



QSAR and Molecular Docking-Based Design of Novel Fatty Acid Amide Hydrolase Inhibitors as Anti-Alzheimer Agents

Smita Jain^{1#}, Gulam Muhammad Khan^{1,2##}, Neha Chauhan¹, Ajita Paliwal¹, Sarvesh Paliwal¹, Swapnil Sharma¹, Shailendra Paliwal⁵, Abhay Bhardwaj³, Seema V. Pattewar⁴

¹Department of Pharmacy, Banasthali Vidyapith, Banasthali-304022, Rajasthan, India

²School of Health and Allied Sciences, Pokhara University, Dhungepatan, Pokhara-30, Nepal

³Department of Pharmacy, KIET School of Pharmacy, Ghaziabad-201206, Uttar Pradesh, India

⁴Department of Pharmacy, Sanjivani College of Pharmaceutical Education and Research, At-Sahajanandnagar, Ahmednagar- 423603, Maharashtra, India

⁵Department of Pharmacy, Lala Lajpat Rai Memorial Medical College, Meerut-250002, Uttar Pradesh, India

#These authors contributed equally to this work

***Corresponding Author:**

Smita Jain

Department of Pharmacy Banasthali Vidyapith Banasthali, India-304022

Email: smitajain1994@gmail.com

***Corresponding Author:**

Gulam Muhammad Khan

Department of Pharmacy Banasthali Vidyapith Banasthali, India-304022

Email: gulamkhan@gmail.com

Abstract

Aim: To examine the role of Fatty acid amide hydrolase (FAAH) inhibitors in the treatment of Alzheimer's disease, attempts have been made to find potent inhibitors of FAAH enzyme by a 2D quantitative structure–activity relationship (QSAR) model.

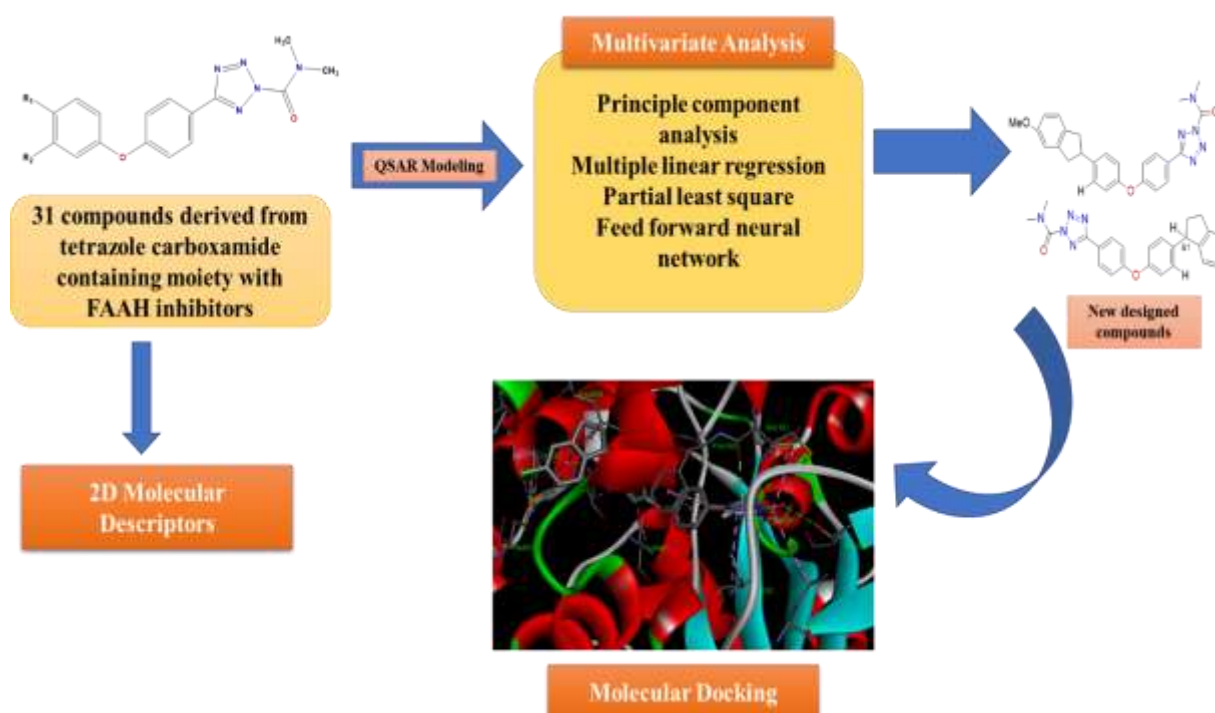
Materials and Methods: QSAR studies were performed on (4-Phenocyclophenyl tetrazolecarboxamide motif, which was aligned for generation of a QSAR based model. The 2D QSAR model was developed using partial component analysis, multiple linear regression (MLR), partial least square (PLS) and forward feed neural network (FFNN). After the development of the robust model, new compounds were designed and molecular docking was performed for understanding the confirmation of binding interactions with the FAAH enzyme.

Results: The best MLR statistical expressions revealed good predictive and authenticated ability with values $s = 0.273971$, $f = 79.0831$, $r = 0.950236$, $r^2 = 0.903$, $r^2_{cv} = 0.857917$. The r^2

(training and test-set) values of MLR, PLS and FFNN were found as 0.903, 0.9004, 0.8737, 0.8263 and 0.9772, 0.9366 respectively, which in turn revealed the soundness of the model. Ten novel compounds as FAAH inhibitors were created based on the results of the models. Molecular docking analysis validated the affinity of novel compound's and impressive binding interactions with the FAAH enzyme's active regions.

Conclusion: The values of standard statistical parameters reveal the predictive power and robustness of this model. We anticipate that this research will be very useful in lead optimization for early drug development of new comparable compounds.

Keywords: Quantitative structure-activity relationship, Descriptors, Lipinski's rule of five, Multiple linear regression, Alzheimer's disease, Molecular docking



Graphical abstract

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative, neuropsychiatric disorder that inescapably causes memory and learning deficits. According to the current statistics, four million Indian suffer from dementia and the numbers rises up to 44 million on a global scale, declaring it an alarming global health crisis. It has also been predicted that Alzheimer's and dementia burden will rise up to an alarming count of 7.5 million by the year 2030 [1].

Histologically, AD concurrently rummages cellular electrical charge transport and neurotransmitter activity that majorly transpires as a result of amyloid beta ($A\beta$) plaques accumulation and neurofibrillary tangles in the brain [2,3]. Gradually, each $A\beta$ molecule cluster up to form plaques resulting in blockage of synaptic cell-to-cell signalling and immune system activation. Further, the formation of neurofibrillary tangles occurs as an outcome of breakdown of tau proteins that is prerequisite for sustaining cellular transportation networks. As a result of plaques and neurofibrillary tangles formation, synaptic dysfunction and brain cell fatality is detected. The above-mentioned plaques and tangles have

been extensively researched which postulated a working theory that justified the cell death and tissue loss observed in AD brain [4] however, the theory still remains to be irrefutably confirmed.

The endocannabinoid (eCB) system reacts as an adaptive response to various pathological situations in the body by activating pleiotropic and pro-homeostatic signalling system. This system operates by functional, morphological and phenotypic changes in microglia and monitors neuroinflammation in several neurodegenerative diseases such as Alzheimer disease, Parkinson disease, Huntington's disease etc. The major parameters of eCB system are cannabinoid type receptors {(CB₁ and CB₂), lipid ligands (N-arachidonylethanolamine (AEA), and 2-Arachidonoylglycerol (2-AG)}, biosynthetic proteins and enzymes. In a recent study, activation of CB₁₋₂ receptors showed A β inhibition, improvement in cognitive impairments in transgenic AD mice model. However, the study reported associated adverse effects such as neurotoxicity, neuroinflammation and gliosis. Earlier studies have suggested that targeting eCB degradation could be considered as reasonable therapeutic options in AD [5,6]. Recently, scientists have reported that FAAH (FAAH, EC 3.5.1.99) degrades AEA, resulting in the elevated expressions of microglia-associated neurotic plaques and astrocytes of post-mortem brains of AD patients [7,8]. The genetic ablation of FAAH reported therapeutic benefits with respect to neuroinflammation, cognitive decline and A β accumulation, reflecting the modulation of microglial phenotypes, analgesic, anxiolytic, anti-depressant, loss of memory and anti-inflammatory activity through commencement of various mechanisms without showing the inadmissible side effects of direct cannabinoid receptor agonists [4,9]

The conventional management of AD includes drug categories such as donepezil, rivastigmine, galantamine and memantine for palliative care [10]. However, their usage does not reverse or prevent any pathologic processes. The usage of these drugs poses serious threats to clinicians and to the quality of life of patient due to observable adverse effects such as diarrhoea, dizziness, headache, muscle cramping, constipation and mental confusion. These medications does not serve as an effective treatment for the mitigation of the disease or improvement in its prognosis. Hence, there is an urgent need to develop some novel candidates for the management of AD.

To overcome the shortcomings of conventional anti-AD therapy, review of binding requirements of anti-Alzheimer's drugs using computational techniques has become an obligation. Quantitative structure-activity relationship analysis (QSAR) is a popular method for studying the connection between ligand structures and binding affinities. It helps researchers minimize money and time while creating novel medications and chemicals. Utilizing the QSAR method may increase the chances of success and decrease the time and money spent on the drug development process [11]. There is a common misconception that QSAR is equivalent to chemo-metrics or multivariate statistical analysis. Occasionally, its use is constrained by the presence of Hansch analysis, which employs multiple regression correlation methods to explore the quantitative association between the biological activity of a set of compounds and their physicochemical substituent or global characteristics representing their hydrophobic, electronic, steric, and other impacts [12]. QSAR studies have predictive ability.

Several FAAH inhibitors have been identified for the treatment of AD and few compounds are under clinical trial. However, the need of introducing some novel pharmacologically active compounds has become essential in the treatment of AD. Therefore, the current study targets to determine the structural elements of (4-phenoxyphenyl) tetrazolecarboxamide motif derivatives. Thus, in the present study we have designed a novel compound i.e., FAAH inhibitor, employing computer aided drug design (CADD) targeting cannabinoid receptor. In addition, we have evaluated the extent of neural networks that complements the regressive study of cannabinoid receptor functioning in the management of AD.

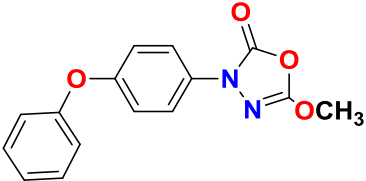
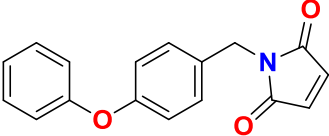
2. EXPERIMENTAL METHODOLOGY

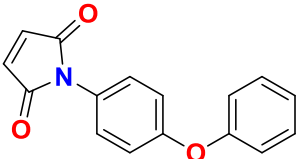
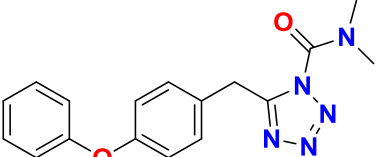
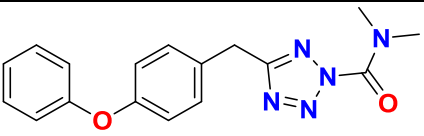
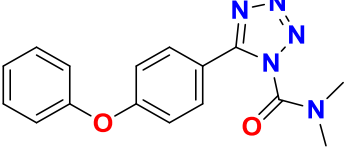
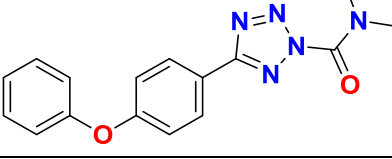
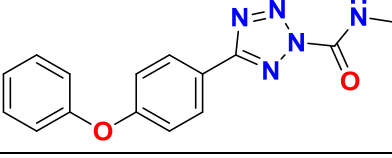
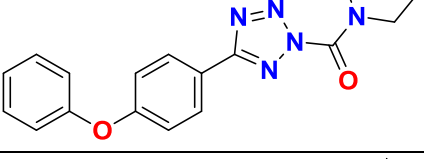
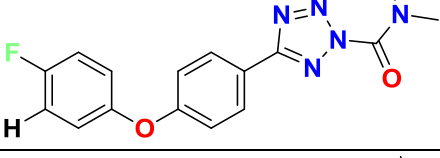
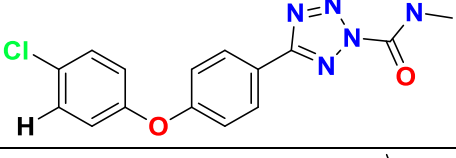
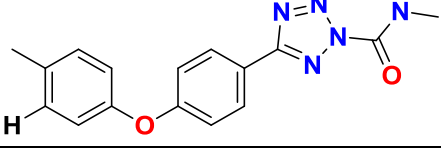
QSAR analysis was executed on (4-Phenocyclophenyl) tetrazolecarboxamide motif that consisted of 32 compounds. After analysis, it was found that one compound lacked an exact IC_{50} value and hence, was excluded from further analysis. Therefore, a data set of final 31 compounds was developed for modelling [13]

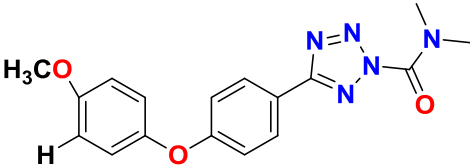
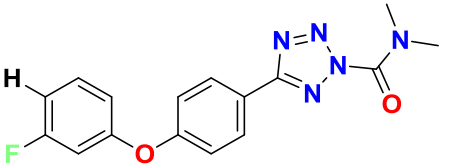
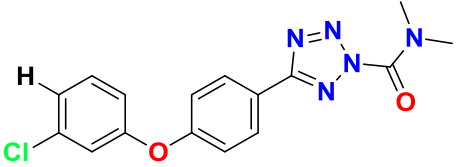
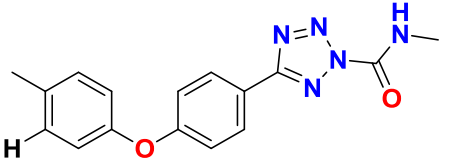
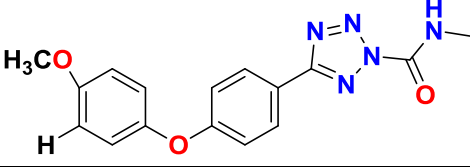
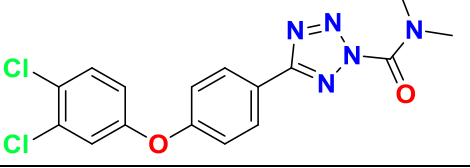
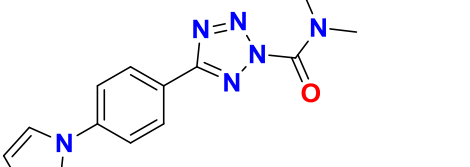
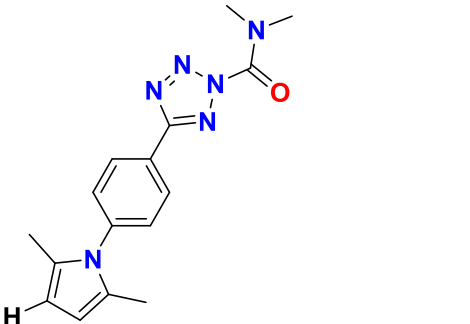
2.1 Input preparation for studies of QSAR

Energy reduction was used after the molecular structures were generated and their geometries were arranged using a stand-alone module of Discovery Studio (Version 2.0). All of the structures were imported into the TSAR (Tool structure activity relationship) (Version 3.3, Accelrys Inc., Oxford, England) spreadsheet and given descriptive names. In addition, a chemical encoding system was proposed, in which molecules were characterised as templates, each of which may have any number of substituents connected to it through a single bond. It is also possible for a single hydrogen atom to function as a substitute. There are several possible ways to use a template representation of a molecule to display its properties after substituents have been added. As shown in Table S1, each substituent was assigned a number that corresponds to its location inside the molecule. Using the Corina-make 3D option, we were able to convert all of the structures and the defined substituents into high-quality 3D representations that account for all of the relevant factors, such as total energy, valence terms (such as bond, bond angle, and torsional potential), and non-bonded terms (electrostatic and vander-waals interactions) [14]. Charges were calculated using charge-2 package available with TSAR.

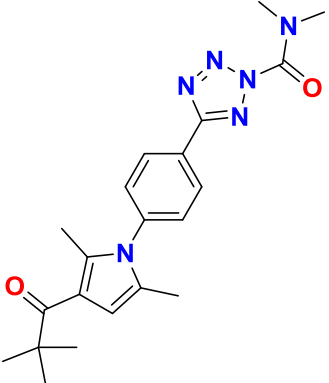
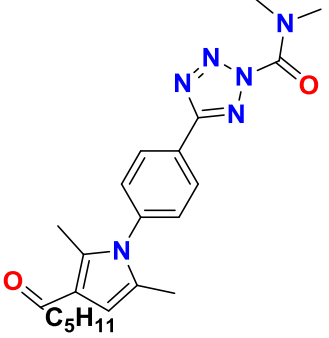
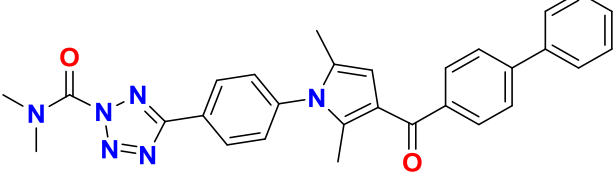
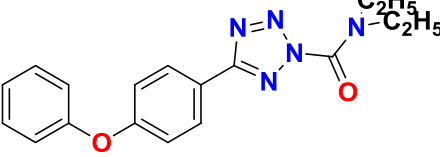
Table S1. Structure and biological activity data of (4-Phenocyclophenyl) tetrazolecarboxamide motif derivative analogues used in QSAR analysis

Compound	Structure	IC_{50} value (μ M)
1		0.72
2		4.9

3		5.3
4		0.006
5		0.099
6		0.007
7		0.012
8		0.021
9		0.015
10		0.014
11		0.016
12		0.024

13		0.013
14		0.010
15		0.109
16		0.013
17		0.015
18		0.011
19		0.005
20		0.057

21		0.009
22		0.008
23		0.012
24		0.14
25		0.088
26		0.019
27		0.016

28		0.029
29		0.011
30		0.040
31		0.037

2.2 Preparation of dataset and calculation of descriptors

In the current study, the IC_{50} (μM) values of selected 31 compounds were converted to negative logarithm of IC_{50} (i.e., $pIC_{50} = \log 1/IC_{50}$). Then the data set was divided randomly into 2 sets which consisted of training and test sets. The purpose of training set was to develop the QSAR model (includes 2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21 and 22 compounds) whereas test set was used for the validation of developed model (includes 1,4,23,24,25,26,27,28,29,30,31 compounds).

Descriptors are numerical amounts generated for reflecting the molecular configuration and grasping the systemic, steric, technological and multilateral dimensions of an operation. It provides complete useful information on every chemical and its substituents in molecular descriptor calculation. The character of descriptors used and their encoding of structural characteristics of biological activity provide a steady decision on the quality of QSAR analysis (use of proper description and molecular descriptors in QSAR studies is an integral step). In the current study, molecular descriptors were determined for the selected compounds with the aid of TSAR 3.3 software. Then the molecular descriptors were

determined for whole molecules and their identified substituents. The following class of descriptors is determined by TSAR: molecular attributes, dipole moments [15], molecular indices, wiener index [16] molecular connectivity indices [17], molecular shape indices [18], electrotopological state indices [19], LogP [20], number of defined atoms (carbons, oxygens etc.), and number of defined groups (methyl, hydroxyl etc.). For both the whole molecule (wholistic descriptors) and substituents, the plurality of these descriptors is calculable. Some descriptors are available that can only be determined for substitutes. The above included criteria apply on verloop [21], dipole bond and lipole bond. Both substances have been discarded descriptors with the same meanings. The remaining descriptors were examined in parallel. The correlation coefficient was determined for each pair of descriptors. If the correlation coefficient was greater than 0.6 for two descriptors, the regression for each of them was checked with biological activity. Descriptor having high regression value were retained and used for building model [22]

2.3 Principal component analysis (PCA)

The principal component analysis (PCA) is design based on data analysis for the transformation of a variable set which is correlated to a new variable set known as principal components. These are few in number but serve independently. These new variables can be used to reduce the system's dimensionality with a minimality in loss of information (XLSTAT software). PCA can also be used for the determination of nonlinearity and non-multicollinearity between variables and to select corresponding descriptors activity.

2.4 Multivariate analysis

Two different categories can be represented for the actual used parameters: Either the structure can be treated as whole entity or it can be treated as combination of fragments (substituents) with clear chemical meaning. This fact entails two kinds of parameterization theories consisting of the whole molecular approach which is always applicable to theory and the latter being the fragment approach, which can be sometimes inapplicable to theory but is practical in several cases. In specific fragment-based approach, the series structure is interpreted by a build-up of common backbone and by variation of its substituents. Only the substituents can be parameterized. On the other hand, the multi-parameter approach is utilized for providing a statistically satisfactory description of the data from fragment-based approach. In the current study, multivariate analysis was performed using MLR, PLS and FFNN.

2.5 Linear regression analysis

A linear mathematical function that relates descriptor values to the activity may be created using multiple linear regressions and partial least square. For n observations and p independent variables the general linear regression model is represented as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

Where $\beta_0, \beta_1, \beta_2, \beta_p$ are regression coefficients, X_1, X_2, \dots, X_p , are the independent variables (numerical descriptors), and Y is the dependent variable of interest (biological activity).

There are several statistical parameters for the selection of the best linear model. The basic method of assessing a QSAR model is to evaluate the r^2 (squared correlation) of the operation and a set of independent descriptors. The r^2 is well known to increase as the number of variables increases, so that more complex models are more precise. However, a good QSAR

model with a high r^2 r^2_{cv} & F is considered. R^2 is the fitness of the model; r^2_{cv} is the cross-validated r^2 (QSAR model quality measurement) [23]. The cross-validated r^2 is obtained through leave-out or leave-out (LOO) as described by Cramer et al. F (Fischer statistics) is in fact the ratio of explicit to unexplained variance for a certain number of degrees of freedom. It therefore indicates a true relation or the degree of importance for the MLR models. If F reached a maximum and a minimum standard error the simulation is considered optimal. With the inclusion of one additional variable, a substantial decline in F may lead to the new descriptor not being as important as anticipated, i.e., its inclusion has endangered the statistical quality of the combination [24]. Another reliability check used was value of standard deviation (S), which should be low. To maximize the predictability of the model, some outliers have been removed from the training set by maximizing the r^2_{cv} values. The outlier's free data set was used to construct the final model and the quality of the model was checked using various statistical parameters.

2.6 Feed forward neural network (FFNN) analysis

FFNN is an integrated layered processing unit system for directed transfer and data processing. FFNN has been stated to be equivalent or superior to MLR in reliable forecasts. By extension, MLR assumes or specifically integrates a linear relationship between binding affinities and molecular descriptors [25]. In comparison, FFNN does not presume that a problem is linear. The key benefit of FFNN is that QSAR can be established without the analytical form of a specific correlation model being defined [26]

The ADME properties were checked for the inhibitors and they revealed that the molecules possessed suitable pharmacokinetic profiles. Lipinski's "rule of five", postulated by Christopher A. Lipinski (1997) was applied on the entire data set. This rule states that under given established circumstances such as medication drugs being lipophilic and relatively smaller molecules [27][28] the drug-likeness is demonstrated as per certain parameters. For a molecule to be more drug-like, it should have a molecular weight >500, log P value >5, the number of hydrogen bond donors >5, the number of hydrogen bond acceptors >10 and it should have poor permeability and absorption rate [27]. The Lipinski's rule of five provides an approach based on molecular properties of a molecule which reflect its pharmacokinetics inside the human body. Under the given factors, the drug likeliness of the selected molecules was calculated by implementation of Lipinski's rule of five and is depicted in Table S2.

Table S2. Values of calculated parameters for Lipinski's rule of five

Compound	ADME weight (Whole molecule)	ADME H bond acceptors	ADME H bond donors	ADME Log P Whole molecule	ADME Rotatable Bond	ADME violations
1	284.29	5	0	4.1502	4	0
2	279.31	3	0	2.7959	4	0
3	265.28	3	0	2.701	3	0
4	323.39	5	0	3.3783	4	0
5	323.39	5	0	3.4937	4	0
6	309.36	5	0	3.3262	3	0
7	309.36	5	0	3.4416	3	0

8	309.36	5	0	3.4416	3	0
9	309.36	5	0	3.4416	3	0
10	327.35	5	0	3.5811	3	0
11	343.8	5	0	3.9596	3	0
12	323.39	5	0	3.9088	3	0
13	339.39	6	0	3.1889	4	0
14	327.35	5	0	3.5811	3	0
15	343.8	5	0	3.9596	3	0
16	323.39	5	0	3.9088	3	0
17	339.39	6	0	3.1889	4	0
18	378.24	5	0	4.4776	3	0
19	312.42	4	0	4.4765	2	0
20	310.4	4	0	3.2089	2	0
21	340.48	4	0	4.5055	2	0
22	313.41	5	0	3.7971	2	0
23	362.48	4	0	5.2429	2	1
24	324.43	4	0	3.0715	3	0
25	346.43	4	0	3.8089	3	0
26	352.44	5	0	2.5174	3	0
27	414.51	5	0	4.4303	4	0
28	394.53	5	0	4.3731	4	0
29	408.56	5	0	4.3345	7	0
30	490.61	5	0	6.1147	5	1
31	337.42	5	0	4.1266	5	0

2.7 Molecular docking

Using PyRx Autodock version 4.0, the discovered compounds were docked to the enzyme fatty acid amide hydrolase (FAAH) [29,30]. The structure of the FAAH enzyme was obtained from Protein Databank (PDB ID 2vya). All of the discovered compounds were assembled in an Autodock version 4.0 (Autodock vina) folder of the PyRx programme, and grid batch docking was also done.

3. Results and Discussion

3.1 Principal component analysis (PCA)

Principle component analysis (PCA) was used to three descriptors that coded for 31 compounds. The first three axes F1, F2 and F3 represent 47.83%, 36.10%, 52.17% of the total variance respectively and the estimation of total information was 83.93% (Figure 1).

The PCA was performed to identify the link between various descriptors and to have better knowledge of the chemical distribution [31]. Table 6 displays the correlation's matrix of the three descriptors. The correlation coefficients in the resulting matrix indicate whether the descriptors have a high or low connection. According to the findings, there was excellent co-

linearity ($r > 0.5$) between the bulk of the variables and the three descriptors that were highly linked with each other.

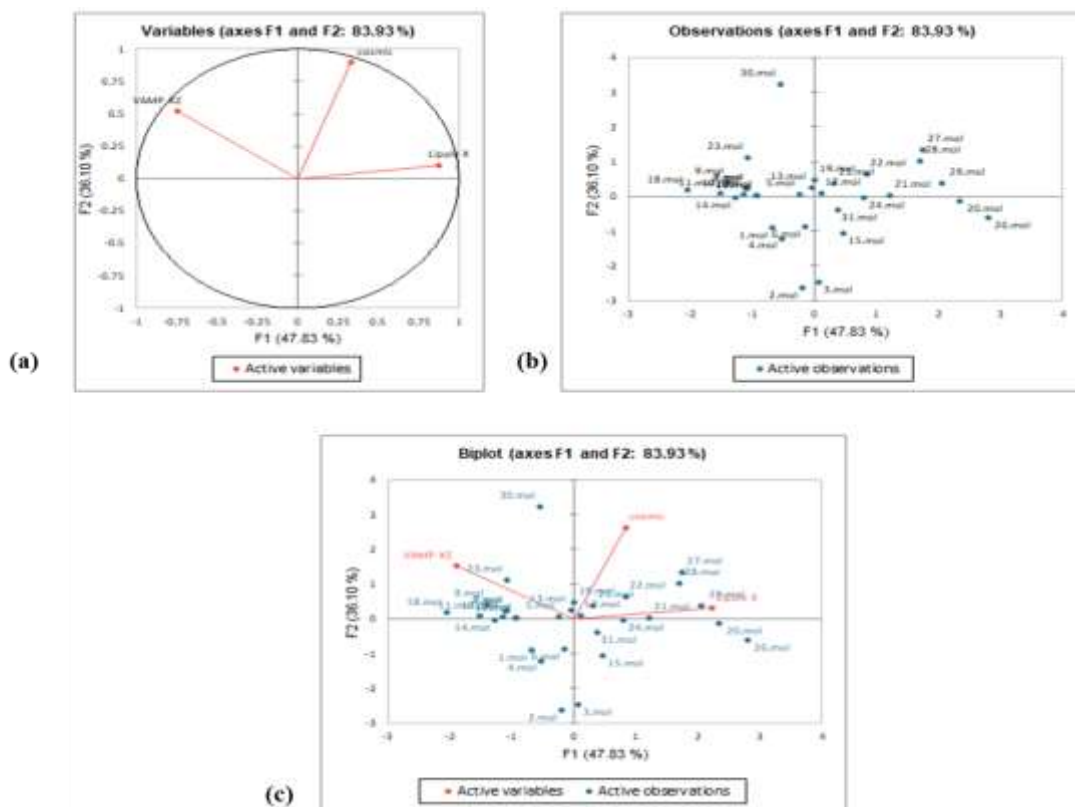


Figure 1. (a) Correlation matrix circle, (b) Observations, (c) Biplot

3.2 Linear multivariate analysis

The QSAR equations were derived by multiple linear regression (MLR) and partial least squares (PLS). Three descriptors were identified after data reduction which were independent of each other and were used to develop regression model.

QSAR outliers can be very significant and interesting, especially when the biological activity observed is greater than the one predicted by the established model (more residual value). Outliers may be present quite reasonably such as incorrect parameter estimation, a mathematical model may not be suitable, there may be a lack of certain descriptors to show the QSAR for all of the compounds, or there may even be a lack of different modes of mechanisms. In the current study, four outliers (4, 23, 25 and 31) were detected with the help of regression line of equation. The statistical characteristics including LOO r^2_{cv} and correlation coefficient r , for the training set were substantially improved after these compounds had been omitted from the training set. These compounds were therefore finally removed from the training set and models of MLR & PLS for the remaining 42 compounds were produced. Best QSAR models (MLR & PLS) obtained with excellent r^2_{cv} and equation-1 illustrate the r values for the training set.

Equation 1 (MLR)

$$Y = 0.046420071 * X_1 - 0.038273897 * X_2 - 0.01517441 * X_3 - 0.54302436$$

Where, X_1 is cosmic total energy (whole molecule), X_2 is lipole X component (whole molecule), X_3 is VAMP polarization XZ (whole molecule)

MLR on the training set compounds with the three specified descriptors exhibited a progressive increase in statistical values, as shown in Table 1 and satisfactory r^2 values of (training and test) proved the model's resilience (Figure 2).

Table 1. Statistical tests and their values obtained after performing MLR analysis

ANALYSIS. Statistical tests	Values
s value	0.273971
F value	79.0831
Regression coefficient, r	0.950236
r^2	0.903
Cross validation, r^2 (cv)	0.857917

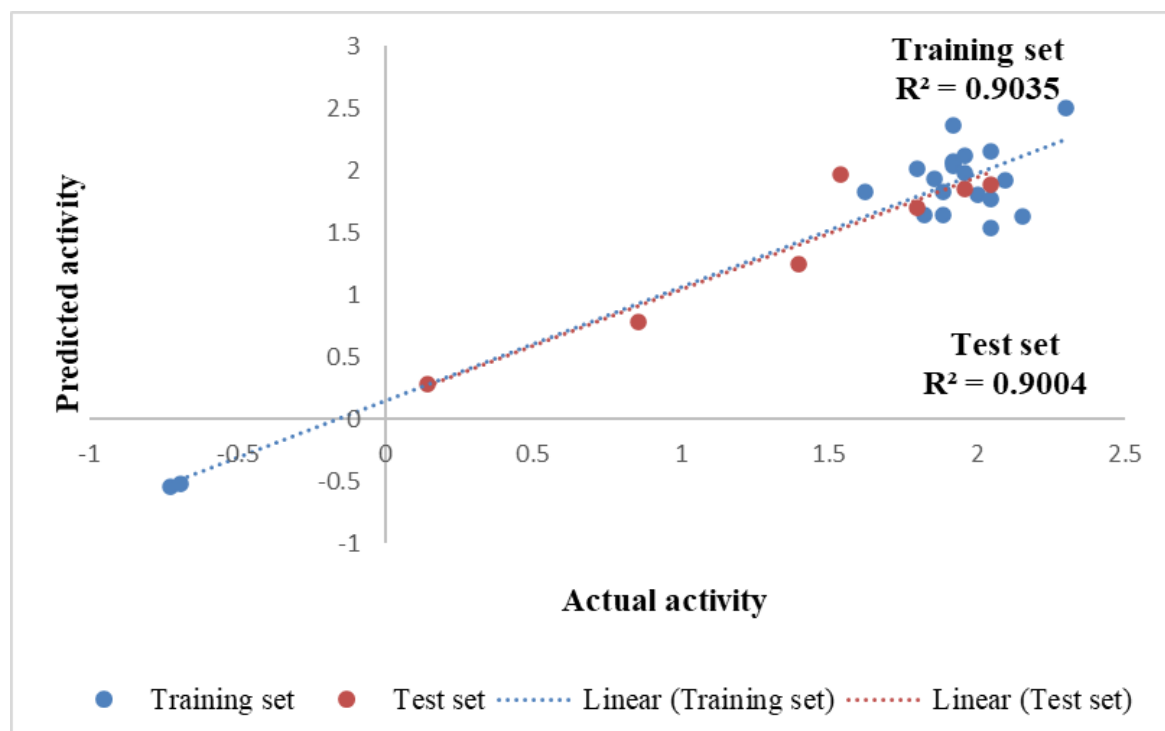


Figure 2. Plot of actual versus predicted activity for the training set and test set of compounds derived from MLR analysis

Results obtained from conventional MLR were checked with PLS analysis using same data set. PLS model (equation 2) complemented the MLR model in terms of model soundness (r^2) and predictability (r^2 cv).

Equation 2 (PLS)

$$Y = 0.046420068 * X1 - 0.038274251 * X2 - 0.015174526 * X3 - 0.51215571$$

Where, X1 is cosmic total energy (whole molecule), X2 is lipole X component (whole molecule), X3 is VAMP polarization XZ (whole molecule)

PLS conducted on the training set compounds with the three specified descriptors resulted in progressive increase in statistical values, as shown in Table 2, and the satisfactory r^2 values of (training and test) supports the resilience of the model (Figure 3)

Table 2. Statistical tests and their values obtained after performing PLS analysis

ANALYSIS. Statistical tests	Values
s value	1.2168
r^2	0.8737
Cross validation, r^2 (cv)	0.85762

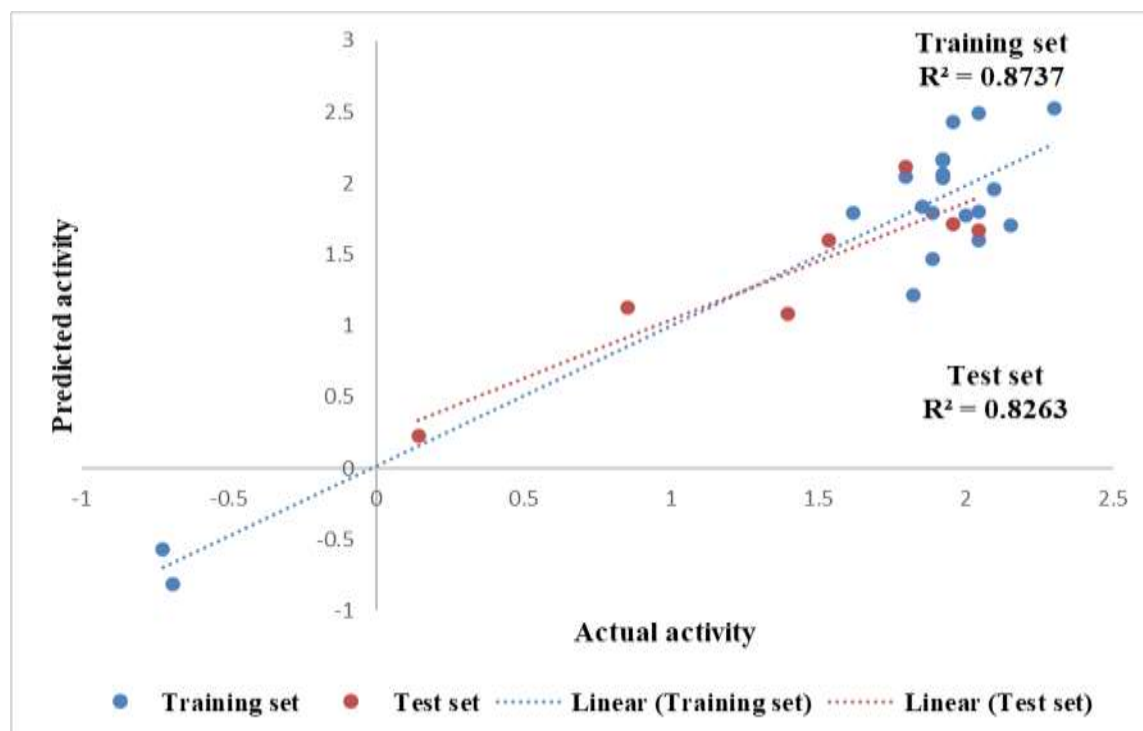


Figure 3. Plot of actual versus predicted activity for the training set and test set of compounds derived from PLS analysis

Since both MLR and PLS should produce comparable results for a well-defined problem, the r^2 and r^2_{cv} values of the MLR and PLS models were calculated and compared with square correlation coefficient value and then were defined in both models of $r^2 = 0.903$ for MLR & $r^2 = 0.8737$ for PLS, which is a relative fitness calculation by the equation of regression. The strong cross-validated square correlation value of $r^2_{cv} = 0.857917$ for MLR & r^2_{cv} value of $= 0.85762$ for PLS also demonstrated strong internal model predictivity.

The standard deviation from the established model was determined by the mean error square representing the shift of residual or regression line variance. Therefore, standard deviation is an absolute indicator of fitness efficiency and should have a low return value. The s-value of the final MLR model was found to be 0.273971, which shows the high magnitude of the model. The f-value was derived from the f-test that implied the likelihood of a true relationship or the strength of the model. The f value for a certain degree of freedom is the ratio of the stated to unexplained variance. The higher the f-value, the higher the likelihood of statistically significant QSAR model. In general, f-values of various MLR models are compared until the standard is reached, which does not indicate any further change in f-value. The f-value of the final MLR model was observed at 79.0831, which clearly showed the significance of the derived model. A statistical significance value of 1.2168 further signified the robustness of developed model.

3.3 Feed forward neural network analysis

The neural network models were used to investigate the form of relationships between the biological data and molecular descriptors. The findings of this study indicated that both approaches can be used to build predictive models more efficiently. At 1608 cycles the best RMS fit for the model was 0.031237. The predictive capacity was assessed by the plot of expected versus experimental affinities between training and tests for model compounds that is shown in Figure 4. The results were shown on a 2D plot of the input output node. A neural network of three layers has been used in this analysis. Input descriptors were the same as multivariate regression (MLR and PLS). The trained neural network architecture provided a close correlation coefficient for the training set (r^2 training=0.9772). Figure S1 (a, b and c) show the dependency plots obtained from FFNN. The study of all the compounds showed that the relationship between biological and three descriptor activities was linear and was close to the study of MLR and PLS. Details of FNN are shown in Table 3.

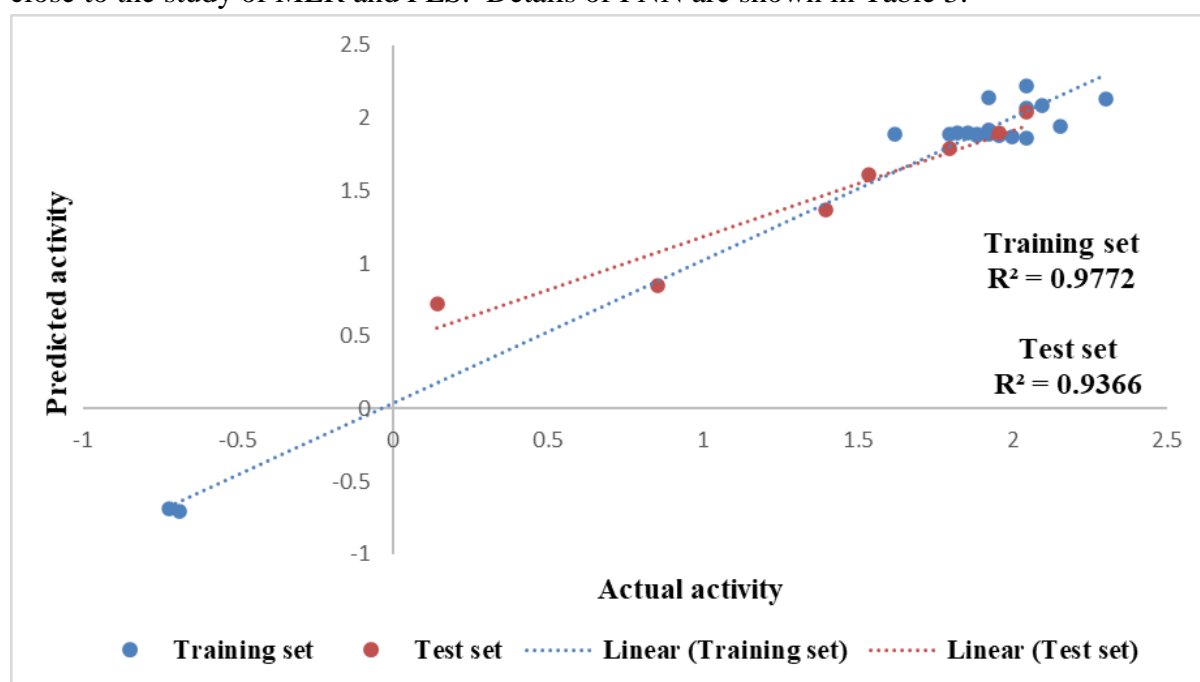


Figure 4. Plot of actual versus predicted activity for the training set and test set of compounds derived from FFNN analysis

Table 3. Details of FNN

Summary of FNN	
No of configuration	3-2-1
Test RMS fit	0.061426
No of cycle	1608
Best RMS fit	0.031237

3.4 Test set prediction

The external predictive capability of QSAR model was also regulated by test sets of compounds excluded during model development. The test set viewed all compounds in a way similar to the control group compounds. The r^2 value of MLR= 0.9004 and PLS= 0.8263

extracted for the test set showed the high predictive capability of the developed model. The predictivity of MLR was compared to FFNN using same external test set and value of r^2 was found to be comparative (0.9366). Table 4 and 5 shows the actual and predicted activity of MLR, PLS and FFNN analysis for compound training and testing and their corresponding graphs are shown in Figure 2-4.

The three parameters with high correlation were left on TSAR sheet and were further employed for the generation of regression equation. They were further analysed for relative impacts for the compound activity which is tabulated on Table 6. The conclusion was made as per t-test, Jackknife SE and Covariance SE values mentioned on Table 7 and was observed to be significant for the best model. Hence, the importance of each descriptor was evaluated and confirmed.

Table 4. Actual and predicted activity data obtained from linear and non-linear analysis for training set of compounds

Compounds No.	Actual Value (Log 1/c)	Predictive value		
		MLR	PLS	FNN
2	-0.6902	-0.519407	-0.80927	-0.70395
3	-0.72428	-0.540578	-0.56536	-0.687871
5	2.04576	1.53257	1.60138	1.86276
6	2.1549	1.63246	1.70959	1.94409
7	1.92082	2.05588	2.15921	1.91596
8	1.92082	2.04133	2.0366	1.91318
9	1.92082	2.07645	2.06671	1.89333
10	1.85387	1.93499	1.83755	1.8959
11	1.79588	2.01469	2.04899	1.8919
12	1.61979	1.82257	1.79073	1.88548
13	1.88606	1.63704	1.46904	1.88179
4	2.00	1.8016	1.77822	1.87006
15	2.04576	1.76708	1.80206	2.06701
16	1.88606	1.82656	1.78969	1.88697
17	1.82391	1.63988	1.21857	1.89815
18	1.95861	2.11417	2.43471	1.88167
19	2.30103	2.50586	2.52947	2.13014
20	2.04576	2.15022	2.4907	2.22381
21	2.09691	1.92166	1.95532	2.08254
22	1.92082	2.36806	2.17115	2.14021

Table 5. Actual and predicted activity data obtained from linear and non-linear analysis for test set of compounds

Compounds No.	Actual Value (Log 1/c)	Predictive value		
		MLR	PLS	FNN
1	0.142668	0.283701	0.224799	0.725111
24	0.853872	0.785342	1.12951	0.853405
26	2.04576	1.88988	1.66772	2.04089
27	1.79588	1.7039	2.11246	1.77993
28	1.5376	1.96916	1.60407	1.60992
29	1.95861	1.85123	1.71556	1.90406
30	1.39794	1.24912	1.0822	1.36917

Table 6. Correlation matrix showing correlation between biological activity and parameters used

	Log value	Cosmic Total Energy (Whole Molecule)	Lipole Component (Whole Molecule)	X	VAMP Polarization XZ (Whole Molecule)
Log value	1	0.65145	-0.23028		-0.035671
Cosmic Total Energy (Whole Molecule)	0.65145	1	0.24378		0.094894
Lipole Component (Whole Molecule)	-0.23028	0.24378	1		-0.40462
VAMP Polarization XZ (Whole Molecule)	-0.035671	0.094894	-0.40462		1

Table 7. t-test values, Jacknife SE and Covariance SE for the selected descriptors

	t-test value	Jacknife SE	Covariance SE
Cosmic Total Energy (Whole Molecule)	7.1557	0.0073159	0.0049188
Lipole Component (Whole Component)	-4.5545	0.01574	0.015789
VAMP Polarization XZ (Whole Component)	-2.8243	0.02526	0.02934

3.5 Interpretation of descriptors entered

The three parameters: cosmic total energy (whole molecule), lipole X- component (whole molecule), VAMP polarization XZ (whole molecule) showing high correlation with the biological activity were used to generate regression equation and were analysed for their relative impact on the activity of the compounds.

3.5.1 Cosmic total energy (whole molecule)

Positive cosmic energy coefficient, which is seen as a reasonable indicator of hydrogen bond strength (more energy strength than the bond), indicates that a greater cosmic energy is averse to operate [32]. The model obtained showed that this behaviour increased as a result of greater values of cosmic energy and hydrophobic parameters. Cosmic energy, the name of the molecular mechanical strength area in TSAR for energy calculations, is the most important descriptor of three parameters for describing the anti-Alzheimer behaviour of (4-Phenocyclohexyl) tetrazolecarboxamide and hydrophobic descriptors that have a fundamental impact on biochemical processes and it also influences the fate of a molecule in the binding site environment, as seen in regression equation 1, which indicates that cosmic radiation is greatly associated with biological activity. The present study (Figure S1a) shows that the anti-Alzheimer behaviour of the sequence is primarily due to cosmic energy and hydrophobic influences. By evaluation of the results, it was established that cosmic energy and hydrophobic descriptors are highly compliant to each other because higher cosmic energy is equivalent to fewer hydrogen bonding groups and more hydrophobic groups contribute positively to increase lipophilicity.

3.5.2 Lipole X component (whole molecule)

Lipophilicity plays a vital role in determining drug distribution inside the body after absorption and it also predicts the extent of rate of metabolism and excretion. Lipophilicity has also been considered as the chief driving force for the binding of drugs to their receptor targets [33]. Lipole serves as a quantitative descriptor of molecular lipophilicity. Lipole X component is a directional component of lipophilicity. A positive contribution of lipole X component towards biological activity provides a better justification about the inhibitory activity of compounds with increment of bulky lipophilic group in whole compound. Since in the FFNN dependence graphs of the developed model (Figure S1b) the value of lipole X component is negatively correlated, an increase in the value of lipole X component resulted in decrease of FAAH inhibitory activity. The most active compound showed a lower value of lipole X component while the least active compound showed high lipole X component therefore, the inverse relation between the biological activity and lipole X component was justified. This is consistent with the negative coefficients found in the final MLR and PLS models.

3.5.3 Vamp polarization XZ (whole molecule)

VAMP polarization XZ provides an optimal semi-empirical molecular orbital package that is employed in the calculation of electrostatic properties and executes structure optimizing. The polarizability is the transformation factor between an applied electric field and the induced dipole moment [34]. The static polarizability linearly correlates to number of valence electrons per atom (i.e., polarizability is parallelly connected to sum of valence electrons in a molecule. VAMP polarization XZ shows a negative correlation with the actual activity as

determined by the regression equation. Therefore, the VAMP parameter value increases with the decrease in actual activity. Since in the FFNN dependence graphs of the developed model (Figure S1c) the value of VAMP polarization XZ (whole molecule) is negatively correlated, an increase in the value of VAMP polarization XZ (whole molecule) resulted in decrease of FAAH inhibitory activity. The most active compound showed a lower value of VAMP polarization XZ (whole molecule) while the least active compound showed higher value of VAMP polarization XZ (whole molecule), thus, justifying the inverse relation between the biological activity and VAMP polarization XZ (whole molecule). This is consistent with the negative coefficients found in the final MLR and PLS models.

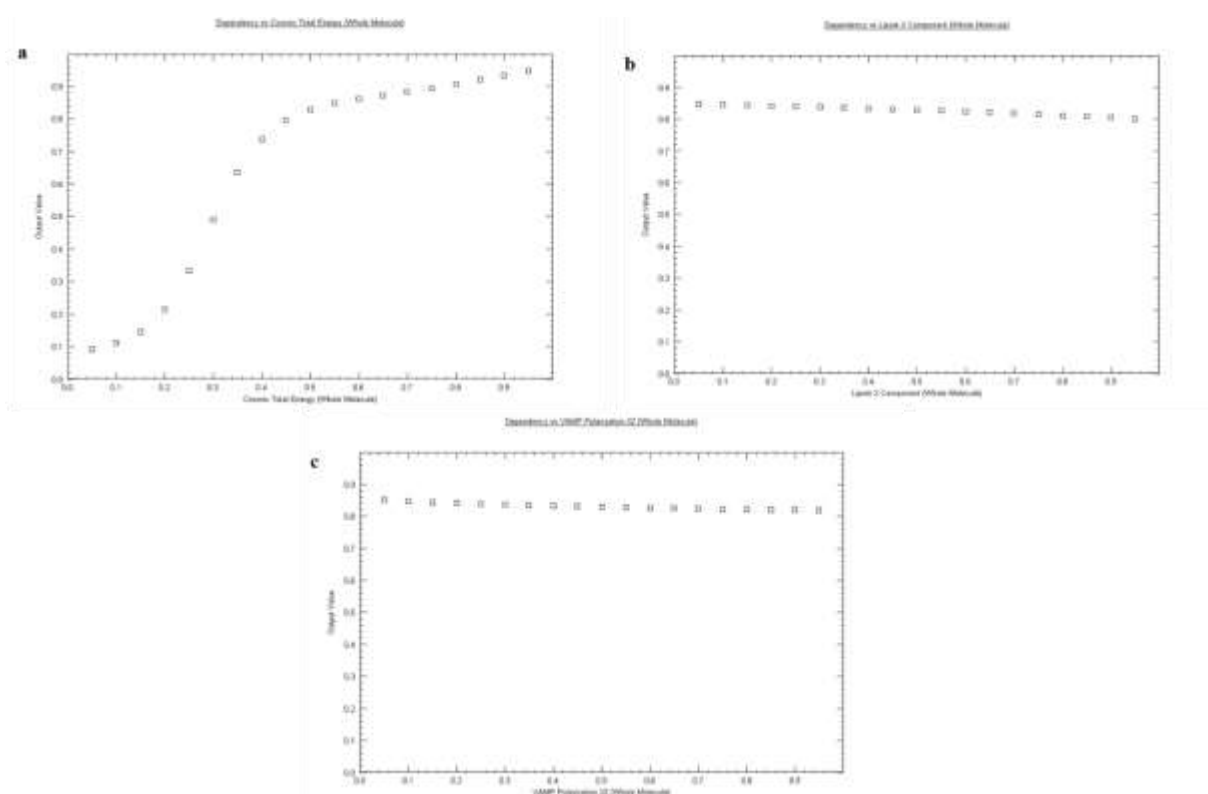
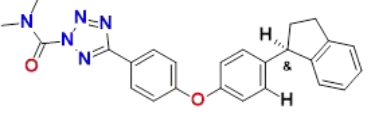
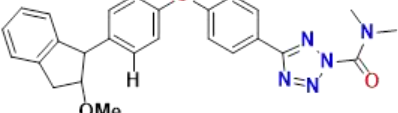
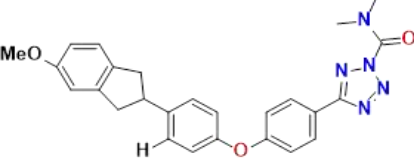
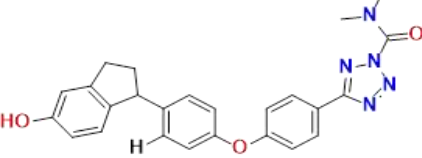
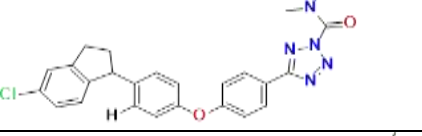
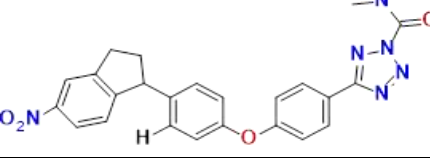
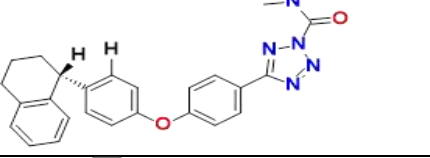
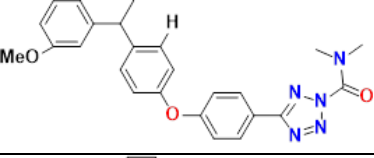
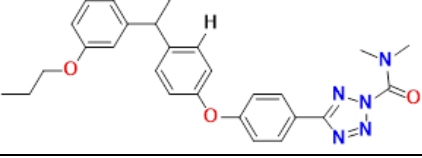


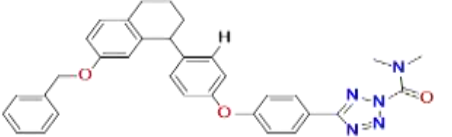
Figure S1. (a) Dependency plot between biological activity and cosmic total energy (whole molecule). (b) Dependency plot between biological activity and lipole X component (whole molecule). (c) Dependency plot between biological activity and VAMP polarization XZ (whole molecule)

3.6 Proposed novel compounds

As a result, using MLR, PLS, and ANN techniques, we designed novel compounds with distinct and better FAAH inhibitor values than the previously examined compounds (Table 8). Using the foregoing data as a guide, we made appropriate replacements and then estimated the activities of the new compounds using the suggested model in (1). The IC_{50} value of novel compounds is less than 0.009. As a result, they are considered as dependable compounds for developing novel compounds with distinct and enhanced activity values compared to the investigated compounds. These findings may be used to medication design, and the creation of new and safer medicines is recommended.

Table 8. Proposed compounds, values of calculated descriptors, and predicted values of pIC₅₀ using MLR and PLS model

Structures	Compounds	Cosmic total energy	Lipole X component	vamp polarization XZ	pIC ₅₀ (MLR)	pIC ₅₀ (PLS)
	C1	49.3533	-14.1429	2.1476	0.00554	0.00516
	C2	53.0537	-9.03516	-7.03536	0.00424	0.00395
	C3	45.745	-8.28413	-5.96482	0.01028	0.00957
	C4	42.2229	-11.2207	2.55557	0.01557	0.0145
	C5	45.8921	-16.8236	2.22872	0.00635	0.00591
	C6	57.1828	-3.23205	-8.15366	0.00438	0.00408
	C7	45.7995	-9.27874	-10.6278	0.00795	0.00741
	C8	44.4222	-8.88982	-10.7077	0.00951	0.00886
	C9	44.1413	-10.1976	-10.9426	0.00866	0.00807

	C10	47.5528	-8.26792	-5.39911	0.00866	0.00806
--	-----	---------	----------	----------	---------	---------

3.7 Molecular docking

Molecular docking evolved into the basic concept of structure-based virtual screening. The sampling algorithm predicts the several conformations, or poses, that the ligand might take inside the binding or active pocket. The produced poses are then sorted according to their binding energies, with the top ranked pose corresponding to the right ligand conformation. The protein from the protein data repository was used in our molecular docking investigation (PDB ID: 2vya). As protein, the cocrystal structure of the FAAH protein complexed with PF-750 in the active site was used. Docking studies were conducted on 10 designed compounds with high projected activities, as well as 21 compounds with high experimental activity, as indicated in Table S1. Compounds C1, C4, C5, C9, C10, and 21 had good binding energies of -11.7, -11.4, -11.6, -11.2, -11.8, and -8.6 kcal/mol, respectively, in all developed compounds. Table S3 also shows the lengths of hydrogen bonds produced between ligands and amino acids of the active receptor pocket in the range $>5\text{\AA}$ (2.11 - 5.0 \AA). As illustrated in Figures 5-11, all developed compounds had a similar interaction with one of the amino acid residues ser241 generated by strong hydrogen bonding between the tetrazole carboxamide group and oxygen groups in the amino acid. Compounds also shown hydrophobic interactions with the protein's active site pockets' Phe192, Ile238, Arg486, Ile530, Leu429, Trp531, Ile407, Leu404, Val491 and Phe192 amino acid residues.

Moreover, compound 21 exhibited an interaction with Ser241 in the active pocket by hydrogen bond as shown in Figure 10.

Similarly, the reference compound URB597 showed binding interactions similar to novel designed compounds. The docking score of the URB597 was -10.36 Kcal/mol. URB597 showed hydrogen bond interactions with Ser241, Gly485 and Leu401 amino acids residue of the active binding site of the receptor and hydrophobic bond interactions with Met191, Ile238, Phe192, Ser193, and Leu404 amino acids residues with the binding sites of the receptor (Figure 11).

These results clearly reveal that the interaction of the reference ligand and 21 are mostly the same as those observed in the five designed ligand complexes with the protein. These results indicate that C1, C4, C5, C9 and C10 ligands as selective inhibitors against FAAH enzyme activity.

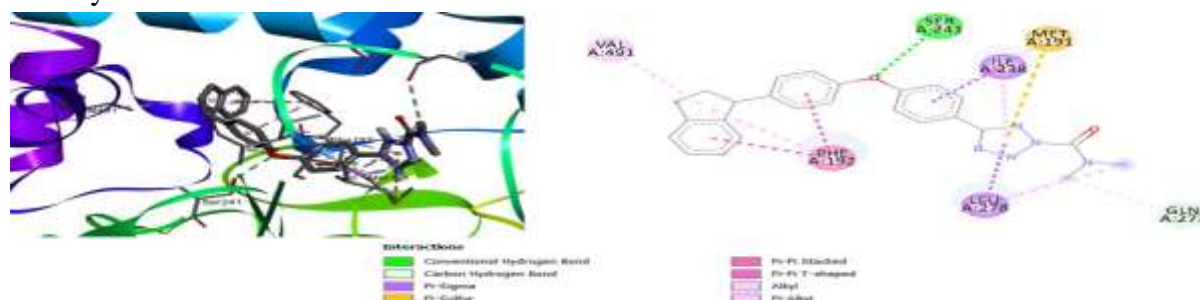


Figure 5. 3D and 2D interactions of the compound C1 with the active sites of the FAAH enzyme (2vya)

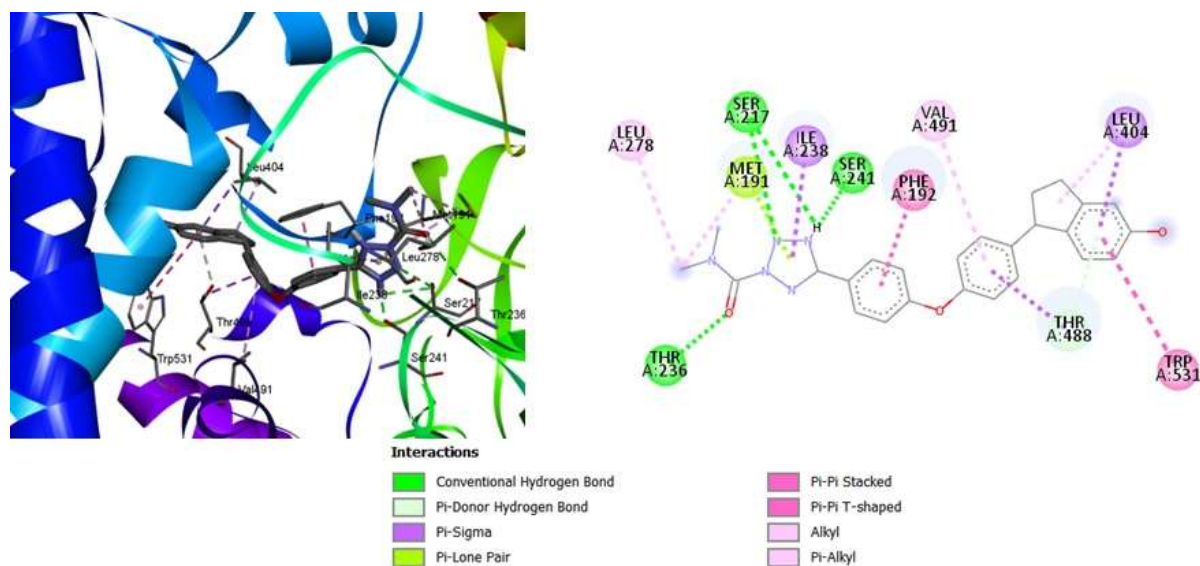


Figure 6. 3D and 2D interactions of the compound C4 with the active sites of the FAAH enzyme (2vya)

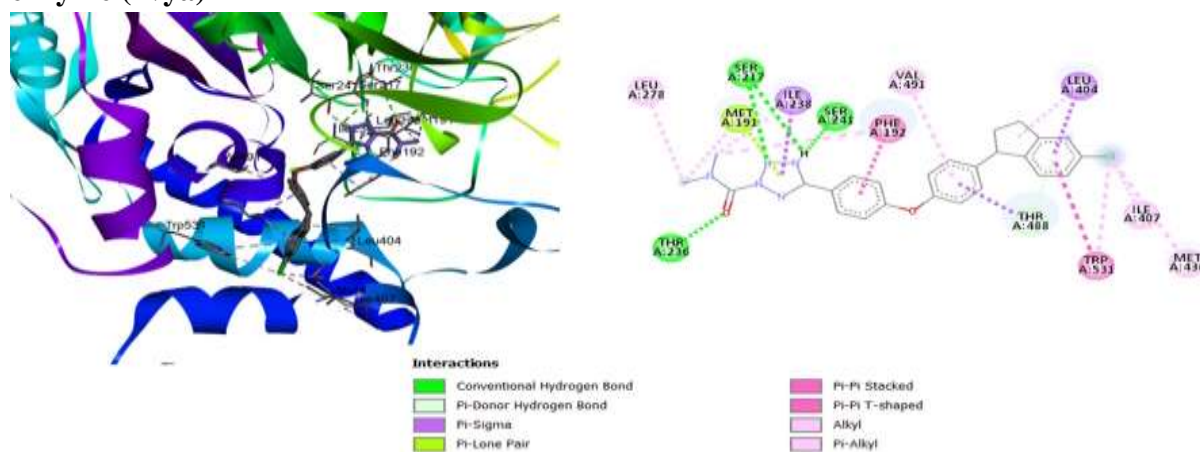


Figure 7. 3D and 2D interactions of the compound C5 with the active sites of the FAAH enzyme (2vya)

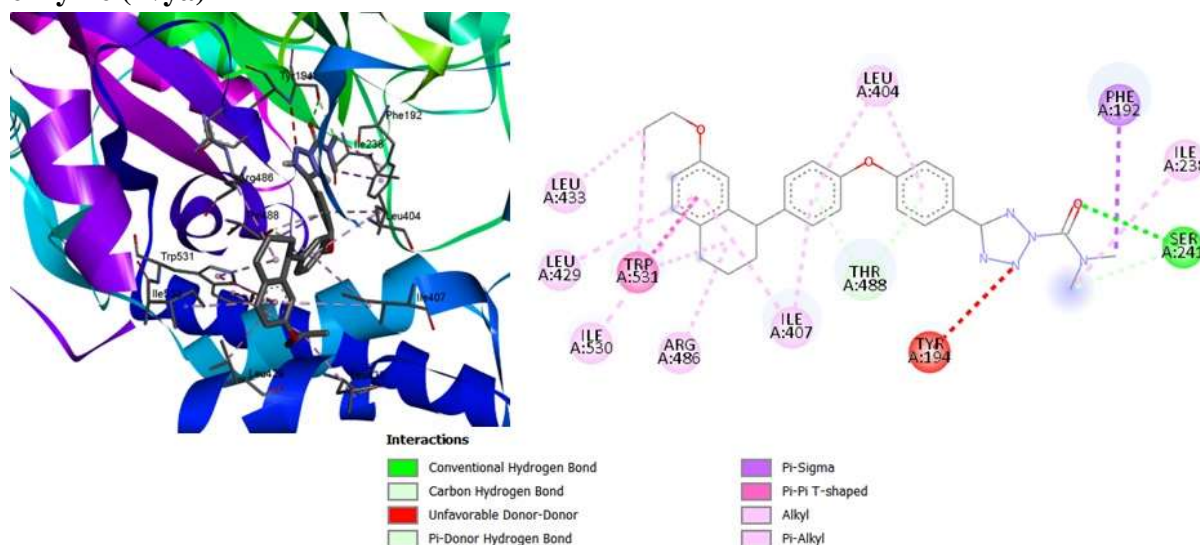


Figure 8. 3D and 2D interactions of the compound C9 with the active sites of the FAAH enzyme (2vya)

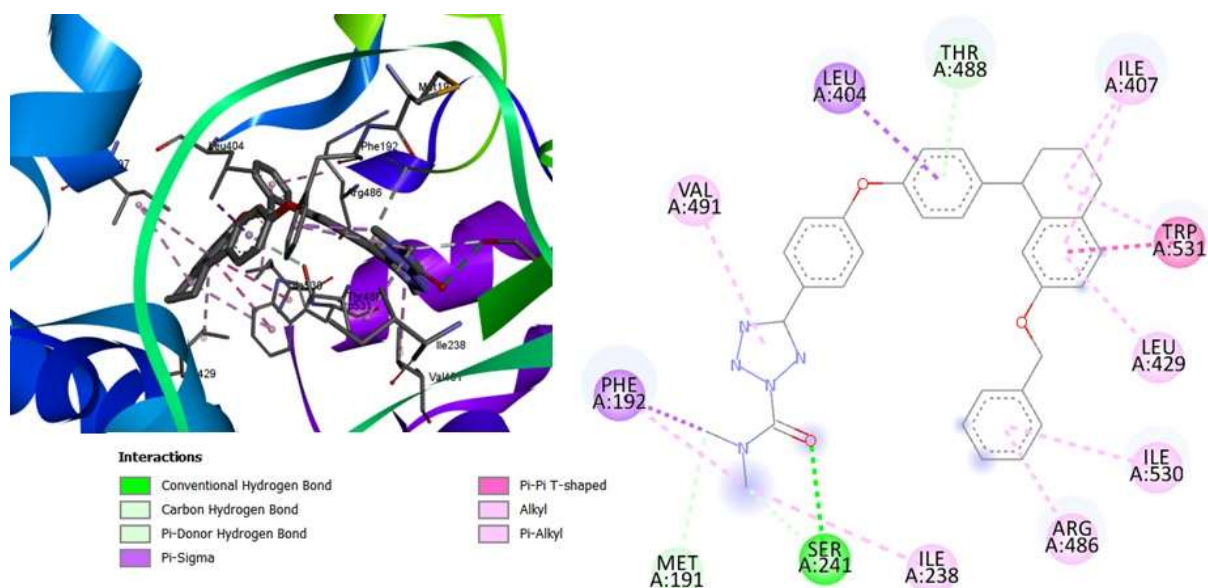


Figure 9. 3D and 2D interactions of the compound C10 with the active sites of the FAAH enzyme (2vya)

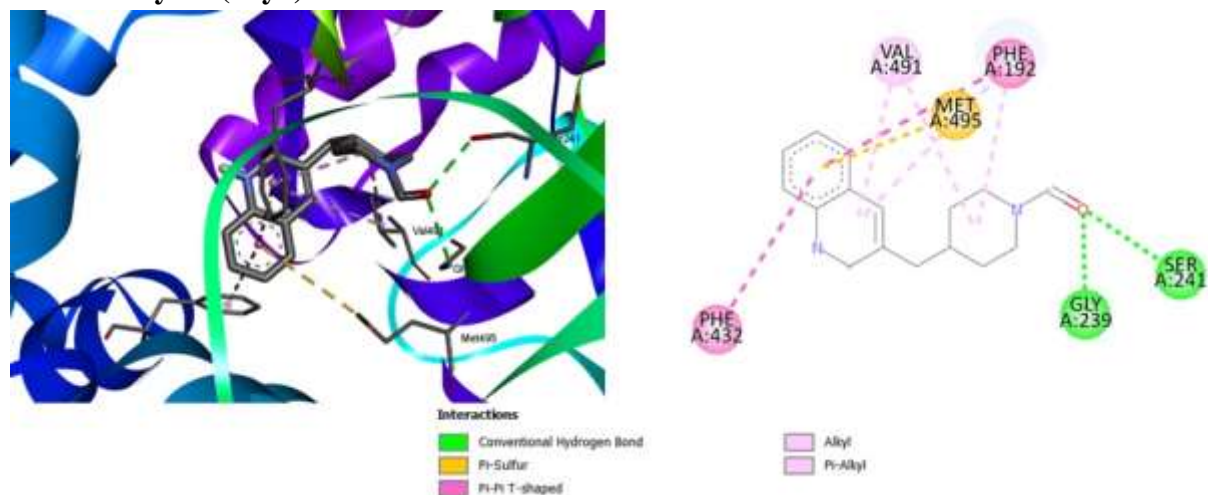


Figure 10. 3D and 2D interactions of the compound 21 with the active sites of the FAAH enzyme (2vya)

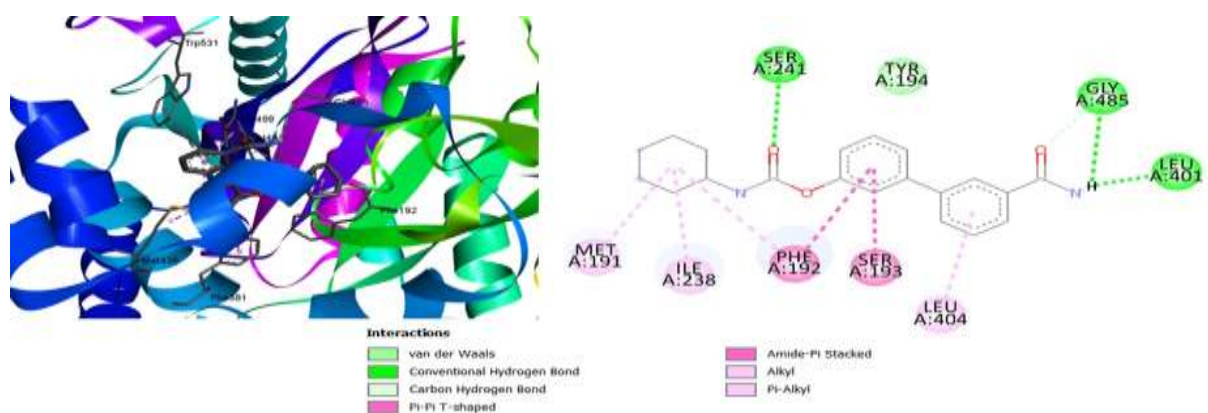


Figure 11. 3D and 2D interactions of the reference compound URB597 with the active sites of the FAAH enzyme (2vya)

Table S3. Interactions Data and Binding Free Energies for Investigated Ligands with the Active Site of the enzyme (FAAH: 2vya)

Compounds	Binding energies (kcal/mol)	Amino acid residues	H-bond distance (Å)
C1	-11.7	C-Gln273 O-Ser241	3.56519 2.70288
C2	-11.0	H-Ser241 H-Ser217 C-Cys400	1.86201 2.59126 3.67459
C3	-11.1	H-Ser241 H-Ser217	2.00638 2.5434
C4	-11.4	H-Ser241 N-Ser217 N-Ser217 O-Thr236	1.48538 3.03674 2.96529 3.1206
C5	-11.6	N-Ser217 H-Ser217 O-Thr236 H-Ser241	3.00031 3.04538 3.14517 1.45687
C6	-9.9	H-Thr488 O-Ser241	3.0455 2.6432
C7	-7.3	C-Thr488	3.14035
C8	-10.2	O-Ser241 C-Ser241 C-Thr488	2.68029 2.7122 4.12075
C9	-11.2	C-Ser241 O-Ser241 C-Thr488 C-Thr488	2.79448 2.91744 3.21422 4.05695
C10	-11.8	O-Ser241 C-Ser241 C-Met191	2.72702 3.03751 3.70748
21	-8.6	O-Ser241 O-Gly239	2.70011 2.83031

4. Conclusion

In summary, the quantitative structure-activity relationship models of the anti-Alzheimer activity of (4-Phenocyclophenyl) tetrazolecarboxamide motif derivatives were developed using statistical analytic techniques. With superior predictive power, the artificial neural network outperformed multiple linear regression and partial least squares regression. We discovered a positive relationship between various descriptors and anti-Alzheimer activity. The findings suggest that the proposed models in this research may reliably predict activity and that the descriptors used are relevant to describe this activity. The suggested models' accuracy and

predictability were shown by comparing the anticipated values of the examined activity for the various models (Table 6). The suggested approaches will minimize the time and cost of synthesis and determination of the anti-alzheimer activity of derivatives of the (4-Phenocyclophenyl) tetrazolecarboxamide motif. Furthermore, the descriptors include enough chemical, electrical, and topological information to encode structural traits that might be employed in the building of QSAR models in conjunction with other descriptors. The findings also revealed that the proposed compounds C1, C4, C5, C9, and C10 have the potential to be FAAH protein inhibitors. Lower binding energies and greater molecular interactions with pockets characterise the ligands. According to the findings, computer-aided drug design (CADD) may be an essential discovery approach for generating a novel treatment for a variety of disorders.

Finally, we conclude that the most powerful compounds identified in this investigation may be synthesised and pharmacologically evaluated to produce very potent FAAH enzyme inhibitors as anti-Alzheimer agents.

ACKNOWLEDGEMENT

The author wishes to express gratitude to vice chancellor late Prof Aditya Shastri for providing the software for the study. The author sincerely appreciates the Islamic Development Bank (IsDB) in Jeddah, Kingdom of Saudi Arabia for the Ph.D. scholarship grant (ID No: 600039032) we gratefully acknowledge this support.

CONFLICT OF INTEREST

Not applicable

ABBREVIATIONS

FAAH: Fatty acid amide hydrolase; **QSAR:** Quantitative structure activity relationship; **MLR:** Multiple linear regression; **PLS:** Partial least square; **FFNN:** Forward feed neural network; **AD:** Alzheimer's disease; **eCB:** Endocannabinoid; **A β :** Amyloid beta; **AEA:** N-arachidonoyl ethanolamine; **2-AG:** 2-Arachidonoylglycerol; **PCA:** Principal component analysis; **CV:** Cross-validation; **CB:** Cannabinoid.

REFERENCES

- [1] E. Aydin, L. Hritcu, G. Dogan, S. Hayta, E. Bagci, The effects of inhaled *Pimpinella peregrina* essential oil on scopolamine-induced memory impairment, anxiety, and depression in laboratory rats, *Molecular Neurobiology* 53 (2016) 6557–6567. <https://doi.org/10.1007/s12035-016-9693-9>.
- [2] J. Liang, S. Tian, J. Han, P. Xiong, Resveratrol as a therapeutic agent for renal fibrosis induced by unilateral ureteral obstruction, *Renal Failure* 36 (2014) 285–291. <https://doi.org/10.3109/0886022X.2013.844644>.
- [3] K. Tak, R. Sharma, V. Dave, S. Jain, S. Sharma, *Clitoria ternatea* mediated synthesis of graphene quantum dots for the treatment of Alzheimer's disease, *ACS Chemical Neuroscience* 11 (2020) 3741–3748. <https://doi.org/10.1021/acschemneuro.0c00273>.
- [4] S. Jain, A. Bisht, K. Verma, S. Negi, S. Paliwal, S. Sharma, The role of fatty acid amide hydrolase enzyme inhibitors in Alzheimer's disease, *Cell Biochemistry Function* (2021). <https://doi.org/10.1002/CBF.3680>.

- [5] M. Grieco, M.G. de Caris, E. Maggi, F. Armeli, R. Coccurello, T. Bisogno, M. D'Erme, M. Maccarrone, P. Mancini, R. Businaro, Fatty acid amide hydrolase (FAAH) inhibition modulates amyloid-beta-induced microglia polarization, *International Journal of Molecular Science* 22 (2021). <https://doi.org/10.3390/ijms22147711>.
- [6] K. Tak, P. Sharma, R. Sharma, V. Dave, S. Jain, S. Sharma, One-pot hydrothermal green synthesis of *Polygala tenuifolia* mediated graphene quantum dots for acetylcholine esterase inhibitory activity, *Journal of Drug Delivery Science and Technology* 73 (2022) 103486. <https://doi.org/10.1016/j.jddst.2022.103486>.
- [7] R.K.P. Tripathi, A perspective review on fatty acid amide hydrolase (FAAH) inhibitors as potential therapeutic agents, *European Journal of Medicinal Chemistry* 188 (2020) 111953. <https://doi.org/10.1016/j.ejmech.2019.111953>.
- [8] S. Jain, N. Chauhan, A. Bhardwaj, G. Yadaw, M.K. Singh, A. Mishra, QSAR modeling of α -ketoazazole motif analogues as potent anti-Alzheimer agents, *Ymer* 21 (2022) 624–640. <https://doi.org/10.37896/YMER21.05/71>.
- [9] S. Patel, J. Shukla, S. Jain, V. Paliwal, N. Tripathi, S. Paliwal, S. Sharma, Repositioning of tubocurarine as analgesic and anti-inflammatory agent: Exploring beyond myorelaxant activity, *Biochemical Pharmacology* 205 (2022) 115248. <https://doi.org/10.1016/j.bcp.2022.115248>.
- [10] S. Montanari, L. Scalvini, M. Bartolini, F. Belluti, S. Gobbi, V. Andrisano, A. Ligresti, V. di Marzo, S. Rivara, M. Mor, A. Bisi, A. Rampa, Fatty Acid Amide Hydrolase (FAAH), Acetylcholinesterase (AChE), and Butyrylcholinesterase (BuChE): networked targets for the development of carbamates as potential anti-alzheimer's disease agents, *Journal of Medicinal Chemistry* 59 (2016) 6387–6406. <https://doi.org/10.1021/ACS.JMEDCHEM.6B00609>.
- [11] M.M. Neaz, M. Muddassar, F.A. Pasha, S.J. Cho, 2D-QSAR of non-benzodiazepines to benzodiazepines receptor (BZR), *Medicinal Chemistry Research*. 18 (2009) 98–111. <https://doi.org/10.1007/s00044-008-9111-6>.
- [12] B. Testa, QSAR: Hansch analysis and related approaches, *Trends in Pharmacological Sciences* 16 (1995) 280. [https://doi.org/10.1016/s0165-6147\(00\)89046-x](https://doi.org/10.1016/s0165-6147(00)89046-x).
- [13] A. Holtfrerich, W. Hanekamp, M. Lehr, (4-Phenoxyphenyl)tetrazolecarboxamides and related compounds as dual inhibitors of fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase (MAGL), *European Journal of Medicinal Chemistry* 63 (2013) 64–75. <https://doi.org/10.1016/j.ejmech.2013.01.050>.
- [14] N. Dwivedi, B.N. Mishra, V.M. Katoch, 2D-QSAR model development and analysis on variant groups of anti-tuberculosis drugs, *Bioinformation* 7 (2011) 82–90. <https://doi.org/10.6026/97320630007082>.

- [15] V.K. Vays, A. Jain, M. Ghate, D. Maliwal, QSAR modeling of some substituted alkylidenepyridazin-3-one as a non-cAMP-based antiplatelet agent, *Medicinal Chemistry Research* 20 (2011) 355–363. <https://doi.org/10.1007/s00044-010-9333-2>.
- [16] W. Harry, Structural determination of paraffin boiling points, *Journal of the American Chemical Society* 69 (1947) 17-20. <http://pubs.acs.org/doi/abs/10.1021/ja01193a005>.
- [17] M. Randić, On Characterization of Molecular Branching, *Journal of the American Chemical Society* 97 (1975) 6609–6615. <https://doi.org/10.1021/ja00856a001>.
- [18] L.B. Kier, A Shape Index from Molecular Graphs, *Quantitative Structure- Activity Relationships*. 4 (1985) 109–116. <https://doi.org/10.1002/qsar.19850040303>.
- [19] M.A. de Brito, C.R. Rodrigues, J.J.V. Cirino, R.B. de Alencastro, H.C. Castro, M.G. Albuquerque, 3D-QSAR CoMFA of a series of DABO derivatives as HIV-1 reverse transcriptase non-nucleoside inhibitors, *Journal of Chemical Information and Modeling* 48 (2008) 1706–1715. <https://doi.org/10.1021/ci8001217>.
- [20] T. Fujita, J. Iwasa, C. Hansch, A new substituent constant, IR, derived from partition coefficients, *Journal of the American Chemical Society* 86 (1964) 5175–5180. <https://doi.org/10.1021/ja01077a028>.
- [21] M. Tanaka, K. Yagyū, S. Sackett, Y. Zhang, Anti-inflammatory effects by pharmacological inhibition or knockdown of fatty acid amide hydrolase in BV2 microglial cells, *Cells* 8 (2019) 491. <https://doi.org/10.3390/cells8050491>.
- [22] Z. Anwer, S. P. Gupta, A QSAR study on a series of pyrrole derivatives acting as lymphocyte-specific kinase (Lck) inhibitors, *Medicinal Chemistry (Los Angeles)* 8 (2012) 649–655. <https://doi.org/10.2174/157340612801216319>.
- [23] A. Schneider, G. Hommel, M. Blettner, Linear regression analysis: part 14 of a series on evaluation of scientific publications, *Deutsches Arzteblatt International* 107 (2010) 776–782. <https://doi.org/10.3238/arztebl.2010.0776>.
- [24] C. Kramer, C.S. Tautermann, D.J. Livingstone, D.W. Salt, D.C. Whitley, B. Beck, T. Clark, Sharpening the toolbox of computational chemistry: A new approximation of critical F-values for multiple linear regression, *Journal of Chemical Information and Modeling* 49 (2009) 28–34. <https://doi.org/10.1021/ci800318q>.
- [25] I. v. Tetko, D.J. Livingstone, A.I. Luik, Neural Network Studies. 1. Comparison of Overfitting and Overtraining, *Journal of chemical information and computer sciences* 35 (1995) 826–833. <https://doi.org/10.1021/ci00027a006>.
- [26] D.M. Hawkins, S.C. Basak, D. Mills, Assessing model fit by cross-validation, *Journal of chemical information and computer sciences* 43 (2003) 579–586. <https://doi.org/10.1021/ci025626i>.

- [27] C.A. Lipinski, Lead- and drug-like compounds: The rule-of-five revolution, *Drug Discovery Today: Technologies* 1 (2004) 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>.
- [28] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Advanced Drug Delivery Reviews* 64 (2012) 4–17. <https://doi.org/10.1016/j.addr.2012.09.019>.
- [29] N. Tripathi, S. Paliwal, S. Sharma, K. Verma, R. Gururani, A. Tiwari, A. Verma, M. Chauhan, A. Singh, D. Kumar, A. Pant, Discovery of Novel Soluble Epoxide Hydrolase Inhibitors as Potent Vasodilators, *Scientific Reports* 8 (2018) 1–11. <https://doi.org/10.1038/s41598-018-32449-4>.
- [30] S. Kesar, S. Paliwal, P. Mishra, K. Madan, M. Chauhan, N. Chauhan, K. Verma, S. Sharma, Identification of Novel Rho-Kinase-II Inhibitors with Vasodilatory Activity, *ACS Medicinal Chemistry Letters* 11 (2020) 1694–1703. <https://doi.org/10.1021/acsmchemlett.0c00126>.
- [31] A. Ousaa, B. Elidrissi, M. Ghamali, S. Chtita, A. Aouidate, M. Bouachrine, T. Lakhlifi, QSTR analysis and combining DFT of the toxicity of heterogeneous phenols, *Journal of Materials and Environmental Science* 8 (2017) 476–484.
- [32] S. Paliwal, D. Yadav, R. Yadav, S. Paliwal, In silico structure-based drug design approach to develop novel pharmacophore model of human peroxisome proliferator activated receptor c agonists, *Medicinal Chemistry Research* 20 (2011) 656–659. <https://doi.org/10.1007/s00044-010-9370-x>.
- [33] A.M. Almerico, M. Tutone, A. Lauria, A QSAR study investigating the potential anti-HIV-1 effect of some acyclovir and ganciclovir analogs, *Arkivoc.* 2009 (2009) 85–94. <https://doi.org/10.3998/ark.5550190.0010.808>.
- [34] M.T.D. Cronin, T.W. Schultz, Development of quantitative structure-activity relationships for the toxicity of aromatic compounds to *Tetrahymena pyriformis*: Comparative assessment of the methodologies, *Chemical Research in Toxicology* 14 (2001) 1284–1295. <https://doi.org/10.1021/tx0155202>.