Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2011, 3(1): 572-579
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X
CODEN (USA): PCHHAX

Quantitative Structure Activity Relationship (QSAR) Analysis on Arylbenzofuran Derivatives as Histamine H3 Antagonists

Sanmati K. Jain* and Priyanka Sinha

**SLT Institute of Pharmaceutical Sciences, Guru Ghasidas Vishwavidyalaya, Bilaspur (CG)*

ABSTRACT

Quantitative structure activity relationship (QSAR) analysis on arylbenzofuran derivatives were performed for their antihistaminic (H_3-receptor antagonist activity) using VlifeQSARPro software. Partial least square (PLS) linear regression analysis coupled with stepwise variable selection method was applied to derive QSAR models which were further validated for statistical significance by internal and external validation. Statistically significant QSAR model generated have squared correlation coefficient (r2) 0.8662, cross validated correlation coefficient (q2) 0.6029 and predictive correlation coefficient (pred_r2) 0.3940. The QSAR model indicated that the T_3_N_5 (count of number of triple bonded atoms separated from nitrogen atom by five bond in a molecule), T_C_C_7 [count of number of Carbon atoms (single or double bonded) separated from any other Carbon atom (single or double bonded) by 7 bonds in a molecule] and T_2_3_5 [count of number of double bonded atoms (i.e. any double bonded atom, T_2) separated from any other triple bonded atom by 5 bonds in a molecule] were the important determinants for H_3-receptor antagonistic activity.

Keywords: QSAR, Histamine H3 Antagonists, PLS, Arylbenzofurans

---

INTRODUCTION

The histamine H3 receptor is a G-protein-coupled receptor described earlier as central histamine modulating autoreceptors [1] and later as heteroreceptors regulating release of other neurotransmitters. Activation of histamine H3 receptor (H3R) by the endogenous ligand, histamine [1-5], reduces neurotransmitter release, while antagonism of the H3R leads to enhanced neurotransmitter release [6-7].

This enhanced neurotransmitter release is thought to be responsible for improvements in cognition, attention [8], wakefulness [9], nasal congestion [10-11], and in some cases an anti-obesity effect [12-14] upon administration of H3R antagonists. Thus, H3 receptor antagonists may be potential therapeutic agents for attention deficit/hyperactivity disorder, Alzheimer's disease, mild cognitive impairment, or schizophrenia and obesity.

The thirst for discovery of new chemical entities of therapeutic interest has been continued since for many years to medicinal chemistry experts. In recent years, a substantial progress that has been made by computational chemistry led new challenges to drug discovery by rational process. As an application of computational chemistry, nowadays, quantitative structure activity relationship (QSAR) has become more popular tool for the prediction of biological activities of molecules. The quantitative relations between the chemical properties of a molecule (physicochemical, structural and conformational) and the biological response assist to understand the driving forces for the drugs action and helps to predict the biological activities of newly designed analogues, contributing to the drug discovery processes [15].

The main objective of the present study is the search for novel arylbenzofuran derivatives that would show a promise to become useful H_3-receptor antagonist. A series of compounds of aryl benzofurans was selected as novel H_3-receptor antagonist for QSAR study [16].

MATERIALS AND METHODS

In the present study a data set of aryl benzofuran derivatives (29 molecules) as human H_3-receptor antagonists has been taken from the literature for QSAR studies (Table 1). The reported Ki values hH_3 binding affinity (μM), determined by using human histamine H_3-receptor was selected and have been converted to the logarithmic scale [pKi (moles)], for undertaking the QSAR study.

All twenty nine compounds were drawn using 2D draw application of QSARPlus [17] and converted to 3D structure. All molecules were optimized for the minimization of energies using Merck Molecular Force Field (MMFF) method until the root mean square (rms) gradient reached 0.01 kcal/mol A° before they were undertaken for 2D QSAR studies.

Number of descriptors was calculated after optimization or minimization of the energy of the data set molecules. Various types of physicochemical descriptors were calculated: Individual (H-Acceptor count, H-Donor count , X logP, SMR, polarisablity, etc.), retention index (Chi), atomic valence connectivity index (ChiV), Path count, Chi chain, Chiv chain, Chain Path Count, Cluster, Path cluster, Kapa, Element count (H, N, C, S, O, Cl, Br, I), Estate numbers (SsCH3 Count, SdCH2 Count, SssCH2 Count, StCH count etc.), Estate contribution (SsCH3-index., SdCH2- index, SssCH2 – index , StCH index) and Polar surface area.

More than 200 alignment independent descriptors were also calculated using the following attributes. A few examples are T_2_O_7, T_2_N_5, T_2_2_6, T_C_O_1, T_O_Cl_5 etc.

Structural descriptors	Selected Attributes
*Topological	2
Range	3
Min – 0	T (any)
Max. - 7	C
	N
	O
	F
	S
	Cl
	Br
	I

Generation of training and test set of compounds: In order to evaluate the QSAR model externally, data set was divided into training and test set using Random selection method, Manual data selection method and Sphere Exclusion methods. Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive effectiveness of the model which is not included in model generation.

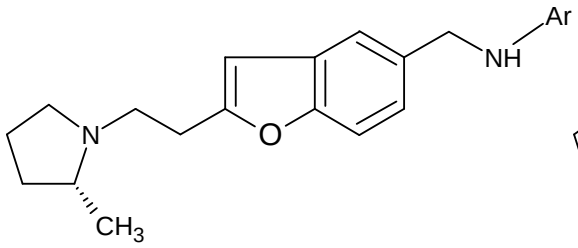
Random selection: In order to construct and validate the QSAR models, both internally and externally, the data sets were divided into training [90%-60% (90%, 85%, 80%, 75%, 70%, 65% and 60%) of total data set] and test sets [10%-40% (10%, 15%, 20%, 30%, 35% and 40%) of total data set] in a random manner. 10 trials were run in each case.

Manual data selection: Whole range of activities was sorted through ascending & descending order and every 4^{th}, 5^{th}, 6^{th},7^{th}, 8^{th}, 9^{th} and 10^{th} compound assigned to the test set.

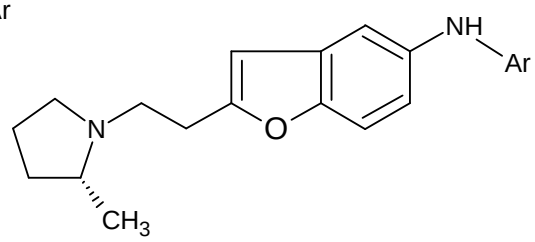
Sphere Exclusion method: In this method initially data set were divided into training and test set using sphere exclusion method. In this method dissimilarity value provides an idea to handle training and test set size. It needs to be adjusted by trial and error until a desired division of training and test set is achieved. Increase in dissimilarity value results in increase in number of molecules in the test set.

Statistical computation: Stepwise multiple linear regression (MLR), partial least square regression (PLSR) and principal component regression (PCR) were used for model generation. All the calculated descriptors were considered as independent variable and biological activity as dependent variable.

Table 01: General structure of the Aryl benzofuran derivatives and their biological activities (data set of 29 molecules)



A (1-18)



B (19-29)

S. No.	Ar substituent	human H_3 pKi
1	2-nitrophenyl	9.50
2	4-nitrophenyl	9.13
3	4-(2,2,2-trifluoroacetyl)phenyl	9.13
4	3-cyanophenyl	9.40
5	6-chloropyridazine-3-yl	9.51
6	3-cyanopyrazin-2-yl	8.98
7	Pyrazin-2-yl	9.53
8	5-bromopyrimidin-2-yl	8.91
9	Pyrimidin-5-yl	9.16
10	5-ethylpyrimidin-2-yl	9.48
11	Pyrimidin-2-yl	9.07
12	5-nitrothiazol-2-yl	9.72
13	3-nitropyridin-2-yl	9.36
14	3,5-dinitropyridin-2-yl	9.15
15	2,6-dicyanopyridin-4-yl	9.88
16	3-cyanopyridin-2-yl	9.13
17	5-cyanopyridin-2-yl	9.46
18	5-nitropyridin-2-yl	9.52
19	2-nitrophenyl	9.66
20	4-cyanophenyl	9.57
21	3-cyanophenyl	9.53
22	5-ethyl-pyrimidin-2-yl	9.21
23	Pyrimidin-2-yl	9.51
24	Pyrimidin-5-yl	10.12
25	Pyrazin-2-yl	9.98
26	5-nitropyridin-2-yl	10.32
27	3-nitropyridin-2-yl	9.87
28	3-cyano-6-methylpyridin-2-yl	8.53
29	5-trifluoromethyl-pyridin-2-yl	9.27

RESULTS AND DISCUSSION

Selected data set (arylbenzofuran derivatives) was subjected to various regression analysis methods (MLR, PLSR, PCR) for model building. The statistically significant model obtained is shown in Table-02.

Table-02: Predictive QSAR model with equation generated

Model	Method	Equation
1	Manual (Biological activity sorted in ascending manner)	pKi = 0.3427 T_3_N_5 - 0.1132 T_C_C_7 - 0.077 T_C_N_1 - 0.888 SsBrE-index - 0.1311 T_2_3_5 + 12.2046 n=20 Degree of freedom = 14 F=24.2698 r^2=0.8662 q^2=0.6029 pred_r^2=0.3940 r^2se=0.1425 q^2se=0.2455 pred_r^2se=0.3813

Test set size = 9 (Compounds no. 1,4,11,16,18,22,24,26,27).

In this equation n is the number of molecules (Training set) used to derive the QSAR model, r^2 is the squared correlation coefficient, q^2 is the cross-validated correlation coefficient, pred_r^2 is the predicted correlation coefficient for the external test set, *F* is the Fisher ratio, reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F–test indicate that the model is statistically significant. r^2 se, q^2 se and pred_r^2se are the standard errors terms for r^2, q^2 and pred_r^2. R^2 is the correlation coefficient for observed vs. predicted biological activity.

The QSAR model was obtained by using manual method (biological activity sorted in ascending manner) of training and test set data selection, where 20 of the total molecules were selected for training set while remaining 9 were selected as test set molecules.

Table- 03: Actual and predicted activity for Training set

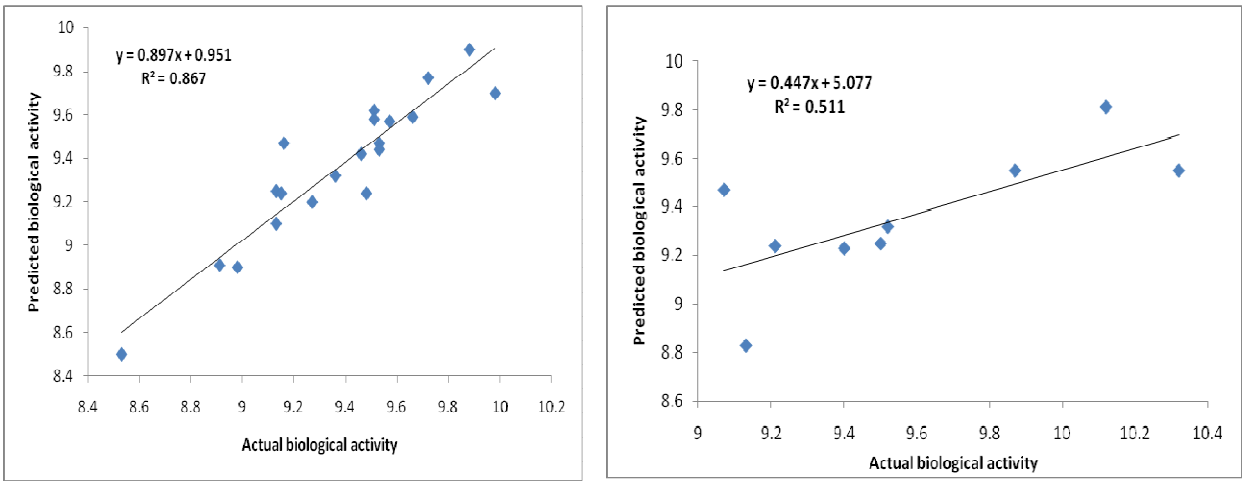
S. No.	Actual	Predicted
2	9.13	9.25242
3	9.13	9.10301
5	9.51	9.62823
6	8.98	8.90874
7	9.53	9.47427
8	8.91	8.91726
9	9.16	9.47427
10	9.48	9.24788
12	9.72	9.77764
13	9.36	9.32486
14	9.15	9.24788
15	9.88	9.90384
17	9.46	9.42331
19	9.66	9.59201
20	9.57	9.57726
21	9.53	9.44618
23	9.51	9.58747
25	9.98	9.70067
28	8.53	8.57414
29	9.27	9.28864

The model explains 86.62 % (r^2= 0.8662) of the total variance in the training set as well as it has internal (q^2) and external (pred_r^2) predictive ability of 60.29% and 39.40 % respectively. The F-

test = 24.2698 shows the statistical significance of 99.98% of the model. In addition randomization test shows confidence of 99.9% that the generated model is not random and hence it can be selected as the QSAR model. Table-03 and 04 represents the predicted biological activity by the model for training and test set respectively. The plot of observed vs predicted activity provides an idea about how well the model was trained and how well it predicts the activity of the external test set. From the plot (Figure-01) it can be seen that the model is able to predict the activity of the training set quiet well as well as external test set, providing confidence of the model.

Table- 04: Actual and predicted activity for Test set

S.No.	Actual	Predicted
1	9.5	9.25242
4	9.4	9.23768
11	9.07	9.47427
16	9.13	8.8363
18	9.52	9.32486
22	9.21	9.24788
24	10.12	9.81386
26	10.32	9.55125
27	9.87	9.55125



(Training set) (Test set)

Figure-01: Graph between actual and predicted biological activity for training set and test set

In the QSAR equation, the positive coefficient value of T_3_N_5 on the biological activity indicated that higher T_3_N_5 value leads to better antihistaminic activity (compound 12,13,15,16). The next most important descriptor influencing activity variation is T_C_C_7. The negative coefficient of this descriptor indicates that, the decrease in the count of number of Carbon atoms (single or double bonded) separated from any other Carbon atom (single or double bonded) by 7 bonds in a molecule will lead to positive effect on the activity (compound 9,17,18).

The negative coefficient of T_C_N_1 and SsBrE-index showed that higher values of these descriptors is detrimental and lower value lead to increase in the activity. Figure-02 represents the contribution chart showing contribution of the various descriptors playing important role in determining the histamine H3 receptor antagonistic activity. It reveals that the descriptors T_3_N_5, T_C_C_7 and T_2_3_5 contributing 33%, 27% and 15 % respectively. Two more descriptors T_C_N_1 (count of number of Carbon atoms (single, double or triple bonded) separated from any Nitrogen atom (single or double bonded) by 1 bond distance in a molecule) and SsBrE-index (Electrotopological state indices for number of bromine connected with one single bond) are contributing 11 % and 13% respectively.

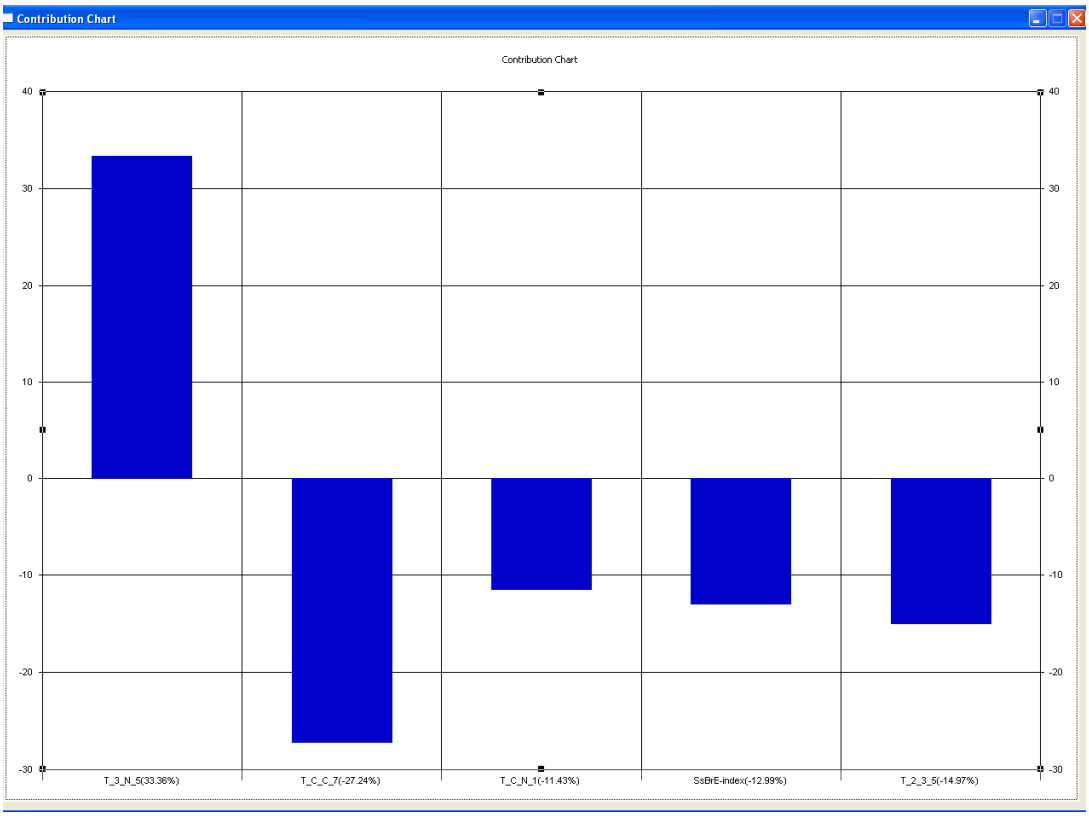


Figure-2: Contribution chart showing contribution of descriptors in biological activity

CONCLUSION

Quantitative structure activity relationship (QSAR) analysis on arylbenzofuran derivatives were performed for their antihistaminic (H_3-receptor antagonist activity) using VlifeQSARPro software. Partial least square (PLS) linear regression analysis coupled with stepwise variable selection method was applied to derive QSAR models which were further validated for statistical significance by internal and external validation. Statistically significant QSAR model generated have squared correlation coefficient (r2) 0.8662, cross validated correlation coefficient (q2) 0.6029 and predictive correlation coefficient (pred_r2) 0.3940. The QSAR model indicated that the T_3_N_5 (33%), T_C_C_7 (27%) and T_2_3_5 (15%) were the important determinants for

H_3-receptor antagonistic activity. Descriptors T_C_N_1 and SsBrE-index are contributing 11% and 13% in biological activity. Structural information obtained can be used for predicting the activity of the newer compounds with more potent activity.

Acknowlwdgement
SKJ is thankful to AICTE, New Delhi for the research project sanctioned under RPS scheme. PS is thankful to AICTE, New Delhi for the award of PG fellowship.

REFERENCES

[1] S. J. Hill, C. R. Ganellin, H. Timmerman, J.-C. Schwartz, N. P. Shankley, J. M. Young, W. Schunack, R. Levi, H. L. Haas, *Pharmacol. Rev.*, **1997**, 49, 253.
[2] J.-M. Arrang, M. Garbarg, , J.-C. Schwartz, *Nature*, **1983**, 302, 832.
[3] R. Leurs, P. Blandina, C. Tedford, H. Timmerman, *Trends Pharmacol. Sci.*, **1998**, 19, 177.
[4] T. W. Lovenberg, B. L. Roland, S. J. Wilson, X. Jiang, J. Pyati, A. Huvar, M. R. Jackson, M. G. Erlander, *Mol. Pharmacol.*, **1999**, 55, 1101.
[5] R. Leurs, R. A. Bakker, H. Timmerman, I. J. P. de Esch, *Nat. Rev. Drug Disc.*, **2005**, 4, 107.
[6] S. Morisset, A. Rouleau, X. Ligneau, F. Gbahou, J. Tardivel-Lacombe, H. Stark, W. Schunack, C. R. Ganellin, J.-C. Schwartz, J.-M. Arrang, *Nature*, **2000**, 408, 860.
[7] A. Rouleau, X. Ligneau, J. Tarvidel-Lacombe, S. Morisset, F. Gbahou, J.-C. Schwartz, J.-M. Arrang, *Br. J. Pharmacol.*, **2002**, 135, 383.
[8] A. A. Hancock, G. B. Fox, In Cognition Enhancing Drugs; Buccafusco, J. J., Ed.; Birkha¨user Verlag: Basel, Switzerland, **2004**, p 97.
[9] A. J. Barbier, C. Berridge, C. Dugovic, A. D. Laposky, S. J. Wilson, J. Boggs, L. Aluisio, B. Lord, C. Mazur, C. M. Pudiak, X. Langlois, W. Xiao, R. Apodaca, N. I. Carruthers, T. W. Lovenberg, *Br. J. Pharmacol.*, **2004**, 143, 649.
[10] R. L. McLeod, C. A. Rizzo, R. E., Jr. West, R. Aslanian, K. McCormick, M. Bryant, Y. Hsieh, W. Korfmacher, G. G. Mingo, L. Varty, S. M. Williams, N.-Y. Shih, R. W. Egan, J. A. Hey, *J. Pharmacol. Exp. Ther.*, **2003**, 305, 1037.
[11] R. Aslanian, M. W. Mutaki, N.-Y. Shih, K. D. McCormick, J. J. Piwinski, P. C. Ting, M. M. Albanese, M. Y. Berlin, X. Zhu, S. C. Wong, Y. J. Rosenblum, R. West, S. She, S. M. Williams, M. Bryant, J. A. Hey, *Bioorg. Med. Chem. Lett.*, **2002**, 12, 937.
[12] A. A. Hancock, *Curr. Opin. Investig. Drugs*, **2003**, 4, 1190.
[13] A. A. Hancock, Y. L. Bennani, E. N. Bush, T. A. Esbenshade, R. Faghih, G. B. Fox, P. Jacobson, V. Knourek-Segel, K. M. Krueger, M. E. Nuss, J. B. Pan, R. Shapiro, D. G. Witte, B. B. Yao, *Eur. J. Pharmacol.*, **2004**, 487, 183.
[14] K. Malmolof, F. Zaragoza, V. Golozoubova, H. H. F. Refsgaard, T. Cremers, K. Raun, B. S. Wulff, P. B. Johansen, B. Westerink, K. Rimvall, *Int. J. Obesity*, **2005**, 29, 1402.
[15] M.M.C.,Ferreira, *J Braz. Chem. Soc.*, **2002**, 13, 742.
[16] M. Cowart, R. Faghih, M.P. Curtis, G.A. Gfesser, Y.L. Bennani, L.A. Black, L. Pan, K.C. Marsh, J.P. Sullivan, T.A. Esbenshade, G.B. Fox, A.A. Hancock, *J. Med. Chem.*, **2005**, 48, 38-55.
[17] VLife Sciences Technology Pvt. Ltd. Pune-411045, Web: www.vlifesciences.com.