International Journal of Pharmacy and Biological Sciences
ISSN: 2321-3272 (Print), ISSN: 2230-7605 (Online)
IJPBS | Volume 8 | Issue 3 | JUL-SEPT | 2018 | 333-341
Research Article | Biological Sciences | Open Access | MCI Approved| ज्ञान-विज्ञान विमुक्तये
|UGC Approved Journal |

MONTE CARLO ALGORITHM BASED QSAR STUDY: EXPLORATION CHEMICAL FUNCTIONALITIES OF BACE1 INHIBITORS FOR THERAPEUTIC APPLICATION IN ALZHEIMER'S DISEASE

Md Lutful Islam[a*], Gulabchand K. Gupta[b]

[a]Research Scholar, Department of Computer Science and Engineering, Shri Jagdish Prasad Jhabarmal Tibrewala University, Vidyanagari, Jhunjhunu - 333 001, Rajasthan, India
[b]Seva Sadan College of Arts, Science & Commerce (Affiliated to University of Mumbai), Seva Sadan Marg, Ulhasnagar-421 003, Dist. Thane, Maharashtra, India

***Corresponding Author Email: lutful.islam@gmail.com**

ABSTRACT

Monte Carlo (MC) algorithm is widely used in the pharmaceutical research. Main objective of the current study was to explore the chemical functionalities of beta-site APP cleaving enzyme1 (BACE1) inhibitors for potential BACE1 inhibition using MC algorithm based QSAR study. In this purpose simplified molecular-input line-entry system (SMILES) based descriptors were used to develop quantitative structure-activity relationship (QSAR) models. BACE1 inhibitors dataset were collected from the binding database and divided into training, test, calibration and external sets. The training set was used to develop models and remaining sets for validation of the generated models. Two approaches were used to develop the QSAR models such as without- and with-considering the influence of cyclic rings of molecules on the inhibitory activity. Different statistical parameters were explored to assess the quality of the models. Data of the statistical parameters clearly explained the selected models were robust and efficient enough to predict the inhibitory activity of the molecules. Comparison of statistical parameters also suggested that cyclic rings of the dataset have positive impact on the inhibitory activity. From details analysis of QSAR models it was observed that molecular fragments found important to increase or decrease inhibitory activity which undoubtedly explained that models have mechanistic interpretation. Therefore, developed QSAR models in the current study can be used to design new promising BACE1 inhibitors for therapeutic application in Alzheimer's disease.

KEY WORDS

Alzheimer's disease, BACE1 inhibitors, QSAR, Monte Carlo algorithm, SMILES

INTRODUCTION

Alzheimer's disease (AD), caused by dementia is incurable neurodegenerative condition and highly dominant in old age worldwide [1-3]. It is known that AD is the most common cause of senile dementia which characterized by impairment of memory, disorientation, difficulty in speaking or writing, loss of reasoning skills, and delusions among other symptoms [4]. Although the direct cause of the AD is not clear but it is reported that both genetic and environmental factors are play major role in the progression of the AD [5]. According to World Health Organization (WHO) there is about 5.7 million people living in United States of America in 2018. As per data of 2017, more than 44 million people have AD or related dementia worldwide. It is also reported that AD most prominently found in Western Europe and least prevalent in Sub-Saharan Africa. In India, more than 4 million people have some form of dementia and AD.

According to report by the leading newspaper The Indian Express (September 21, 2016), the AD affected will be double in India by 2030. There is high risk of other age-related diseases including hypertension, dyslipidemia, metabolic syndrome and diabetes with the progression of AD. To date the only pharmacological therapeutics for such devastating disease is attenuate the symptoms and do not affect the mechanisms underlying disease development.

Gathering and deposition of amyloid β (Aβ) is widely accepted hypothesis for the growth of AD [6]. Aβ is a neurotoxic species produced by the successive breakdown of β-amyloid precursor protein (APP) by two aspartyl protease, beta-site APP cleaving enzyme1 (BACE1) and finally by γ secretase [7]. BACE1 has already been recognized as a significant target for AD intervention as its inhibition would halt the development of Aβ at the very beginning of β-APP processing [5]. Therefore, slow down the development of AD by inhibiting Aβ formation at an early stage is ideal approach to treat the AD. Experimentally it has been proved that BACE1 enzyme could be clinically feasible with few mechanistic side effects [8-10]. Therefore, limiting the production of Aβ through successful inhibition of BACE1 may represent modifying treatment for AD.

Therefore, keeping in mind, the urgent need of potential drug candidates for the proper treatment of AD, the current study was considered Monte Carlo (MC) algorithm based quantitative structure activity relationship (QSAR) study for exploration of important chemical functionalities and design new lead chemical agents for therapeutic application of AD. The QSAR can be defined as statistically validated and mathematical association between molecular descriptors obtained from chemical structures with biological activities. The molecular descriptors are the numerical values of the experimental or calculated properties obtained from molecules which describe specific information of the studied chemical structure. Well validated and robust QSAR models can give intuitions into the crucial structural information of the small molecules which contribute to biological activity [11]. The molecular descriptors are mainly divided into physico-chemical, topological and electronic, geometric and structural, and simple indicator parameter. Additionally, descriptors are also be categorized based on dimensionality (0D, 1D, 2D or 3D) [12]. Calculation of geometry-based descriptors are generally tough and required high computational costs and long computational calculation time. Subsequently 0D, 1D and 2D are known as the conformation-independent descriptors based on the constitutional and topological molecular features of compounds have been established as a substitute method [13, 14]. Most of the time the QSAR models are generated using the descriptors based on molecular graph [15-17] but the simplified molecular input-line entry system (SMILES) representation of the molecular structure can also be used [18-20] for molecular descriptor generation followed by development of QSAR models. SMILES notation based descriptors are based on both on the molecular structure and the property under analysis irrespective of details from the 3D-molecular geometry [11]. Hence, SMILES based molecular descriptors can be used to generate QSAR models [21-23]. Several research groups from industry and academia have already proven the reputation of the methodology, which was proficient of developing QSAR models with a similar or improved quality to the ones built with descriptors containing thousands of 0D–3D descriptors [24-30].

MATERIALS AND METHODS

Dataset

More than thousand BACE1 inhibitors were collected from Binding DB (http://www.bindingdb.org/) with inhibition constant (K_i) activity in nM range. Duplicate and without activity molecules were deleted. Further the Lipinski's rule of five [31] and Viber's [32] rules were checked and only considered molecules those followed the above two rules. On filtering using above criteria finally 411 molecules were used for the study. The experimental inhibitory activity (K_i) of entire dataset were converted into logarithm value [pK_i = log((1/K_i) x10^7)]. The SMILES structure along with pK_i values of molecules are given in Supplementary file (Tables S1 and S2). The whole dataset was randomly divided into training, calibration, test and external sets. Each of the set has specific role in QSAR formulation. The training set was used for model development and calibration and test sets were used to check the predictive ability of developed model. The external set was used for final estimation of the model using the compounds those were unseen during model generation that is no information of validation set was involved for model advance.

Optimal descriptors

To calculate the molecular descriptors the entire dataset was converted into SMILES format. The SMILES format of chemical structures is one of the useful representations and can be used to select optimal molecular descriptors which are mathematical functions of so-called correlation weights (CW) that is "Descriptors of Correlation Weights" (DCW). In the current study, the Monte Carlo algorithm was implemented to derive the DCW. The DCW was calculated using two approaches viz. without and with considering the influence of cyclic rings to the inhibition constant. The following equation (1) was used to calculate the DCW and influence of cyclic rings on inhibitory activity not considered.

$$DCW_1(SMILES,T,N_{epoch}) = \alpha\sum CW(S_k) + \beta\sum CW(SS_k) + \gamma\sum CW(SSS_k) + x\cdot CW(NOSP) + y\cdot CW(HALO) + z\cdot CW(BOND) + t\cdot CW(PAIR) \quad (1)$$

The T represents threshold which is defines as coefficient for classifying various molecular features extracted from SMILES into two classes such as active, in which CW is involved in the modelling process and rare, where CW is not involved in the modelling process. The number of epochs in Monte Carlo optimization represented by N_{epoch} which offers the best statistical results of the calibration set. The one symbol can be identified separately and represented by S_k. The SS_k and SSS_k are represented for combination of two or three respectively. Descriptors based on presence or absence of different elements are represented by NOSP, HALO, BOND and PAIR. NOSP explain the nitrogen, oxygen, sulphur and phosphorus; HALO represents halogen atoms such as fluorine, chlorine and bromine; BOND offers double (=), triple (#) or stereochemical bonds (@ or @@); and PAIR refers the probable grouping of pair atoms and/or SMILES attributes (for example double, triple, and stereochemical bonds) that takes place in the structure together. α, β, γ, x, y and t are discrete coefficient with values 0 and 1. Detail calculation of the above descriptors with example can be found in Worachartcheewan et al. [28].

The optimal descriptors with influence of cyclic rings can be calculated using following equation (2).

$$DCW_2(SMILES,T,N_{epoch}) = \alpha\sum CW(S_k) + \beta\sum CW(SS_k) + \gamma\sum CW(SSS_k) + x\cdot CW(NOSP) + y\cdot CW(HALO) + z\cdot CW(BOND) + t\cdot CW(PAIR) + CW(C3) + CW(C4) + CW(C5) + CW(C6) + CW(C7) \quad (2)$$

The C3, C4, C5, C6 and C7 are denoted by three-, four-, five-, six- and seven-membered cyclic rings. Details explanation of such descriptors are can be found somewhere else [33].

The well-known MC algorithm was adopted to calculate the CW which must give the best statistical results for the test set. To find out the preferable threshold (T*) and number of epochs (N*), range of T and N_{epoch} were selected from 1 to 10 and 1 to 30 respectively. The statistical results were analysed and the best (N*, T*) selected for final model development. The selected best statistics of calibration set makes possible to obtain the endpoint value using numerical values of correlation weights from the training set as follows:

$$\text{Endpoint} = C_0 + C_1 \times \text{DCW}(\text{SMILES, T, N}_{\text{epoch}}) \quad (3)$$

The endpoint represents the biological activity and, C_0 and C_1 are constant.

Validation

It is worth approach to validate the statistical model to explore the predictive ability and reliability of the selected model. The QSAR model can be validated by the help of a) internal validation using training set compounds; b) external validation using test compounds; and c) Y-scrambling or randomization of data. Research from worldwide [23, 24, 26, 27, 30] used these validation methodologies on SMILES notation optimal descriptor based QSAR models. In the current research the cross-validated correlation coefficient (Q^2) and error of estimation (s) were calculated based on predicted activity of training compounds. It is reported that high Q^2(>0.5) and low s explained better predictive ability of the model [34]. In order to check the good predictive capability of the training set molecules, the modified R^2 ($R^2_{m(LOO)}$) developed by Roy et al.[35, 36] the R^2_m was calculated which is the measure of the degree of deviation of the predicted activity from the observed ones. In order to verify the chance correlation **Y**-scrambling described by Ojha and Roy [37] was also performed in which ten probes of calculation were carried out. In one probe of calculation, X and Y represent the vectors of experiment and the vector of

prediction. First of all, exchange of random N1 and random N2 from row X (Y is not modified) were performed thousand times. Further, from above probes the $R^2_{(X, Y)}$ was calculated and represented as R^2_r. The $^CR^2_p$ was finally calculated according to the equation (4).

$$^C R^2_p = R \times (R^2 - R^2_r)^{1/2} \quad (4)$$

Where R^2 and R^2_r were utilized from the non-randomized and randomized model respectively. For acceptance of QSAR model the threshold value of $^CR^2_p$ should be greater than 0.5.

RESULTS AND DISCUSSION

In order to explore the important chemical functionalities of BACE1 inhibitors SMILES-based attributes were extracted and correlated with the inhibition constant. Total 411 BACE1 inhibitors with K_i value were collected from Binding DB. Monte Carlo algorithm based CORAL software (http://www.insilico.eu/coral/) was used to develop the robust QSAR model. In this research two approaches were used such descriptor generation with- and without-considering the influence of cyclic rings of the molecular systems.

Selection of optimal T and N_{epoch}

To develop the robust QSAR model it is necessary to identify optimal T and N_{epoch}. In this regards the "Search for preferable model" option of the CORAL was used for the threshold values in the range of 1 to 10 and the number of epochs ranging from 1 to 30. The statistical parameters, epoch numbers and corresponding threshold values are given in Tables 1 and 2 in case without- and with-considering the influence of cyclic rings on inhibitory activity respectively. The correlation coefficient of training, calibration and test sets were explored and optimal T and N_{epoch} (T* and N^*_{epoch}) selected. The (T* and N^*_{epoch}) were found to be (5, 4) and (5, 6) in case of without- and with-considering the influence of cyclic rings on inhibitory activity respectively. From Tables 1 and 2 it can be observed that some epoch numbers may have higher correlation coefficient for training, test and calibration sets but the $R_m{}^2_{avg}$ values not up to the mark (≥0.5). Hence above selected optimal T and N_{epoch} were used to develop QSAR model.

Table 1: Statistical parameters of training, calibration and test set to search T* and N^*_{epoch} for without influence of cyclic rings on inhibitory activity

Epoch no.	R_{tr}^2	s_{tr}	R_c^2	s_c	R_{ts}^2	s_{ts}	$R_m{}^2_{av}$	T
2	0.639	0.695	0.659	0.764	0.585	0.771	0.557	1
3	0.697	0.637	0.703	0.726	0.586	0.770	0.531	1
4	0.675	0.660	0.702	0.732	0.618	0.761	0.535	5
5	0.681	0.654	0.703	0.731	0.622	0.757	0.534	5
6	0.716	0.617	0.729	0.717	0.607	0.769	0.518	3
7	0.702	0.632	0.722	0.0720	0.622	0.767	0.509	5
8	0.717	0.616	0.736	0.698	0.611	0.782	0.490	5
9	0.762	0.565	0.762	0.660	0.514	0.873	0.378	1
10	0.720	0.612	0.740	0.705	0.592	0.814	0.449	5
11	0.726	0.606	0.740	0.693	0.599	0.798	0.463	5
12	0.747	0.582	0.747	0.692	0.546	0.837	0.419	3
13	0.736	0.595	0.748	0.690	0.556	0.836	0.424	4
14	0.742	0.588	0.748	0.691	0.561	0.829	0.431	4
15	0.739	0.592	0.743	0.692	0.554	0.830	0.427	4
16	0.741	0.590	0.753	0.689	0.560	0.831	0.426	4
17	0.739	0.591	0.752	0.693	0.581	0.844	0.405	5
18	0.746	0.583	0.748	0.688	0.553	0.843	0.412	4
19	0.749	0.580	0.753	0.685	0.547	0.855	0.400	4
20	0.732	0.600	0.753	0.692	0.600	0.819	0.434	5
21	0.725	0.607	0.739	0.702	0.586	0.821	0.447	6
22	0.734	0.598	0.746	0.694	0.582	0.825	0.432	5
23	0.726	0.606	0.746	0.707	0.601	0.810	0.459	7
24	0.751	0.578	0.757	0.696	0.512	0.892	0.357	4

25	0.730	0.602	0.745	0.698	0.573	0.826	0.437	6
26	0.602	0.731	0.634	0.796	0.566	0.786	0.537	4
27	0.755	0.574	0.755	0.697	0.536	0.870	0.378	3
28	0.721	0.612	0.751	0.697	0.557	0.850	0.412	8
29	0.732	0.599	0.751	0.695	0.576	0.830	0.428	6
30	0.721	0.611	0.753	0.704	0.567	0.829	0.431	6

R_{tr}^2: Correlation coefficient of training set; s_{tr}: standard error of training set; R_c^2: Correlation coefficient of calibration set; s_c: standard error of calibration set; R_{ts}^2: Correlation coefficient of test set; s_{ts}: standard error of test set; $R_m^2{}_{av}$: Modified correlation coefficient; *T*: Threshold

Table 2: Statistical parameters of training, calibration and test set to search T* and N*$_{epoch}$ for with influence of cyclic rings on inhibitory activity

Epoch no.	R_{tr}^2	s_{tr}	R_c^2	s_c	R_{ts}^2	s_{ts}	$R_m^2{}_{av}$	*T*
2	0.598	0.734	0.662	0.765	0.601	0.769	0.550	5
3	0.676	0.659	0.690	0.746	0.588	0.774	0.533	1
4	0.685	0.650	0.705	0.739	0.608	0.760	0.543	4
5	0.681	0.654	0.706	0.732	0.615	0.770	0.530	5
6	0.704	0.630	0.716	0.717	0.623	0.768	0.513	5
7	0.722	0.611	0.733	0.705	0.575	0.810	0.455	4
8	0.740	0.591	0.741	0.698	0.559	0.817	0.446	3
9	0.741	0.589	0.743	0.703	0.534	0.843	0.413	3
10	0.742	0.589	0.742	0.696	0.554	0.829	0.429	3
11	0.751	0.577	0.751	0.685	0.553	0.837	0.417	3
12	0.756	0.572	0.756	0.667	0.410	0.846	0.405	2
13	0.736	0.595	0.743	0.698	0.564	0.825	0.459	4
14	0.727	0.605	0.748	0.704	0.586	0.831	0.425	5
15	0.740	0.591	0.747	0.697	0.578	0.815	0.444	4
16	0.731	0.601	0.745	0.690	0.578	0.827	0.431	5
17	0.743	0.587	0.746	0.693	0.569	0.836	0.420	4
18	0.738	0.593	0.750	0.695	0.572	0.835	0.420	5
19	0.749	0.580	0.750	0.681	0.566	0.840	0.414	4
20	0.734	0.597	0.748	0.696	0.589	0.820	0.435	5
21	0.754	0.574	0.754	0.685	0.550	0.847	0.402	3
22	0.741	0.589	0.748	0.694	0.596	0.830	0.416	5
23	0.740	0.590	0.754	0,690	0.581	0.835	0.416	5
24	0.740	0.590	0.754	0.690	0.581	0.835	0.416	5
25	0.748	0.581	0.753	0.686	0.543	0.849	0.409	4
26	0.722	0.611	0.742	0.705	0.594	0.824	0.460	9
27	0.727	0.604	0.752	0.706	0.579	0.835	0.429	8
28	0.729	0.602	0.739	0.702	0.575	0.839	0.425	6
29	0.718	0.615	0.746	0.702	0.589	0.822	0.449	8
30	0.741	0.589	0.752	0.688	0.585	0.829	0.422	5

R_{tr}^2: Correlation coefficient of training set; s_{tr}: standard error of training set; R_c^2: Correlation coefficient of calibration set; s_c: standard error of calibration set; R_{ts}^2: Correlation coefficient of test set; s_{ts}: standard error of test set; $R_m^2{}_{av}$: Modified correlation coefficient; *T*: Threshold

Without considering influence of various cyclic rings

The best model was developed without influence of cyclic rings of the molecular systems on the inhibitory activity as given below. The best model was selected based on best Monte Carlo optimization runs.

$$pK_i = 0.001(\pm 0.010) + 0.016(\pm 0.0001) \times DCW(5,4) \quad (5)$$

Training set: n = 203; R^2 = 0.643; s = 0.692; F = 362; Q^2 = 0.637; R_m^2 = 0.535; ${}^cR^2_p$ = 0.642

Calibration set: n = 69; R^2 = 0.685; s = 0.742; F = 145; R_m^2 = 0.521; ${}^cR^2_p$ = 0.676

Test set: n = 69; R^2 = 0.616; s = 0.761, F = 107; R_m^2 = 0.539; ${}^cR^2_p$ = 0.609

External set: n = 70; R^2 =0.581; s = 0.683, F = 314; R_m^2 = 0.556

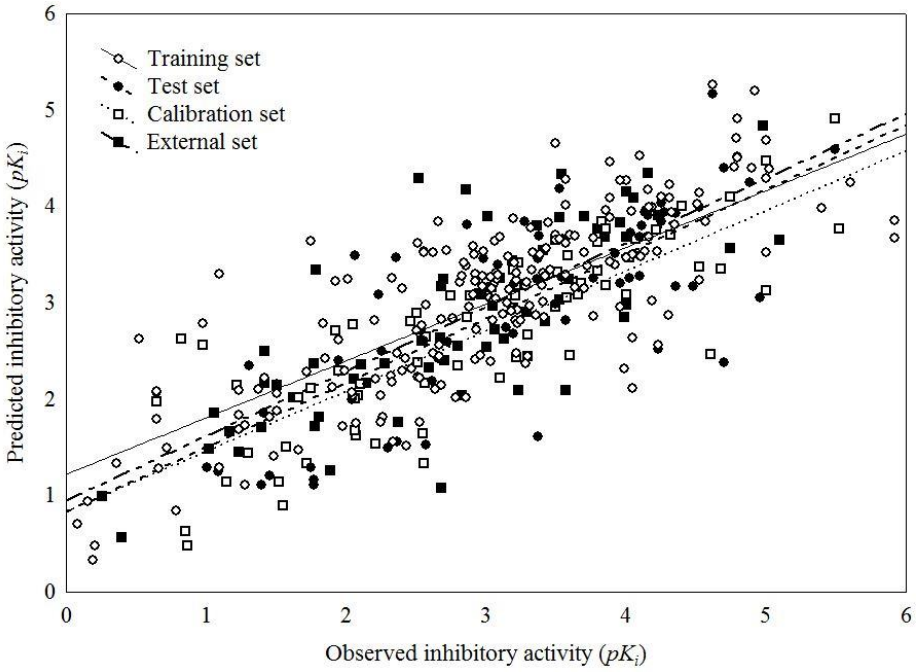

Fig. 1: Observed and predicted inhibitory activity

The statistical value of the model undoubtedly explained that model is statically robust in nature and capable enough to predict the inhibitory activity of the external set of molecules. The experimental and predicted inhibitory activity according the model are given in the Fig. 1 and Table S1 (Supplementary file). The fitness of experimental and predicted activity can be explained by plotting radar plot. For the training, test, calibration and external molecules the radar plots are given in Fig. 2. The radar plot clearly showed the closeness between the experimental and predicted activity.

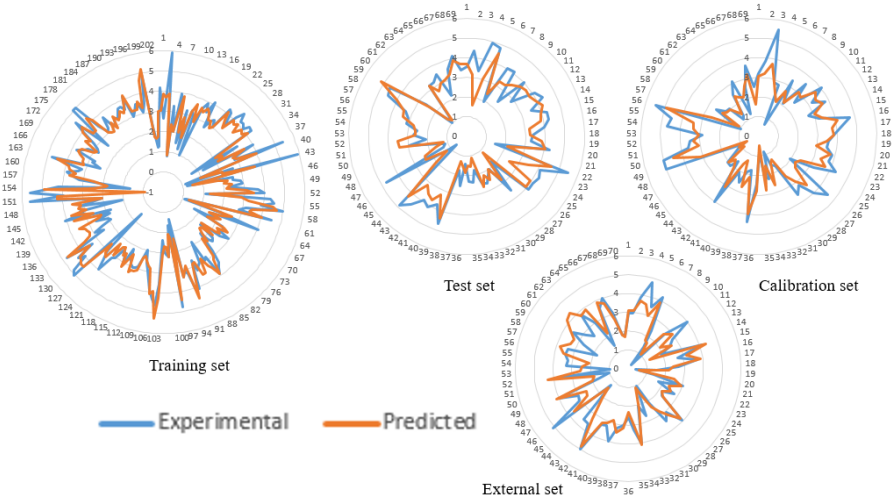

Fig. 2: Radar plot showing fitness of observed and predicted inhibitory activity of training, test, calibration and external sets

The detailed exploration of DCW from the best model developed without considering any influence of cyclic rings on the inhibitory activity explained that components "+ + + +F- - -N= = =", "+ + + +Cl- - - S = = =" and "+ + + +S- - - SB3 = = =" were found to be positive impact, while "+ + + +F- - - S = = =" showed negative impact on the inhibitory activity. The "BOND10000000" component was also showed positive influence towards the inhibition of BACE1. Presence of chlorine which characterized by the component "HALO01000000" increases the pK_i but presence of fluorine, bromine and iodine showed nothing considerable impact on pK_i. Presence of nitrogen and oxygen together give positive influence on inhibition of BACE1. On other hand impact

of nitrogen, oxygen and sulphur together ("NOSP11100000") showed negative impact towards the pK_i.

With considering influence of various cyclic rings

The best QSAR model with the influence of cycling rings on inhibitory activity was developed and given below. The best model was selected based on best Monte Carlo optimization runs.

$$pK_i = -0.020(\pm 0.009) + 0.029(\pm 0.0001) \times DCW(5,6) \quad (6)$$

Training set: n = 203; R^2 = 0.693; s = 0.639; F = 458; Q^2 = 0.690; R_m^2 = 0.581; ${}^cR^2_p$= 0.693
Calibration set: n = 69; R^2 = 0.718; s = 0.711; F = 171; R_m^2 = 0.593; ${}^cR^2_p$ = 0.714
Test set: n = 69; R^2 = 0.589; s = 0.781; F = 96; R_m^2 = 0.567; ${}^cR^2_p$ = 0.576
External set: n = 70; R^2 =0.660; s = 0.621; F = 52; R_m^2 = 0.588

The quality of the model was adjudged by the statistical parameters. The correlation coefficient between the experimental and predicted activity of training, test, calibration and external sets were found to be 0.693, 0.718, 0.589 and 0.660 respectively. The cross-validated correlation coefficient (Q^2) of training set was found to be 0.690 which clearly indicated the robustness of the model. The experimental and predicted inhibitory activity of the training, test, calibration and external set are given in Fig. 3 and Table S2 (Supplementary file). The radar plot (Fig. 4) also explained closeness between the experimental and predicted activity.

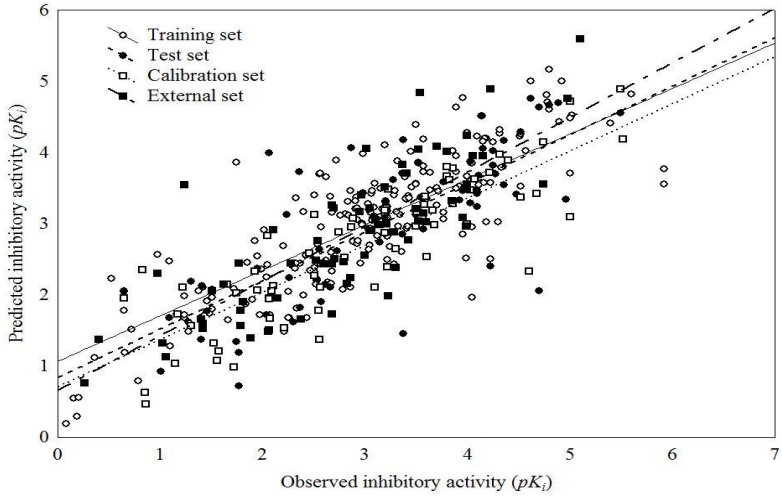

Fig. 3: Observed and predicted inhibitory activity

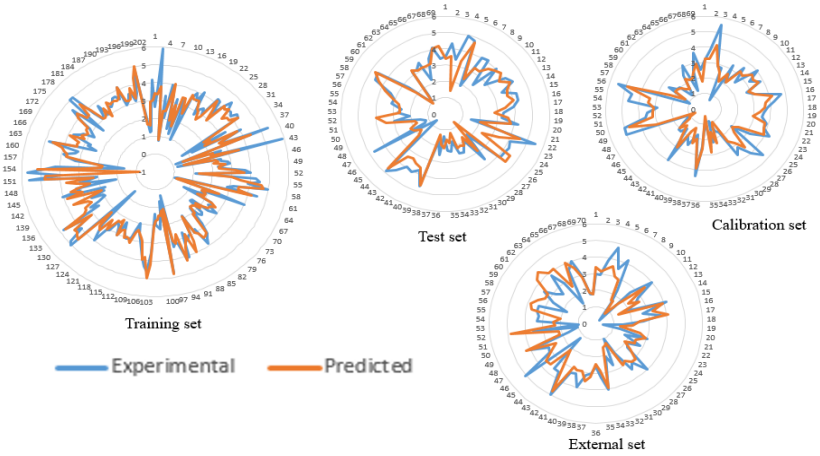

Fig. 4: Radar plot showing fitness of observed and predicted inhibitory activity of training, test, calibration and external sets

After model development the DCW were explored in details. It was observed that "+ + + +F- - -S= = =" showed negative influence for inhibition of BACE1 whereas as "+ + + +O- - -S= = =" increases the value of pK_i. Both ""BOND10000000" and "BOND11000000" give positive impact on pK_i. Only presence of fluorine was found to be favorable. High positive DCW value of all three components, "NOSP10000000", "NOSP11000000" and "NOSP11100000" explained that presence of nitrogen, oxygen and sulphur are crucial for the inhibition of BACE1 enzyme.

In comparison of both models it can be found that the correlation coefficient of training, test, calibration and external sets showed higher value for the model developed with considering the influence of cyclic rings compare to the model developed without considering influence of cyclic rings on the inhibitory activity. The improved R_{m^2avg} value also found in case of model developed with considering the cyclic rings. Hence from the statistical parameters it can be undoubtedly explained that presence of cyclic rings in the BACE1 inhibitors have importance for inhibition of BACE1.

CONCLUSION

The well validated QSAR models were developed from a large dataset of BACE1 inhibitors collected from the Binding database. The online CORAL software which is based on Monte Carlo algorithm was used to develop the QSAR models. The SMILES molecular representation was considered for the molecular descriptors. Two models were developed such as with- and without-influence of cyclic rings to the inhibitory activity. Statistical parameters obtained from the both models explained that cyclic rings are crucial of the BACE1 inhibitors for the potential inhibition of BACE1 enzyme. Both models have mechanistic interpretation in terms of molecular fragments on the DCW for increase or decrease the inhibitory activity. Finally, it can be concluded that explained important molecular fragments can play important role to design new promising BACE1 inhibitors for the therapeutic application in Alzheimer's affected community.

ACKNOWLEDGEMENT

One of the authors, Md L. Islam would like to thank Head of the Department, Department of Computer Engineering, MH Saboo Siddik College of Engineering, Byculla, Mumbai for providing research facilities.

REFERENCES

[1] Santana, I., Farinha, F., Freitas, S., Rodrigues, V., Carvalho, A. [The Epidemiology of Dementia and Alzheimer Disease in Portugal: Estimations of Prevalence and Treatment-Costs]. Acta Med Port. 2015, 28, 182-8.

[2] Alzheimer's, A. 2015 Alzheimer's disease facts and figures. Alzheimers Dement. 2015, 11, 332-84.

[3] Chiang, K., Koo, E.H. Emerging therapeutics for Alzheimer's disease. Annu Rev Pharmacol Toxicol. 2014, 54, 381-405.

[4] Selkoe, D.J. Alzheimer's disease: genes, proteins, and therapy. Physiol Rev. 2001, 81, 741-66.

[5] Ghosh, A.K., Osswald, H.L. BACE1 (beta-secretase) inhibitors for the treatment of Alzheimer's disease. Chem Soc Rev. 2014, 43, 6765-813.

[6] Vassar, R., Bennett, B.D., Babu-Khan, S., Kahn, S., Mendiaz, E.A., Denis, P., et al. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. Science. 1999, 286, 735-41.

[7] Cui, H., Hung, A.C., Klaver, D.W., Suzuki, T., Freeman, C., Narkowicz, C., et al. Effects of heparin and enoxaparin on APP processing and Abeta production in primary cortical neurons from Tg2576 mice. PLoS One. 2011, 6, e23007.

[8] Luo, Y., Bolon, B., Kahn, S., Bennett, B.D., Babu-Khan, S., Denis, P., et al. Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation. Nat Neurosci. 2001, 4, 231-2.

[9] Christensen, M.A., Zhou, W., Qing, H., Lehman, A., Philipsen, S., Song, W. Transcriptional regulation of BACE1, the beta-amyloid precursor protein beta-secretase, by Sp1. Mol Cell Biol. 2004, 24, 865-74.

[10] Nishitomi, K., Sakaguchi, G., Horikoshi, Y., Gray, A.J., Maeda, M., Hirata-Fukae, C., et al. BACE1 inhibition reduces endogenous Abeta and alters APP processing in wild-type mice. J Neurochem. 2006, 99, 1555-63.

[11] Zivkovic, J.V., Trutic, N.V., Veselinovic, J.B., Nikolic, G.M., Veselinovic, A.M. Monte Carlo method based QSAR modeling of maleimide derivatives as glycogen synthase kinase-3beta inhibitors. Computers in biology and medicine. 2015, 64, 276-82.

[12] Roy, K., Kar, S., Das, R.N. A Primer on QSAR/QSPR Modeling Fundamental Concepts, Springer International Publishing, 2015.

[13] Duchowicz, P.R., Comelli, N.C., Ortiz, E.V., Castro, E.A. QSAR study for carcinogenicity in a large set of organic compounds. Current drug safety. 2012, 7, 282-8.

[14] Talevi, A., Bellera, C.L., Di Ianni, M., Duchowicz, P.R., Bruno-Blanch, L.E., Castro, E.A. An integrated drug development approach applying topological descriptors. Current computer-aided drug design. 2012, 8, 172-81.

[15] Randic, M., Basak, S.C. A new descriptor for structure-property and structure-activity correlations. Journal of

chemical information and computer sciences. 2001, 41, 650-6.

[16] Junkes Bda, S., Arruda, A.C., Yunes, R.A., Porto, L.C., Heinzen, V.E. Semi-empirical topological index: a tool for QSPR/QSAR studies. Journal of molecular modeling. 2005, 11, 128-34.

[17] Ivanciuc, O. Chemical graphs, molecular matrices and topological indices in chemoinformatics and quantitative structure-activity relationships. Current computer-aided drug design. 2013, 9, 153-63.

[18] Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. Journal of chemical information and modeling. 1988, 28, 31-6.

[19] Weininger, D., Weininger, A., Weininger, J.L. SMILES. 2. Algorithm for generation of unique SMILES notation Journal of chemical information and modeling. 1989, 29, 97-101.

[20] Weininger, D. SMILES. 3. Depict. Graphical depiction of chemical structures. Journal of chemical information and modeling. 1990, 30, 237-43.

[21] Toropov, A.A., Benfenati, E. SMILES in QSPR/QSAR Modeling: results and perspectives. Current drug discovery technologies. 2007, 4, 77-116.

[22] Toropov, A.A., Benfenati, E. SMILES as an alternative to the graph in QSAR modelling of bee toxicity. Computational biology and chemistry. 2007, 31, 57-60.

[23] Veselinovic, A.M., Veselinovic, J.B., Zivkovic, J.V., Nikolic, G.M. Application of SMILES Notation Based Optimal Descriptors in Drug Discovery and Design. Current topics in medicinal chemistry. 2015, 15, 1768-79.

[24] Toropov, A.A., Toropova, A.P., Diaza, R.G., Benfenati, E., Gini, G. SMILES-based optimal descriptors: QSAR modeling of estrogen receptor binding affinity by correlation balance. Structural Chemistry. 2012, 23, 529–44.

[25] Veselinovic, A.M., Milosavljevic, J.B., Toropov, A.A., Nikolic, G.M. SMILES-based QSAR models for the calcium channel-antagonistic effect of 1,4-dihydropyridines. Archiv der Pharmazie. 2013, 346, 134-9.

[26] Veselinovic, A.M., Milosavljevic, J.B., Toropov, A.A., Nikolic, G.M. SMILES-based QSAR model for arylpiperazines as high-affinity 5-HT(1A) receptor ligands using CORAL. European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences. 2013, 48, 532-41.

[27] Veselinovic, J.B., Toropov, A.A., Toropova, A.P., Nikolic, G.M., Veselinovic, A.M. Monte carlo method-based QSAR modeling of penicillins binding to human serum proteins. Arch Pharm (Weinheim). 2015, 348, 62-7.

[28] Worachartcheewan, A., Mandi, P., Prachayasittikul, V., Toropova, A.P., Toropov, A.A., Nantasenamat, C. Large-scale QSAR study of aromatase inhibitors using SMILES-based descriptors. Chemometrics and Intelligent Laboratory Systems. 2014, 138, 120-6.

[29] Li, Q., Ding, X., Si, H., Gao, H. QSAR model based on SMILES of inhibitory rate of 2, 3-diarylpropenoic acids on AKR1C3. Chemometrics and Intelligent Laboratory Systems. 2014, 139, 120-6.

[30] Toropova, A.P., Toropov, A.A., Veselinovic, J.B., Miljkovic, F.N., Veselinovic, A.M. QSAR models for HEPT derivates as NNRTI inhibitors based on Monte Carlo method. European journal of medicinal chemistry. 2014, 77, 298-305.

[31] Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews. 2001, 46, 3-26.

[32] Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., Kopple, K.D. Molecular properties that influence the oral bioavailability of drug candidates. Journal of medicinal chemistry. 2002, 45, 2615-23.

[33] Toropova, M.A., Toropov, A.A., Raska, I., Jr., Raskova, M. Searching therapeutic agents for treatment of Alzheimer disease using the Monte Carlo method. Comput Biol Med. 2015, 64, 148-54.

[34] Kubinyi, H., Hamprecht, F.A., Mietzner, T. Three-dimensional quantitative similarity-activity relationships (3D QSiAR) from SEAL similarity matrices. Journal of medicinal chemistry. 1998, 41, 2553-64.

[35] Roy, K., Mitra, I., Kar, S., Ojha, P.K., Das, R.N., Kabir, H. Comparative studies on some metrics for external validation of QSPR models. Journal of chemical information and modeling. 2012, 52, 396-408.

[36] Ojha, P.K., Mitra, I., Das, R.N., Roy, K. Further exploring rm2 metrics for validation of QSPR models. Chemometrics and Intelligent Laboratory Systems. 2011, 107, 194-205.

[37] Ojha, P.K., Roy, R. Comparative QSARs for antimalarial endochins: importance of descriptor-thinning and noise reduction prior to feature selection. Chemometrics and Intelligent Laboratory Systems. 2011, 109, 146-61.

Received:04.05.18, Accepted: 08.06.18, Published:01.07.2018

***Corresponding Author:**
Md Lutful Islam*
Email: lutful.islam@gmail.com