EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 4
Issue: 11
648-652
Year: 2017

QSAR AND ANTIMALARIAL DRUG DESIGN OF PHENOTHIAZINE COMPOUNDS UTILIZING MOLECULAR DESCRIPTORS

Priyanka Mukadam* and Dipika Sharma

School of Studies in Chemistry and Biochemistry, Vikram University, Ujjain (M.P.).

***Corresponding Author: Priyanka Mukadam**
School of Studies in Chemistry and Biochemistry, Vikram University, Ujjain (M.P.).

Article Received on 21/09/2017 **Article Revised on 10/10/2017** **Article Accepted on 01/11/2017**

ABSTRACT
An attempt has been made for the development of quantitative structure-activity relationship (QSAR) models for a series of phenothiazine derivatives having activity (IC_{50}). A number of highly descriptive and predictive QSAR models for these compounds were obtained by leave out method using stepwise multiple linear regression analysis. Model validation is performed by incorporating training and test sets approach and calculating CV (Coefficient of Variance), R (Correlation Coefficient), R^2A (Adjustable R^2), F (fisher's statistics).

KEYWORDS: phenothiazine derivatives, Regression analysis, Topological indices.

INTRODUCTION

Malaria is one of the most dangerous diseases in developing countries. Nearly 100 millions (10 crores) of people, all over the world, are attacked by malaria every year of which about 1% (mostly children) die. Malaria is caused by a protozoon, belonging to the sub-phylum of sporozoa and the genus plasmodium. Four species of plasmodium are clinically important:(1) P. vivax (2) P. falciparum (3) P. ovale and (4) P. malariae. Out of them, P. Vivax and p. falciparum are important. The chemotherapy of malaria has been based on drugs developed more than half a century ago. These drugs are continuously losing their efficacy, mainly due to multi-drug resistance developed by the malaria-causing parasite. Malaria parasites invade RBC→ causes hemolysis→ this is a major cause of anaemia due to malaria. Hemolytic jaundice can develop. P. falciparum induced hemolysis is particularly more severe. In P. falciparum malaria, RBCs containing schizoints cling to the walls of capillaries of vital organs like brain/kidney. This can cause hypoxia of the organ concerned. Further, rupture of these schizoints containing RBCs cause further hypoxia leading to damage of the organ→ death may result. In the last three decades, artemisinin and artemisinin-like compounds have proven to be efficient alternatives to the chemotherapeutic control of malaria.[1-2] These facts have led to an increasing interest in the development of Quantitative Structure Activity Relationship (QSAR) models for these compounds.

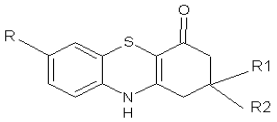
New drugs with still better relief giving potential are synthesized employing the application of mathematical concept called graph theory and topology.[3] All such new compounds are possible to prepare employing combinatory and throughout chemistry.[4] Such new synthesis is possible by the correlation analysis in that correlation is investigated by correlating the activity with topological indices. Needless to state that topological index is a numerical representation of molecular structure. Hence, quantitative structure-activity relationship (QSAR) is only possible using topological indices as molecular descriptors.[5] In the present study we have attempted QSAR study on the modeling of anti-malarial activity of phenothiazine derivatives using stepwise regression analysis

MATERIALS AND METHODS

Biological activity data

A number of 16-phenonthizine compounds having promising antimalarial activities have been considered in the present study. Table 1 contains structural substituent along with biological activities of 16 compounds. In view of this we have for the first-time used the new topological indices developed by Balaban-Khadikar-Sufia.[6-7] These indices (F and G) are the 'Shape' indices and therefore, we can use them for investigating the influence of 'Shape' on anti-malarial activity. In addition, we have also used the well- known Balaban J index.[8] This J index is supposed to be most discrimindly index known and developed in chemical graph theory.

Structural details, activity (IC_{50}) of the phenothiazine used in the present study



Sr.No.	R	R_1	R_2	Log(IC_{50})
1	Cl	H	H	1.602
2	Cl	CH_3	CH_3	1.4771
3	Cl	H	C_6H_5	1
4	Cl	H	3-$CH_3OC_6H_4$	1.301
5	Cl	H	4-$CH_3OC_6H_4$	1
6	Cl	H	2,3-$(CH_3O)2C_6H_3$	1
7	Cl	H	3,4-$(CH_3O)2C_6H_3$	1.4771
8	Cl	H	4-ClC_6H_4	0.602
9	F	H	2,4-ClC_6H_4	1
10	F	CH_3	CH_3	1.7781
11	F	H	C_6H_5	1.301
12	F	H	3-$CH_3OC_6H_4$	1.301
13	F	H	2,3-$(CH_3O)2C_6H_3$	1.301
14	F	H	2,4-Cl_2C6H3	1
15	F	H	4-ClC_6H_4	0.6989
16	F	H	3,4-$(CH_3O)_2C_6H_3$	1.301

Structure Optimization

All the structures of 16-phenonthizine compounds were drawn using Graph theory or topological approaches. Topology is a combination of two words “topos” and “logy” topos means place &logos means study. So Topology is a diagrammatic representation of molecular graph. The graph so obtained is called carbon-hydrogen suppressed molecular graph. Numbers reflecting certain structural features of organic molecules that are obtained from the respective molecular graphs, are usually called "topological indices". Such a number of usually obtained by imposing certain conditions on vertices (atoms), edges (bonds) or both. In this paper we have calculated topological indices/molecular descriptors by the using of distance matrix.[9-13]

Calculation of theoretical molecular descriptor

Balaban-Khadikar-Sufia has introduced two new Balaban type indices viz F and G using the expression of Balaban J index.

$J = (E/R+1) \sum (di.dj)^{-1/2}$

all edgs

Where E is the number of edges, di and dj are the distance sums and R is cyclomatic number defined by the following expression

$R = E-N+1$

Where N is the number vertices (atoms) in the molecular graph.

The two new indices F and G are defined by the following expression

$F = E \sum (di.dj)^{-1/2}$

$= J (R+1)$

and $G = [N^2E/N+R+1] \sum (di.dj)^{-1/2}$

$= [N^2F/R+R+1]$

The value of J, F, G along with the value of log IC_{50} are presented in Table 2.

Statistical Methodology

Stepwise-Multiple Linear Regression Analysis: Quantitative structure–activity relationships are regression models having significant role in the biochemical sciences and engineering. QSAR regression models relate a set of predictor variables (X) calculated from the chemical structures to the potency of the response variable (Y) which is biological activity. Simple linear regression is the method of choice when the research question is to predict the value of a response (dependent) variable, denoted Y, from an explanatory (independent) variable X. The regression model is

$Y=a+bX$

The extension of simple regression to two or more independent variables is straightforward. For example, if four independent variables are being studied, the multiple regression models is.

$Y = a +b_1X_1+b_2X_2+b_3X_3+b_4X_4$

where X_1 is the first independent variable and b_1 is the regression coefficient associated with it, X_2 is the second independent variable and b_2 is the regression coefficient associated with it, and so on. This arithmetic equation is called a linear combination; thus, the response variable Y can be expressed as a (linear) combination of the explanatory variables.[14-17]

After variable selection, multiple linear regression (MLR) method has been used to derive a number of training QSAR models, but in this paper we have made

the model by stepwise regression analysis. The statistical parameters obtained are very useful to investigate the participation of each of the descriptors for modeling the activity. The quality of each model is denoted by R^2 (R is the square root of multiple R-square for regression), CV (Coefficient of Variance), R (Correlation Coefficient), R^2A (Adjustable R^2), F (fisher's statistics). For a QSAR model, the value of R^2 should be more than 0.5.[18-20] Finally, the proposed QSAR models will be cross validated by leave-on-out procedure.

RESULTS AND DISCUSSION

The results obtained in the present study for modeling antimalarial compounds on the basis of topological indices J,R,F,G. The value of J, R, F,G has given in table 2.

Table 2.

Sr. No.	J	R	F	G	Log IC_{50}
1	1.705	3	6.82	87.296	1.602
2	1.73	3	6.92	101.913	1.4771
3	1.422	4	7.11	127.453	1
4	1.407	4	7.035	139.73	1.301
5	1.377	4	6.885	136.75	1
6	1.451	4	7.255	186.557	1
7	1.396	4	6.975	179.357	1.4771
8	1.402	4	7.01	132.439	0.602
9	1.426	4	7.13	141.617	1
10	1.73	4	6.92	101.913	1.7781
11	1.422	4	7.11	127.453	1.301
12	1.407	4	7.035	139.73	1.301
13	1.451	4	7.255	186.557	1.301
14	1.426	4	7.13	155.48	1
15	1.402	4	7.01	132.439	0.6989
16	1.396	4	6.975	179.357	1.301

Application of rule of thumb for the set of 16-phenonthizine compounds used in the present investigation has indicated that we can use at the most four parameters (topological indices) for modeling the anti-malarial activity. By the using of leave out method the data presented in table 2 can be divided into two different categories. The category 1 contains 16 compounds and has four regression models. While category 2 contain 14 or 15 compounds and has five Regression models. The category can be divided by the minimum or maximum of residuals and absolute error. The most appropriate model for modelling the anti-malarial activity in category 1 and 2 has given in table 3.

In category 1 the model does not contain any shape index indicating that their anti-malarial activity is independent of the shape. If explains only 59% variance in the anti malarial activity indicated by the Ridge parameter. This model is free from the defect due to co-linearity. In category 2 the model also contains no shape parameter and therefore the anti malarial activity of compounds belonging to this category is independent of shape. It accounts for 57% variation in the activity and is free from the defect due to colinearity.

It is normally recommended that for obtaining appropriate model N should be as large as possible. The coefficient of Variance, CV should be as small as possible. R^2 is actually square of correlation coefficient, in obtaining the best model R^2 should be as near to 1 as possible. When the value of R^2 is multiplied by100 it gives % variance in the activity, by which model can be perform, if R^2 = 0.998, then the models indicate 99.8% variance in the activity. The ratio of R/Se gives quality factor Q and decides the quality of regression. If the value of Q is higher, then the quality of the model will be best. Obviously one should have R as high as possible and Se as low as possible. This factor accounts for the predictive power of the model.

In order to confirm the results obtained through model discussed above, we have estimated anti-malarial activity and correlated with the observed i.e. experimental activity. Predicted activities for the test compounds generated by this two QSAR models respectively are given in Table 4.

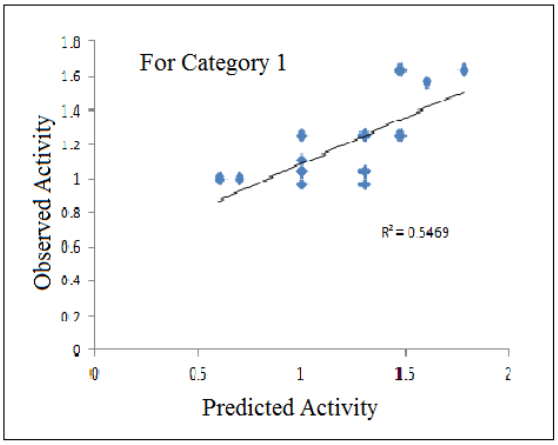
Table 3.

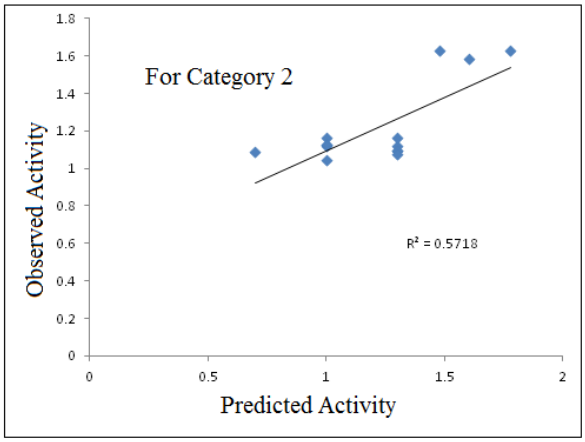
Category	No. Of Compounds	Model Equation	Statistical parameters related to quality of the model				
			N	CV	R^2	R^2A	F
1	16	Anti-malarial activity {Log (IC_{50})} =2.6269+0.8486(±0.2364) NSK-0.8130(±0.2210) NBO	16	0.1917	20.5469	0.4772	7.846
2	14	Anti malarial activity {Log (IC_{50})}= -1.2297+1.6517(±0.4127) J	14	0.1603	0.5718	0.5361	10.772

Table 4.

Category 1 (For 16 compounds)			Category (For 14 compounds)		
Test Compound	Observed Activity	Predicted activity (Using model 1)	Test Compound	Observed Activity	Predicted activity (Using model 2)
1	1.602	1.571522	1	1.602	1.586517
2	1.4771	1.642839	2	1.4771	1.62781
3	1	0.972515	3	1	1.119074
4	1.301	1.043833	4	1.301	1.094297
5	1	1.043833	5	1	1.044745
6	1	1.257785	6	1	1.166974
7	1.4771	1.257785	9	1	1.125681
8	0.602	1.008174	10	1.7781	1.62781
9	1	1.043833	11	1.301	1.119074
10	1.7781	1.642839	12	1.301	1.094297
11	1.301	0.972515	13	1.301	1.166974
12	1.301	1.043833	14	1	1.125681
13	1.301	1.257785	15	0.6989	1.086039
14	1	1.11515	16	1.301	1.076128
15	0.6989	1.008174			
16	1.301	1.257785			

Such a correlation is shown in following fig - Indicating that both observed and estimated activity are close to each other.





CONCLUSION

The results obtained here in indicate that the newly introduced F index in combination with J index can be successfully used for modelling anti-malarial activity of phenonthizine used. This is to convey that there is no specific theoretical modeling for these compounds so far as done, therefore, studies in this direction for exploration of essential structural features of the phenonthizine congeners under the frame work of molecular descriptors may help to design more potent and active phenonthizine derivative which will be developed for the use of different malarials.

ACKNOWLEDGEMENT

The authors are thankful to Late Dr. P. V. Khadikar for providing necessary guidance.

REFERENCES

1. Rabinowitz JL, Myerson RM, (Topics in Medicinal Chemistry) Intersicne Pub, Newyork, 1967.
2. Agrawal VK, Sharma R, Khadikar PV, Bioorg. Med. Chem, 2002; 10: 1361.
3. Trinjastic N, Chemical Graph Theory, CRC, Press, Newyork, 1993.
4. Miertus S, Fassina G, Combinational Chemistry and Techonolgy, Principles, Methods and Application, Marcel Dekker, New York, 2000.
5. Allison PD, Multiple Regressions. A Primer, Pine Forge Press, London, 1998.
6. Gutman I, Graph Theory Notes, New York, 1994; 27: 9-115.
7. Balaban AJ, Khadikar PV, Aziz S. Iranian J. Math. Chem, 2010; 1: 43-60.
8. Balaban AT, Chem. phys. Lett, 1982; 80: 399-404.
9. Marshall C.W, *Applied graph theory*; Wiley Interscience N.Y., 1971.
10. Deo N, *Graph theory with applications to engineering and computer science*; Prentice Hall, New York, 1974.
11. Tang AC, Kiang YS, Yan GS, *Graph theoretical molecular orbitals*; Science Press, Beijing, 1986.
12. Randic M, On characterization of three-dimensional structures; Int. *J. Quant. Chem*, 1998**;** 15: 201.
13. Randic M, On computation of optimal parameters for multivariate analysis of structure- property-relationship, *J. Comput. Chem*, 1991; 12: 970.
14. Todeschini R, Consonni V, Molecular Descriptors for Chemoinformatics, Revised and Enlarged Edition. 2ndedition.Weinheim: Wiley-VCH, 2009.
15. Katritzky AR, Petrukhin R, Tatham D, Basak S, Benfenati E, Karelson M, Maran U, J. Chem. Inf.Comput. Sci, 2001; 41: 679–685.
16. Tropsha A, Gramatica P, Gombar VJ, QSAR Comb. Sci, 2003; 22: 69–77.
17. Draper NR, Smith H, 3rd Ed., John Wiley & Sons, New York, 1998.
18. Roy PP, Roy K, Eur. J. Med. Chem, 2009; 44: 2913-2922.
19. Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN, J. Comput. Chem, 2013; 34: 1071-1082.
20. Chaterjee A, Hadi AS, Price B.; *Regression Analysis by Examples*, Wilely, 2000.