Journal of Computational Methods in Molecular Design, 2015, 5 (4):77-83

Scholars Research Library
(http://scholarsresearchlibrary.com/archive.html)

ISSN : 2231- 3176
CODEN (USA): JCMMDA

Investigation of some acetamido-N-benzylacetamide derivatives as potent anti-convulsant inhibitors

Usman Abdulfatai*, Adamu Uzairu and Sani Uba

Department of Chemistry, Ahmadu Bello University, Zaria-Nigeria

---

ABSTRACT

PM3 semi-empirical method was used to develop the quantitative structure-activity relationship (QSAR) for predicting the anticonvulsant activity of some acetamido-N-benzylacetamide derivative. In order to to find the optimized geometry of the studied molecules, three types of molecular descriptors were used in deriving quantitative relation between anticonvulsant activity and structural properties. The relevant molecular descriptors were selected by Genetic Function Algorithm (GFA). The optimum model has squared correlation coefficient (R)^2 value of 0.98, adjusted squared correlation coefficient (R^2_{adj}) value of 0.98, Leave one out (LOO) cross validation coefficient (Q^2) value of 0.96. The external set used for confirming the predictive power of the model has its R^2_{pred} = 0.89. These confirm the stability, robustness and predictability of the model.

Keywords: Convulsant, QSAR, GFA, Semi Empirical (PM3 method),

---

INTRODUCTION

Epilepsy is generally coursed by neurological issue and is influencing around 1% of the world's population [1]. At present the accessible anticonvulsants are valuable in diminishing the seriousness and repetitive unconstrained seizures in fewer than 70% of patients. Moreover, their treatment is connected with unfavorable symptoms. Therefore, the quest for sheltered and more intense anticonvulsant pharmaceuticals remains the need of medication configuration and the proceeded with quest for the more secure and more successful antiepileptic medications is desperately important [2, 3]. Antiepileptic medications apply their activity by distinctive systems. For example, upgrade of the GABA-ergic neurotransmission, consequences for neuronal voltage-gated sodium and/or calcium channels [4]. Past correlations of the basic attributes of anticonvulsant medication have recognized a typical example characterized by a nitrogen heteroatom framework, no less than one carbonyl gathering, together with two or one phenyl bunch [5-7]. Quantitative structure-action connections (QSAR), as a main consideration in medication outline, are scientific mathematical statements relating concoction structure to their natural movement [8]. Amid the most recent two decade, QSAR models have picked up a broad acknowledgment in sciences [9-12]. Moreover, anticonvulsant specialists have been the point of numerous QSAR studies [13-16]. The reason for this examination is to perform a quantum compound QSAR study on fifteen acetamido-N-benzylacetamide subordinates which were tentatively accepts to had anticonvulsant movement and to get a straight model by utilizing Genetic Function Approximation (GFA) system.

MATERIALS AND METHODS

Data collection

15 molecules of acetamido-N-benzylacetamide derivatives were used as anticonvulsant activity were selected from the literature and used for the present study [17]. The observed structures and the biological activities of these compounds are presented in Fig.1 and Table.1, respectively.

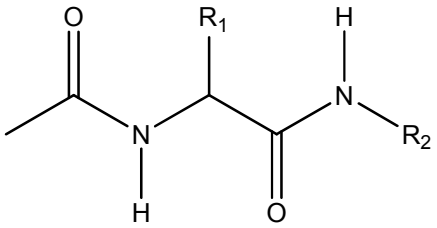

Figure 1: General structure of α_substituted acetamido-N-benzylacetamide derivatives

Table 1: Biological activities of training and test set derivatives

S.NO	R1	R2	IC_{50}	Pred.IC50
1[a]	CH_3	CH_2-Ph	1.88	1.55
2[b]	2-Furanyl	CH_2-Ph-o-F	1.60	1.10
3[b]	2-Furanyl	CH_2-Ph-m-F	1.12	1.85
4[a]	2-Furanyl	CH_2-2,5-C_6H_6	1.38	1.51
5[b]	2-Furanyl	CH_2-2,5-C_6H_6	1.80	1.75
6[b]	3-Allyl	CH_2-Ph	1.53	1.05
7[b]	2-tetrahydrofuranyl	CH_2-Ph	1.71	1.31
8[a]	Ph	CH_2-Ph	1.31	1.55
9[b]	2-Furanyl	CH_2-Ph	1.01	1.15
10[b]	2-Furanyl-5-CH_3	CH_2-Ph	1.28	1.05
11[b]	2-Pyrrolyl-5-CH_3	CH_2-Ph	1.56	1.89
12[a]	3-Thienyl	CH_2-Ph	1.94	1.52
13[b]	1-Pyrazole	CH_2-Ph	1.22	1.54
14[b]	2-Pyridyl	CH_2-Ph	1.03	1.00
15[b]	C(S)NH_2	CH_2-Ph	1.94	1.07

Training set; [b] Test set; [a]

Biological activity

The logarithm of measured IC_{50} (μM) against anticonvulsant activity as pIC_{50} (pIC_{50} = log $1/IC_{50}$) was used as dependent variable, consequently correlating the data linearly to the independent variable/descriptors.

Molecular Modeling

All molecular modeling studies were carried out using Spartan'14 version 1.1.2 [18] and PaDEL-Descriptor version 2.18 [19] running on Toshiba Satellite, Dual-core processor window eight (8) operating system. The molecular structures of the compounds in the selected series were drawn in the graphic user interface of the software. 2D application tool was used to build the structures and exported in 3D format. All 3D structures were geometrically optimized by minimizing energy. Calculation of the structural electronic and other descriptors of all the 15 acetamido-N-benzylacetamide derivatives was conducted by means of Semi- Empirical using the PM3 method. The lowest energy structure was used for each molecule to calculate their physicochemical properties (molecular descriptor).

The semi-empirically optimized structures from the Spartan'14 version 1.1.2 [18] Quantum chemistry package were saved in sdf format, and all the 1D, 2D and 3D descriptors were calculated using PaDEL-Descriptor version 2.18 tool kit [19].

Procedures

The generated descriptors (1D-3D) of the 15 data sets from the PaDEL version 2.18 tool kit [19] were divided into training and test sets. The training set was used to generate the model, while the test set were used for the external

validation of the model. The correlation between activity values of the molecules against neurotransmitter and the calculated descriptors were obtained through correlation analysis using the Material studio software version 8. Pearson's correlation matrix was used as a qualitative model, in order to select the suitable descriptors for regression analysis. The generated descriptors from the PaDEL version 2.18 tool kit [19] were subjected to regression analysis with the experimentally determined activities as the dependent variable and the selected descriptors as the independent variables using Genetic Function Approximation (GFA) method in material studio software version. The number of descriptors in the regression equation was 3, and Population and Generation were set to 500 and 500, respectively.

The number of top equations returned was 4. Mutation probability was 0.1, and the smoothing parameter was 0.5. The models were scored based on Friedman's LOF. In GFA algorithm, an individual or model was represented as one-dimensional string of bits. It was a distinctive characteristic of GFA that it could create a population of models rather than a single model. GFA algorithm, selecting the basic functions genetically, developed better models than those made using stepwise regression methods. And then, the models were estimated using the LOF, which was measured using a slight variation of the original Friedman formula, so that the best fitness score can be received.
In materials studio version 8, LOF is measured using a slight variation of the original Friedman formula [20]. The revised formula is:

$$LOF = SSE / (1 - \frac{C+dp}{M})^2 \quad (1)$$

where:

SSE is the sum of squares of errors, c is the number of terms in the model, other than the constant term, d is a user-defined smoothing parameter, p is the total number of descriptors contained in all model terms (ignoring the constant term) and M is the number of samples in the training set.

Unlike the commonly used least squares measure, the LOF measure cannot always be reduced by adding more terms to the regression model. While the new term may reduce the SSE, it also increases the values of c and p, which tends to increase the LOF score. Thus, adding a new term may reduce the SSE, but actually increases the LOF score. By limiting the tendency to simply add more terms, the LOF measure resists over fitting better than the SSE measure (Materials Studio 8.0 Manual).The significant regression is given by F-test, and the higher the value, the better the model [21] is, as showed in Table 2.

Quality Assurance of the model
The fitting ability, stability, reliability and predictive ability of the developed models were evaluated by internal and external validation parameters. The validation parameters were compared with the minimum recommended value for a generally acceptable QSAR model [22] showed in Table 2.

Table 2: Minimum recommended value of Validation Parameters for a generally acceptable QSAR model

Symbol	Name	Value
R^2	Coefficient of determination	≥ 0.6
$P_{(95\%)}$	Confidence interval at 95% confidence level	< 0.05
Q^2	Cross validation coefficient	< 0.5
R^2 - Q^2	Difference between R^2 and Q^2	≤ 0.3
$N_{ext. test set}$	Minimum number of external test set	≥ 5
R^2_{ext}	Coefficient of determination for external test set	≥ 0.6

The square of the correlation coefficient (R^2) describes the fraction of the total variation attributed to the model. The closer the value of R^2 is to 1.0, the better the regression equation explains the Y variable. R^2 is the most commonly used internal validation indicator and is expressed as follows:

$$R^2 = 1 - \frac{\sum(Yobs - Ypred)2}{\sum(Yobs - Ytraining)2} \quad (2)$$

Where, Yobs; Ypred; Ytraining are the experimental property, the predicted property and the mean experimental property of the samples in the training set, respectively [23]. Adjusted R^2 (R^2_{adj}) value varies directly with the

increase in number of repressors i.e. descriptors, thus, R^2 cannot be a useful measure for the goodness of model fitness. Therefore, R^2 is adjusted for the number of explanatory variables in the model. The adjusted R^2 is defined as:

$$R^2_{adj} = 1\text{-}\ (1 - R^2)\frac{n-1}{n-p-1} = \frac{(n-1)R^2-P}{n-p+1} \quad (3)$$

Where p = number of independent variables in the model (23).
The leave one out cross validation coefficient (Q^2) is given by;

$$Q^2 = 1 - \frac{\sum(Yp-Y)^2}{\sum(Y-Ym)^2} \quad (4)$$

Where Yp and Y represent the predicted and observed activity, respectively, of the training set and Y_m is the mean activity value of the training set [24].

RESULTS AND DISCUSSION

Four QSAR models was built using GFA algorithm, but due to the statistical significance, model 1 was selected, reported and its statistical parameters were as well calculated. The name and symbol of the descriptors used in the QSAR optimization model and Pearson's correlation matrix for descriptors used in the model were shown in the Tables 3 and 4 respectively. Likewise, Table 5 gives the result of Validation of the Genetic Function Approximation (GFA) of model 1 that was generated from material studio; also Table 6 shows the Contributions of the individual descriptors in the model

Table 3: List of some descriptors used in this studied

S/NO	Discriptors Symbol	Names of Discriptor(s)	Class
1	VC-5	Valence cluster, order 5	2D
2	saaCH	Sum of atom-type E-State: :CH:	2D
3	FNSA-3	PNSA-3 / total molecular surface area	3D

Table 4: Pearson's correlation matrix for descriptors used in QSAR model for the activities of anticonvulsant molecules

	IC50	VC-5	SaaCH
IC50	1		
VC-5	0.718241	1	
SaaCH	-0.77979	-0.55967	1
FNSA-3	-0.12934	-0.61438	0.488442

Table 5: Validation of the genetic function approximation from material studio

S/NO		Equation 1
1	Friedman LOF	0.01192900
2	R-squared	0.98255700
3	Adjusted R^2	0.97508100
4	Cross validated R-squared	0.95827600
5	Significant Regression	Yes
6	Significance-of-regression F-value	131.43249200
7	Critical SOR F-value (95%)	4.52386100
8	Lack-of-fit points	7
9	Replicate points	0
10	Min expt. error for non-significant LOF (95%)	0.03560200

Table 6: Contributions of the individual descriptors in the model

	Coefficients	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	3.000523	0.163673	3.56E-0	2.613498	3.387547
VC-5	14.94449	1.382037	1.27E-0	11.67649	18.21249
SaaCH	-0.08307	0.007443	1.03E-0	-0.10067	-0.06547
FNSA-3	22.84202	2.241243	1.89E-0	17.54233	28.14172

The result in this study revealed that model 1 was the best model which is as follows:

$$pIC_{50} = 14.94448 * VC - 5 - 0.083068355 * SaaCH + 22.842023033 * FNSA - 3 + 3.00052.$$

N = 11,R^2 =0.98, R_a = 0.975, Q^2_{cv} =0.9583 and the external validation was found to be 0.89. Model 1 gives the best QSAR model among the four models generated based on statistical significance as it has the highest R^2, R^2adj, and Q^2. Based on this analysis, Model 1 was selected and reported as the best optimized model.

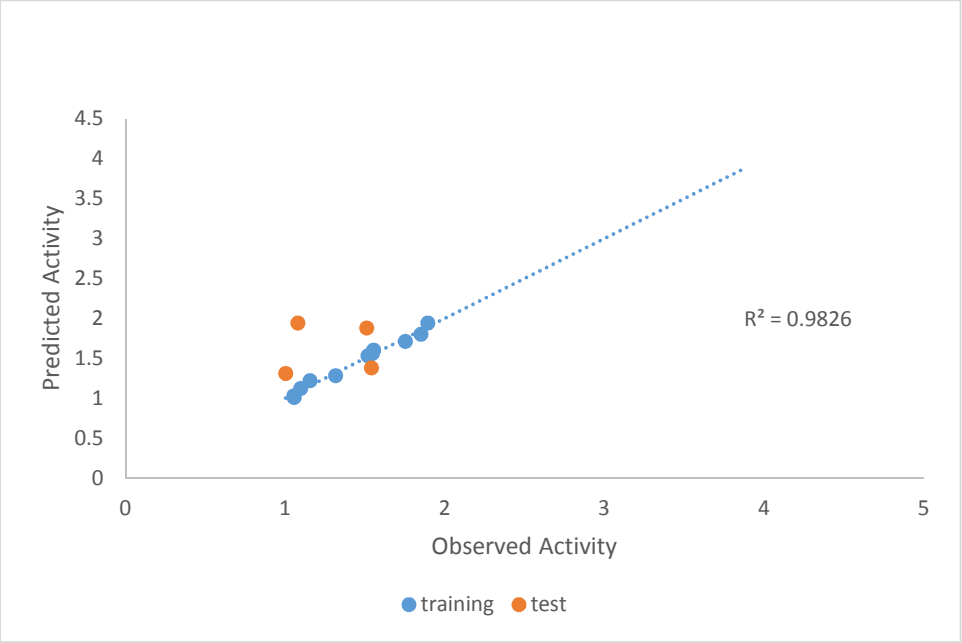

Figure 2: Training and Test Sets plot of model 1

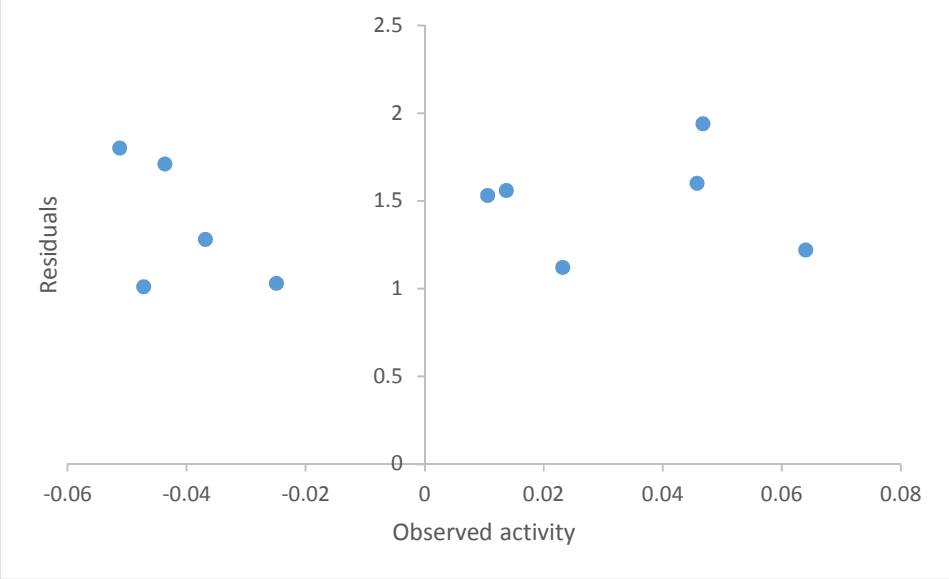

Figure 3: Residual plot of model 1

Figure 2 gives the plot of predicted activities of both training and test sets against observed activities on Microsoft excel package, the R^2 value of 0.98 recorded in this study was in agreement with GFA derived R^2 value, this further confirms the reliability of the model. Also in Fig. 3, the plot of residual activities versus observed activities indicated that there was no systemic error in model development as the spread of residuals was pragmatic on both sides of zero [24].

Table 7: Univariate Analysis of the Inhibition data

Statistical Parameters	
Number of sample points	11
Range	0.9300
Maximum	1.9400
Minimum	1.0100
Mean	1.4364
Median	1.5300
Variance	0.0936
Standard deviation	0.3209
Mean absolute deviation	0.2767
Skewness	0.0363
Kurtosis	-1.6150

A univariate analysis is performed on the inhibition efficiency data in Table 1 as a tool to assess the quality of the data available and its suitability for next statistical analysis. Data in Table 1 showed acceptable normal distribution. Statistical parameters presented in Table 4 have been discussed in details in our previous study [21].

Correlation matrix in Table 4 above, shows clearly that the correlation coefficients between each pair of descriptors is very low, thus, it can be inferred that there exist no significant inter-correlation among the descriptors used in building the model. Table 2 shows that the parameters are in agreement with the standard as $R^2 = 0.98$, $R_a = 0.98$, $Q^2_{cv} = 0.96$ and $R^2_{pred} = 0.89$. These actually confirmed the robustness of the model.

The presence of the two 2D descriptors in the model (VC-5 and saaCH *)* suggests that these types of descriptors are able to characterize better anticonvulsant activities of compounds.

Descriptors used in the generated of different QSAR models were classified in Tables 4 and 5. The contributions of each descriptor (standardized regression coefficients) in the MLR models were determined, and are provided in Table 6.The significance of the descriptors involved in each model decreases in the following order:

FNSA-3 (22.84) > VC-5 (14.94) > saaCH (-0.08)

CONCLUSION

The model with good descriptors presents a satisfactory correlation with the anticonvulsant activity, while the models with 2D and 3D descriptors are of higher excellence. The combination of 2D and 3D descriptors produce a better model to predict the anticonvulsant activity of these compounds. The QSAR results found were able to achieve a higher excellence model when compared to models obtained by other researchers. Also, this study provides a valuable approach for pharmaceutical as well as medicinal chemists to design and synthesis new anticonvulsant drugs that will be more efficient in inhibiting neurotransmitter.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this paper. Also, they declare that this paper or part of it has not been published elsewhere.

CONTRIBUTION OF THE AUTHORS

Usman Abdulfatai initiated the collaborative project, collected the data and wrote the statistical analysis plan. Adamu Uzairu monitored data collection, revised the draft paper. Sani Uba analyzed the data and revised the draft paper. All authors revised the draft paper, read and approved the final version.

REFERENCES

[1] OJ McNamara; JG Hardman; LE Limbird, AG Gilman, Eds. *The Pharmacological Basis of Therapeutics*. New York, McGraw-Hill, **2002**, p. 554.
[2] R Diaz-Arrastia; MA Agostini; PC Van Ness, *JAMA*, **2002**, 287, 2917.
[3] CW Bazil; TA Pedley, *Ann. Rev. Med.*, **1998**, 49, 135.
[4] T Anger; DJ Madge; M Mulla; D Riddall, *J. Med. Chem*., **2001**, 44, 115-137.
[5] B Hemmateenejad; R Miri; M Tabarzad; M Jafarpour; F Zand, *J. Mol. Struct. (Theochem)*, **2004**, 684, 43–49.
[6] SM Tassoa; LE Bruno-Blanch; SC Moon; GL Estiú, *J. Mol. Struct. (Theochem)*, **2000**, 504, 229–240.
[7] MD Carter; V C Stephenson; D F Weaver, *J. Mol. Struct. (Theochem)*, **2003**, 638, 57–62.
[8] P Krogsgaard-Larsen; T Liljefors; Ulf Madsen, Eds. *Textbook of Drug Design and Discovery. London*: Taylor and Francis, **2002**. p. 554.
[9] C Hansch; A Leo, *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*. American Chemical Society. Washington, DC, **1995**.
[10] H Kubinyi, *Drug Disc. Tod.*, **1997**, *2*, 457-467.
[11] O Ivanciuc, 3D QSAR Models, in: *QSPR/QSAR Studies by Molecular Descriptors*. M. V. Diudea, Ed. Nova Science. Huntington, NY. **2001**.
[12] O Ivanciuc; T Ivanciuc; D Cabrol-Bass, *Analysis*, **2000**, *28*, 637-642.
[13] S Marone; I Rozas; DF Weaver, *J. Mol. Struct. (Theochem)*, **1999**, 467, 25-30.
[14] M Marder; G Estiu; LB Blanch; H Viola; C Wasowski; JH Medina, *Bioorg. Med. Chem*., **2001**, 9, 323-335.
[15] H Verli; MG Albuquerque; R Bicca de Alencastro; EJ Barreiro, *Eur. J. Med. Chem*., **2002**, 37(Pt 3), 219-29.
[16] AY Jin; H Kohn; C Béguin; SV Andurkar; JP Stables; DF Weaver, *Can. J. Chem*., **2005**, 83, 37-45.
[17] H Kohn; JD Conley; JD Leander, Brain Res., **1998**, 457(Pt 2), 371-375.
[18] Wavefunction. Inc. Spartan'14. V. 1.1.2. Irvine, California, USA. **2013.**
[19] Yap Chun Wei, PaDEL-Descriptor. V. 2.18. **2011.**
[20] KF Khaled; *Corrosion Science*, **2011,** 53, 3457-3465.
[21] JF Friedman, *Multivariate Adaptive Regression Splines Laboratory for Computational Statistics, Department of Statistics*. Stanford University. Stanford Technical Report. No. 102, **1990.**
[22] V Ravinchandran Rajak; H Jain; A Sivadasan; S Varghese; CP Kishore-Agrawal; R. *Int. J. of Drug Design and Discov*., **2011,** 2, 511-519.
[23] Brand.on-Vaughn; KA Orr, omprehensive R archive network (CRAN): http:// CRAN.R-project.org. retrieved. **2015.**
[24] MJ Heravi; A Kyani; *J. Chem. Inf. Comput. Sci*., **2004,** 44, 1328–1335