual and predicted values of training and test set for multiple linear regression

Table 1: Confidence Table

Descriptor	Coefficient	Jackknife SE	Covariance SE	*t*-value	t-probability
Dipole moment Y component (subs 2)	–0.528	0.094	0.078	–6.711	1.21815e-006
Bond lipole (subs 2)	0.099	0.027	0.023	4.169	0.000433714
Number of hydrogen bond acceptor (subs 1)	1.247	0.202	0.177	7.013	6.34679e-007
Constant	–3.339	0.280			

Table 2: Data set of compounds with value of dependent and independent variable for regression analysis

Compound name	Ki (dependent variable)		Descriptors-independent variables		
	Log (1/Ki)	Pred. Ki by MLR	Dipole moment Y component (subs 2)	Bond lipole (subs 2)	Number of hydrogen bond acceptor (subs 1)
[b]1	8.522	7.404	0.205	6.078	1
2	9	9.236	−0.480	8.308	2
[b]3	8.301	7.587	−0.007	6.784	1
4	9	9.280	−0.551	8.373	2
[b]5	9	7.538	0.3957	8.441	1
6	7.853	7.608	0.263	8.438	1
7	7.221	7.643	0.363	9.326	1
[b]8	6.437	6.603	−0.199	8.416	0
9	6.397	7.075	0.979	6.882	1
10	9	8.573	0.352	6.073	2
[b]11	8.154	7.298	0.921	8.820	1
[b]12	9	8.249	1.485	8.832	2
13	7.376	7.314	1.052	9.671	1
14	8.698	8.516	0.015	3.701	2
15	7.698	6.939	0.939	5.310	1
[b]16	9	8.129	1.072	5.431	2
17	6.920	7.283	0.217	4.925	1
18	8.698	8.628	−0.783	0.583	2
19	6.508	6.050	2.618	5.292	1
20	9	8.775	−0.165	5.353	2
21	7.136	6.863	2.421	12.414	1
22	8.154	8.230	0.095	1.259	2
[b]23	7.346	6.087	2.266	3.791	1
24	7.886	8.177	5.96E-08	0.217	2
25	8.522	8.028	−1.793	1.729	1
[b]26	8.154	7.230	−0.496	0.597	1
[b]27	8.096	7.241	0.162	4.218	1
[#]28	8	6.614	0.908	1.871	1
29	7.275	7.327	−0.725	0.354	1
30	6.619	6.960	0.422	2.772	1
31	6.013	5.988	2.148	2.170	1
32	5.920	5.723	2.564	1.721	1
33	5.236	5.906	2.548	3.468	1
[b]34	7.136	6.930	0	0.218	1
35	7.022	7.012	0.398	3.168	1
36	9	8.756	0.449	8.419	2
37	8.522	8.786	0.394	8.431	2
[#]38	9	10.482	−0.478	8.303	3

[b]Test set compounds, [#]Not included for deriving correlation. MLR: Multiple linear regression

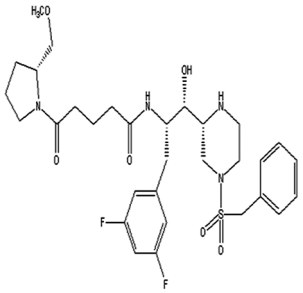
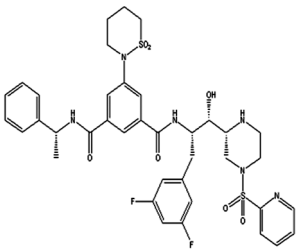
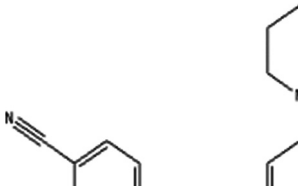
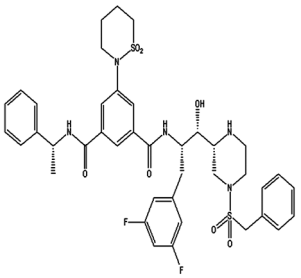
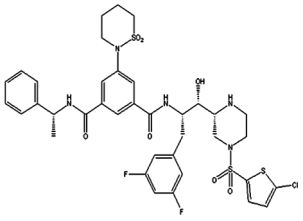
Predicted Compounds And Molecular Fingerprint

The structural modifications at different position were done for designing novel molecules as BACE-1 inhibitors. Five compounds were predicted in this paradigm, considering relevance of descriptors for target. The results of regression analysis prompted the designing of molecules by introducing hydrogen bond acceptor groups at substitution 1. The

Table 3: Correlation matrix of descriptors with the dependent variable

Descriptor	Dipole moment Y component (subs 2)	Bond lipole (subs 2)	Number of hydrogen bond acceptor (subs 1)	Dependent variable
Dipole moment Y component (subs 2)	1	0.195	−0.446	−0.732
Bond lipole (subs 2)	0.195	1	−0.015	0.186
Number of hydrogen bond acceptor (subs 1)	−0.446	−0.015	1	0.793

Table 4: Structure of predicted compounds with their calculated binding constant

Name	Structure of pred. compound	Dipole moment Y component (subs 2)	Bond lipole (subs 2)	Number of hydrogen bond acceptor (subs 1)	Calculated Ki from equation (1)
A1.mol		−2.11	1.988	2	9.46
A2.mol		0.029	3.137	3	9.705
A3.mol		1.109	3.138	4	10.35
A4.mol		−1.098	2.061	3	10.19
A5.mol		−0.739	2.378	3	10.02

binding of these predicted compounds was calculated as per equation (1) and compared with that of compound of highest and lowest Ki value of training set. All the predicted compounds exhibited improved binding constant with the BACE-1 receptor. To check novelty of the predicted compound by fingerprinting technique, the tanimoto index for all compounds is <0.11. The compound A3 with highest binding constant has tanimoto index of 0.087 [Table 4].

CONCLUSION

To identify the most relevant descriptors for receptor binding of BACE-1 inhibitors, the regression analysis was performed by multiple linear regression method. The model was internally validated by test set predictions with value of r^2 as 0.78. The model reflected H-bond acceptor (subs 1) as the most relevant positively correlated descriptor. The predicted compounds showed improved calculated binding constant with the receptor BACE-1. Among all the predicted compounds, the tanimoto index calculated for compound A3 is 0.087. In conclusion, the overall affinity of the compounds was expected to increase by introducing hydrogen acceptor group at substitution. The structure optimization of predicted compounds can be further explored to identify novel BACE-1 inhibitors with improved binding with the receptor.

ACKNOWLEDGMENT

Computational softwares were provided by Banasthali University, and all the authors thank the Vice Chancellor, for extending all the necessary facilities.

REFERENCES

1. Kovács T. Therapy of alzheimer disease. Neuropsychopharmacol Hung 2009;11:27-33.
2. Singh S, Kushwah AS, Singh R, Farswan M, Kaur R. Current therapeutic strategy in alzheimer's disease. Eur Rev Med Pharmacol Sci 2012;16:1651-64.
3. Cole SL, Vassar R. BACE1 structure and function in health and alzheimer's disease. Curr Alzheimer Res 2008;5:100-20.
4. Li Q, Südhof TC. Cleavage of amyloid-beta precursor protein and amyloid-beta precursor-like protein by BACE 1. J Biol Chem 2004;279:10542-50.
5. Sun J, Roy S. The physical approximation of APP and BACE-1: A key event in alzheimer's disease pathogenesis. Dev Neurobiol 2018;78:340-7.
6. Wiltfang J, Esselmann H, Cupers P, Neumann M, Kretzschmar H, Beyermann M, *et al.* Elevation of beta-amyloid peptide 2-42 in sporadic and familial alzheimer's disease and its generation in PS1 knockout cells. J Biol Chem 2001;276:42645-57.
7. Hung SY, Fu WM. Drug candidates in clinical trials for alzheimer's disease. J Biomed Sci 2017;24:47.
8. Evin G. Future therapeutics in alzheimer's disease: Development status of BACE inhibitors. BioDrugs 2016;30:173-94.
9. Pal T, Somani RR, Shirodkar PY, Bhanushali UV. Quantitative structure activity relationship (QSAR) studies of a series of 2,5-disubstituted-1,3,4-oxadiazole derivatives as biologically potential compounds. Med Chem Res 2011;9:1550-5.
10. Cumming J, Babu S, Huang Y, Carrol C, Chen X, Favreau L, *et al.* Piperazine sulfonamide BACE1 inhibitors: Design, synthesis, and *in vivo* characterization. Bioorg Med Chem Lett 2010;20:2837-42.
11. Frimayanti N, Yam ML, Lee HB, Othman R, Zain SM, Rahman NA, *et al.* Validation of quantitative structure-activity relationship (QSAR) model for photosensitizer activity prediction. Int J Mol Sci 2011;12:8626-44.
12. Simões RS, Maltarollo VG, Oliveira PR, Honorio KM. Transfer and multi-task learning in QSAR modeling: Advances and challenges. Front Pharmacol 2018;9:74.
13. Duchowicz PR. Linear regression QSAR models for polo-like kinase-1 inhibitors. Cells 2018;7: pii: E13.
14. Martínez-Santiago O, Marrero-Ponce Y, Vivas-Reyes R, Rivera-Borroto OM, Hurtado E, Treto-Suarez MA, *et al.* Exploring the QSAR's predictive truthfulness of the novel N-tuple discrete derivative indices on benchmark datasets. SAR QSAR Environ Res 2017;28:367-89.
15. Kiralj R, Ferreira MC. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. J Braz Chem 2009;20:4.
16. Kaushik S, Gupta SP, Sharma PK, Anwar Z. A QSAR study on some series of HIV-1 integrase inhibitors. Med Chem 2011;7:553-60.
17. Dong Y, Xiang B, Du D. PVLOO-based training set selection improves the external predictability of QSAR/QSPR models. J Chem Inf Model 2017;5:1055-67.

Source of Support: Nil. **Conflict of Interest:** None declared.