

QSAR and molecular docking studies on a series of spirocyclic BACE-1 inhibitors

Abstract

The most prominent irreversible neurodegenerative disease, Alzheimer's disease (AD), involves an enzyme, known as β -site APP (amyloid precursor protein) cleaving enzyme-1 (BACE-1). Prediction of potent BACE-1 inhibitors is a potential therapeutic option for the management of AD. The present study deals with the modeling of spirocyclic BACE-1 inhibitors using molecular descriptors to further contribute to better clinical management of AD. A multiple linear regression (MLR) analysis has shown that physicochemical and 2D autocorrelation descriptors of these compounds are the major influencing factors of their activity and a docking study has shown that compounds can form hydrogen bonds with the receptor and have effective steric interactions. The MLR model fits the training set with $R^2 = 0.857$. New potential compounds with higher therapeutic potency than the existing compounds were also docked the enzyme, BACE-1, to validate the formation of hydrogen bonds and hydrophobic interactions. The predicted models will be a constructive approach for the clinical management of Alzheimer's disease.

Keywords: alzheimer's disease, BACE-1, QSAR, docking analysis, pharmacokinetic studies

Volume 11 Issue 1 - 2022

Bashirulla Shaik,¹ Tabassum Zafar,² Vijay K Agrawal,³ Satya P Gupta⁴

¹Department of Applied Sciences, National Institute of Technical Teachers' Training & Research, Bhopal-462002, Madhya Pradesh, India

²Department of Bioscience, Barkatullah University, Bhopal, Madhya Pradesh, India

³Department of Chemistry, Awadhesh Pratap Singh University, Rewa, Madhya Pradesh, India

⁴Department of Pharmaceutical Technology, Meerut Institute of Engineering and Technology, Meerut-250005, India

Correspondence: Satya P Gupta, Department of Pharmaceutical Technology, Meerut Institute of Engineering and Technology, Meerut-250005, India, Tel + 91- 8889534745, Email spgbits@gmail.com

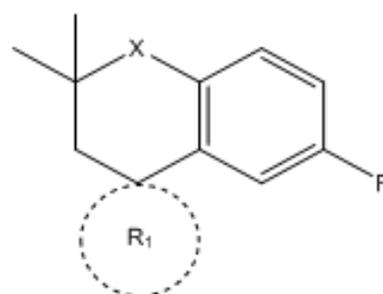
Received: February 22, 2022 | **Published:** February 28, 2022

Introduction

Alzheimer's disease (AD) is an age-dependent, irreversible, neurodegenerative disorder without any complete cure to date. Gradual impairment of cholinergic neuron activity due to nerve cell death leads to a decline in cognitive function, which alters the convenience of daily routines and slowly enables the patient to depend on assistance and monitoring.¹ Dr. Alois Alzheimer was the first German psychiatrist and neuropathologist who described a dementing condition that later became known as Alzheimer's Disease (AD). According to the World Health Organization (WHO) Bulletin, dementia (gradual memory loss) is the prime clinical symptom in AD patients. Dementia due to AD progression is very common in older people from 60 years upwards. More than 47 million dementia patients exist in this world; according to WHO, more than half of them belong to underdeveloped countries. Alzheimer's management involves the downregulation of various pathways that slow down the progression of the disease and offer symptomatic relief.² The symbol of AD is the brain deposition of amyloid-beta ($A\beta$), which is a peptide containing 36–43 amino acids that is likely a primary driver of neurodegeneration. amyloid-beta is produced by the sequential cleavage of amyloid precursor protein (APP) by β -Site amyloid precursor protein cleaving enzyme 1 (BACE-1) and γ -secretase in the amyloidogenic pathway. BACE-1 works as a rate-limiting enzyme that catalyzes the cleavage of amyloid precursor protein (APP) to produce neurotoxic amyloid β ($A\beta$) protein, which leads the plaque formation and results in neurodegeneration. Prediction of potent BACE-1 inhibitors are potential therapeutic options for the management of Alzheimer's Disease (AD).³

The enzyme BACE-1 present in humans and encoded by the BACE-1 gene is an aspartic-acid protease essential in the development of myelin sheaths in peripheral nerve cells. Therefore, inhibition of BACE-1 denotes an attractive beneficial target to slow or prevent Alzheimer's disease (AD).⁴ In the past, several attempts have been made to build QSAR models for BACE-1 Inhibitors. Hunt et al.⁵ explored the structure-activity relationships

of core changes, P3 moieties, and Asp binding functional groups in order to optimize BACE-1 affinity, cathepsin D selectivity, and blood-brain barrier penetration. These authors also demonstrated a pharmacokinetics/pharmacodynamic (PK/PD) relationship between free drug concentrations in the brain and cerebrospinal fluid (CSF) $A\beta$ lowering.⁵ Jadhav et al. developed a 2D-QSAR model using multiple linear regression analysis for the prediction of BACE-1 antagonistic activity for theoretically calculated descriptors to propose some new potential compounds. The model developed by them could be useful to provide an idea about the relationship between the physicochemical descriptors like logP, logD, and PSA with BACE-1 inhibitory activity and to design novel and potent BACE-1 inhibitors.⁶ The aim of the present study is to analyze a series of spirocyclic BACE-1 inhibitors (**I**), studied by Hunt et al.,⁵ in relation to physicochemical properties and drug-receptor interactions in order to find therapeutically most useful compounds of this class. The methods applied in this study are linear and nonlinear QSAR (Quantitative Structure-Activity Relationship) models, support vector machine (SVM), and artificial neural network (ANN) approaches.



Materials and methods

All the compounds for the study have been taken from Hunt. et al.⁵ All the compounds are listed in Table 1 along with their physicochemical

properties and BACE-1 inhibition activity. For QSAR studies, out of total 54 compounds, 75% of them (41 compounds) were selected for training set by random selection using Statistical Data miner software for generation of the models, and the remaining 25% (13 compounds) were used for the test set for evaluating the predictability of the developed models. The chemical structures were drawn using ACD/ChemSketch software⁷ (version 2018.1.1) and the physicochemical/topological descriptors were calculated using DRAGON software.⁸ Among all the calculated descriptors, only those descriptors are listed in Table 1 which were found to be correlated with the activity. In the Table, the test set compounds are marked with superscript ‘b’ and the compounds that behaved as outliers, and thus excluded from the correlation, are marked with superscript ‘c’. The most significant structural descriptors that were found to govern the activity of the compounds were as follows:

MW: Molecular weight

MATS6m = Moran autocorrelation of lag 6 weighted by mass

MATS8p = Moran autocorrelation of lag 8 weighted by polarizability

MATS6i = Moran autocorrelation of lag 6 weighted by ionization potential

GATS2v = Geary autocorrelation of lag 2 weighted by van der Waals volume

The above parameters characterize either the bulk of the molecules or their electronic their electronic properties.⁹⁻¹¹

Results and discussion

QSAR study

A multiple linear regression (MLR) analysis was performed using Statistica Dataminer,¹² which was found to be quite successful in many cases,¹³⁻¹⁶ on the training set compounds to establish a correlation between activity and various descriptors of the compounds. The most significant correlation obtained was as shown by eq. (1).

$$pIC_{50} = -7.3562(\pm 2.5477)MATS6m - 6.1804(\pm 2.9649)GATS2v + 7.3441(\pm 4.3423)MATS6i + 0.0111(\pm 0.0061)MW + 7.2258 \quad (4)$$

$$n = 39, r^2 = 0.7366, r_{cv}^2 = 0.6034, r_{pred}^2 = 0.3792, s = 0.4805, F_{4,34} = 23.76(3.93)$$

$$pIC_{50} = -6.6684(\pm 2.9334)MATS6m - 5.3125(\pm 3.4070)GATS2v + 11.6693(\pm 4.2233)MATS6i + 11.2604 \quad (5)$$

$$n = 39, r^2 = 0.6315, r_{cv}^2 = 0.5462, r_{pred}^2 = 0.2801, s = 0.5601, F_{3,35} = 20.00(4.40)$$

$$pIC_{50} = -6.4851(\pm 3.2745)MATS6m + 10.5968(\pm 2.2954)MATS6i + 7.2314 \quad (6)$$

$$n = 39, r^2 = 0.5260, r_{cv}^2 = 0.4455, r_{pred}^2 = 0.0788, s = 0.6264, F_{2,36} = 19.98(5.25)$$

$$pIC_{50} = 11.1808(\pm 5.5096)MATS6i + 8.0147 \quad (7)$$

$$n = 39, r^2 = 0.3136, r_{cv}^2 = 0.1960, r_{pred}^2 = 0.0115, s = 0.7435, F_{1,37} = 16.91(7.37)$$

Thus, from the above results, it is clear that eq. (1) has a noteworthy correlation between the BACE-1 inhibitory activity and the structural descriptors of the compounds. Although the correlation does not have any mechanistic aspects, but it has good predictive ability. A graph drawn between the predicted and observed activities for both the training and test sets (Figure 1) further shows that the model has good

$$pIC_{50} = -5.9149(\pm 1.9837)MATS6m - 4.5975(\pm 2.2987)GATS2v + 3.3755(\pm 1.2986)MATS8p + 6.2716(\pm 3.2725)MATS6i + 0.0154(\pm 0.0049)MW + 3.8047 \quad (1)$$

$$n = 39, r^2 = 0.8574, r_{cv}^2 = 0.7974, r_{pred}^2 = 0.5204, s = 0.3589, F_{5,33} = 39.68(3.64)$$

In eq. (1), n refers to the number of data points used in the correlation, r^2 is the square of the correlation coefficient, r_{cv}^2 is the square of cross-validated correlation coefficient obtained by leave-one-out (LOO) jackknife procedure, and r_{pred}^2 is the square of correlation coefficient obtained for test set compounds to judge the external validity of the correlation. Values of r_{cv}^2 and r_{pred}^2 are calculated according to eqs. (2) and (3), respectively, where $y_{i,obsd}$ in eq. (2) refers to the observed activity of compound i in the training set and that in eq.(3) to the compound i in test set. Similarly, $y_{i,pred}$ in eq.(2) refers to the predicted activity of compound i in the training set obtained in leave-one-out jackknife procedure and that in eq.(3) to that predicted for the test set compounds by model obtained for the training set. However, $y_{av,obsd}$ in both the equation refers to the average activity of the training set compound.

$$r_{cv}^2 = 1 - \left[\frac{\sum_i (y_{i,obsd} - y_{i,pred})^2}{\sum_i (y_{i,obsd} - y_{av,obsd})^2} \right] \quad (2)$$

$$r_{pred}^2 = 1 - \left[\frac{\sum_i (y_{i,obsd} - y_{i,pred})^2}{\sum_i (y_{i,obsd} - y_{av,obsd})^2} \right] \quad (3)$$

The correlation is supposed to be valid and has the good internal predictive ability if $r_{cv}^2 > 0.60$. Similarly, the external predictive ability of the model is supposed to be good if its $r_{pred}^2 > 0.5$. From both the parameters, the correlation expressed by eq. (1) is found to be quite valid. Among the remaining two statistical parameters, s and F , s is the standard deviation and F is the Fischer-ratio between the variances of calculated and observed activities. The figures within the parentheses with \pm sign refer to the 95% confidence intervals. The F -value given in parenthesis refers to the standard F -value at the 99% level. A higher value of F than this indicates a good correlation. Also, all the descriptors used in this correlation are found to be quite significant as if we remove them one by one, the significance of the correlation is appreciably dropped (eqs. 4-7).

predictive ability. Figure 1 shows that except 1 or two points, all other points lie near the straight line. Using this MLR model [eq. (1)], we have predicted some new compounds, as shown in Table 2, where each compound has a higher activity value than any compound in the existing series (Table 1).

Docking analysis

A molecular docking study was carried out for the predicted compounds (Table 2) using FlexX software to see the binding of these compounds with the enzyme, BACE-1. The ability of a molecule to interact with an enzyme decides its potency. For the study of molecular docking, the crystal structure of the related enzyme is vital which can be now retrieved from the RCSB protein data bank. We selected the enzyme with PDB entry code 4JOO (<http://www.pdb.org>).¹⁷ All the predicted compounds as listed in Table 2 were docked in this enzyme and the docking results are shown in Table 3.

Table 3 Docking results of predicated molecules

Compd. No.	No. of Hydrogen Bonds	H-bonds	H-bonds Length (Å)	Score
1	3	H(48)-Asp276	3.29	-28.2518
		H(49)-Asp80	4.7	
2	4	H(64)-Asp80	8.3	-28.3562
		H(49)-Asp276	4.7	
		H(50)-Asp80	4.7	
		H(64)-Asp80	8.3	
3	3	O(34)-Arg283	1.86	-23.0708
		H(49)-Asp276	4.52	
		H(50)-Asp80	4.7	
4	3	H(66)-Asp80	8.3	-24.5289
		H(48)-Asp276	3.93	
		H(49)-Asp80	4.7	
5	5	H(66)-Asp80	8.3	-26.4259
		H(47)-Asp276	3.97	
		H(47)-Asp276	0.05	
		H(48)-Asp80	4.7	
		H(66)-Asp80	8.3	
6	5	O(33)-Arg283	0.11	-28.2523
		H(47)-Asp287	3.1	
		H(48)-Asp276	0.09	
		H(48)-Asp80	3.62	
		H(61)-Asp80	8.3	
7	5	O(33)-Arg283	7.66	-28.303
		H(47)-Asp276	2.97	
		H(48)-Asp276	0.09	
		H(48)-Asp80	3.01	
		H(61)-Asp80	8.3	
8	4	O(33)-Arg283	7.66	-24.4916
		H(48)-Asp80	4.7	
		H(49)-Asp276	2.22	
		H(69)-Asp276	4.85	
		O(33)-Trp124	4.7	
9	5	H(47)-Asp276	3.1	-28.2511
		H(48)-Asp80	3.62	
		H(48)-Asp276	0.09	
		H(61)-Asp80	8.3	
		O(33)-Arg283	7.66	
10	5	O(33)-Arg283	7.66	-25.973
		H(49)-Asp80	4.66	
		H(48)-Asp276	0.25	
		H(48)-Asp276	4.7	
		H(61)-Asp80	8.3	
11	4	O(32)-Trp124	5.39	-20.5751
		H(49)-Asp80	4.59	
		H(50)-Asp276	4.7	
		H(67)-Asp276	7.21	
		O(32)-Arg283	4.7	

Table Continued...

Compd. No.	No. of Hydrogen Bonds	H-bonds	H-bonds Length (Å)	Score
12	4	H(49)-Asp80	4.63	-20.9623
		H(50)-Asp276	4.7	
		H(73)-Asp276	7.63	
		O(32)-Trp124	4.7	
13	4	H(49)-Asp80	4.63	-20.9623
		H(50)-Asp276	4.7	
		H(73)-Asp276	7.63	
		O(32)-Trp124	4.7	
14	5	H(48)-Asp80	4.7	-23.2141
		H(49)-Asp80	4.7	
		H(70)-Asp80	8.09	
		O(33)-Trp124	4.7	
15	6	O(34)-Trp124	3.41	-24.171
		O(32)-Trp124	0.07	
		O(33)-Trp124	4.29	
		H(48)-Asp80	3.18	
		H(49)-Asp276	4.7	
		H(68)-Asp276	5.22	
		H(67)-Gly59	1.95	

Table 4 Pharmacokinetic properties of the proposed compounds (Table 3)

Compd. No.	ClogP	MW	HA	HD	NRB
1.	4.4491	487.2	7	1	2
2.	5.0551	522.65	7	1	2
3.	4.697	522.65	7	2	3
4.	4.091	487.2	7	2	3
5.	3.485	451.75	7	2	3
6.	3.5594	455.74	8	2	2
7.	3.0542	439.29	8	2	2
8.	3.9129	463.76	7	1	4
9.	3.5594	455.74	8	2	2
10.	4.1654	491.79	8	2	2
11.	3.1523	486.75	8	2	4
12.	2.6471	470.3	8	2	4
13.	3.1333	456.32	8	2	4
14.	2.7894	444.31	8	2	4
15.	2.0998	448.3	9	3	4

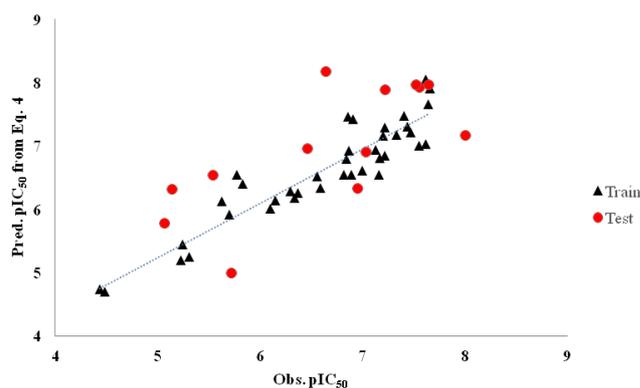


Figure 1 A plot between predicted and observed activities of compounds of Table 1.

The molecular docking analysis was carried on all the predicted compounds in the enzyme. Here we cited the only compounds 10 and 2 (Figure 2 and 3), the first compound 10 being the compound having

the highest predicted activity and the second having the highest docking score (Table 3), just to illustrate the best possible interactions between the inhibitors and the BACE-1 enzyme. From these Figure 2 and 3, it is clear that the predicted compounds have good interactions with the enzyme. They all undergo hydrogen bondings as well as steric interactions, in which several moieties of compounds are surrounded by the different active clefts of the enzyme. The penetration of any moiety of any inhibitor in any cavity of the enzyme will depend on its flexibility. All these steric interactions might involve dispersion interactions, which is a sort of electronic interaction.

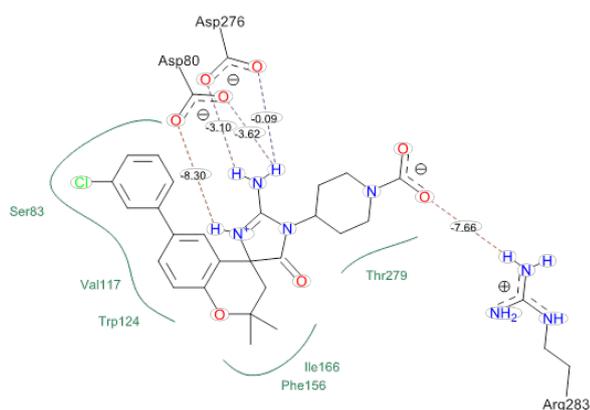


Figure 2 A representation of the binding of predicted compd **6** (Table 3) in BACE-1 (PDB entry Code 4JOO).

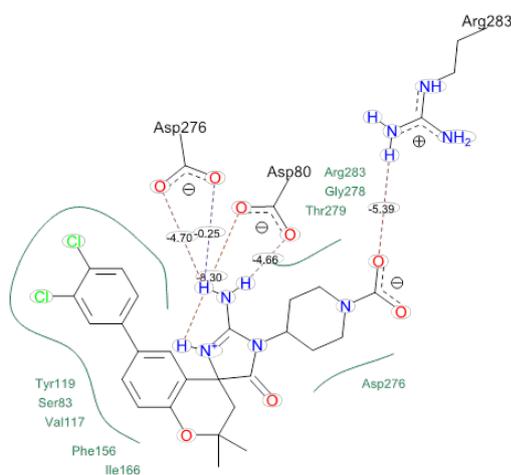


Figure 3 A representation of the binding of predicted compd **10** (Table 3) in BACE-1 (PDB entry Code 4JOO).

Pharmacokinetic studies

The pharmacokinetic profiles of the predicted compounds have been obtained using Data Warrior software¹⁷ and the results are shown in Table 4. These pharmacokinetic profiles include molecular weight (MW), ClogP, number of hydrogen bond acceptors (HAs), number of hydrogen bond donors (HDs), and number of rotatable bonds (NRBs). According to Lipinski's rule of five, compounds having MW < 500 and ClogP < 5 are supposed to have good absorption and permeation abilities.^{18,19} Similarly, according to Veber's rule, compounds having NRB < 10 are supposed to have good oral bioavailability [1-24]. Thus, almost all the predicted compounds have excellent pharmacokinetic profiles.

Conclusion

The BACE-1 inhibition activity of a series of spirocyclic compounds has been found to be well correlated with various physicochemical properties. Using the correlation obtained, some new spirocyclic compounds possessing better activity have been predicted. Through docking study of the predicted compounds, it has been shown how the compounds interact with the enzyme BACE-1. All the compounds are found to have a number of hydrogen-bondings with receptor as well as to involve their bulky groups in significant steric interactions with some sites of the enzyme. The study of the pharmacokinetic profiles of the predicted compounds has shown that they have good pharmacokinetic properties.

Acknowledgments

None.

Conflicts of interest

The author declares there is no conflict of interest.

References

1. Yiannopoulou KG, Papageorgiou SG. Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord.* 2013;6(1):19–33.
2. Neuner SM, Wilmott LA, Hoffmann BR, et al. Hippocampal proteomics defines pathways associated with memory decline and resilience in normal aging and Alzheimer's disease mouse models. *Behav Brain Res.* 2017;30:288–298.
3. Moussa-Pacha NM, Abdin SM, Omar HA, et al. BACE1 inhibitors: Current status and future directions in treating Alzheimer's disease. *Med Res Rev.* 2020;40(1):339–384.
4. Das B, Yan R. A Close Look at BACE1 Inhibitors for Alzheimer's disease treatment. *CNS Drugs.* 2019;33(3):251–263.
5. Hunt KW, Cook AW, Watts RJ, et al. Spirocyclic β -site amyloid precursor protein cleaving enzyme 1 (BACE1) inhibitors: from hit to lowering of cerebrospinal fluid (CSF) amyloid β in a higher species. *J Med Chem.* 2013;25;56(8):3379–403.
6. Jain P, Jadhav HR. Quantitative structure activity relationship analysis of aminoimidazoles as BACE-I inhibitors. *Med Chem Res.* 2013;22: 1740–1746.
7. ACD/Chemsketch, version 2018.1, Advanced Chemistry Development, Inc., Toronto, ON, Canada, 2019.
8. Taletsl. *Dragon software for molecular descriptors calculation*, version 6.0.
9. R Todeschini, V Consonni. *Wiley -vch, weinhiem, Germany, in the serie "Methods and principles in medicinal chemistry"*. 2000, p. 667.
10. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika.* 1950;37:17–23.
11. Geary RC. The contiguity ratio and statistical mapping. *The icorporated statistician.* 1954;5:115–145.
12. *Statistica (Data analysis software sstem) Version 10.*
13. Shaik B, Zafar T, Agarawal VK. Estimation of anti-HIV activity of HEPT analogues using MLR, ANN and SVM techniques. *Int J Med Chem.* 2013; Article ID 795621:1-8.
14. Louis B, Singh J, Shaik B. QSPR and ANN studies on prediction of aqueous solubility of heterogenous set of organic compounds. *Int J Pure and applied chem.* 2008;3(4):259–274.
15. Louis B, Agrawal VK, Khadikar PV. Prediction of intrinsic solubility of generic drugs using MLR, ANN and SVM analyses bioavailability. *Eur J Med Chem.* 2010;45(9):4018–4025.

16. Shaik B, Zafar T, Agrawal VK. Estimation of Anti-HIV Activity of HEPT Analogues Using MLR, ANN, and SVM Techniques. *Int J Med Chem*. 2013;(5112):795621.
17. Sander T, Freyss J, Korff MV, Rufener, C. Data Warrior: An open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model*. 2015;55:460–473.
18. Lipinski CA, Lombardo F, Dominy BW, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;1;46(1-3):3–26.
19. Lipinski CA. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov Today Technol*. 2004;1:337–341.
20. Veber DF, Johnson SR, Cheng S-Y, et al. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. 2002; 45: 2615–2623.