Morocean Journal of Chemistry

Modeling of some cyclic peroxy ketals for their Antimalarial Activities

Shweta Sharma[(a)], K. Anita [(a)], Vijay K. Agrawal[(b)] and Sarita Shrivastav[(c)]*

[(a)]Laboratory of Chemistry, Career College, Barkatullah University , Bhopal, India
[(b)] Laboratory of Chemistry, A.P.S. University, Rewa, India
[(c)] Laboratory of Chemistry, Govt. M.V.M., Barkatullah University , Bhopal, India

**Corresponding author. E-mail : E-mail : shwetasharma2703@yahoo.com*

Received 31 Dec 2014, Revised 02 Fev 2015, Accepted 08 Mars 2015

Abstract

In this work a set of some cyclic peroxy ketals were tested for their antimalarial activities. Quantitative structure activity relationship (QSAR) analysis was applied to 20 organic compounds of the above mentioned derivatives using Physicochemical, informational and 2D-autocorelation parameters and modeled their antimalarial activity ($logIC_{50}$) values. The multiple regression analysis clearly indicates that ^{5}BIC ,^{1}IC, MATS4v and ST parameters yielded the best model having R^2 value of 0.9515. The predictive powers of the models were explained using LOO (Leave-One-Out) Cross validation procedure. The results are also discussed on the basis of ridge regression.

Keywords: *QSAR, Physicochemical, MLR, Ridge regression, 2D-autocorelation, LOO.*

1. Introduction

Malaria is a very serious infectious disease which is caused by protozoans of the genus *Plasmodium* and is transmitted through the bite of infected female *Anopheles* mosquitoes. Every year, over one million people die from malaria, especially in tropical and subtropical areas. Most of the deaths are attributed to the parasite species *Plasmodium falciparum*. Many drugs have been investigated for their efficacy in the treatment of the disease, but strains of *P. falciparum* resistant to some of these drugs have appeared. Hence, the discovery of new classes of more potent compounds to treat the disease is necessary [1–6]. In the evolution of computational chemistry, the use of molecular modeling (MM) has been one of the most important advances in the design and discovery of new drugs. Currently, MM is an indispensable tool in not only the process of drug discovery but also the optimization of existing prototypes and the rational design of drug candidates [7–10]. According to IUPAC, MM is the investigation of molecular structures and properties using computational chemistry and graphical visualization techniques to provide a three-dimensional representation of the molecule under a given set of circumstances [8]. QSAR studies use chemometric methods to describe how a given biological activity or a physicochemical property varies as a function of the molecular descriptors describing the chemical structure of the molecule. Thus, it is possible to replace costly biological tests or experiments using a given physicochemical property (especially those involving hazardous and toxically risky materials or unstable compounds) with calculated descriptors that can, in turn, be used to predict the responses of interest for new compounds [11]. In this study an attempt has been made to model antimalarial activity (log IC_{50}) of a set of 20 cyclic peroxy ketals derivative reported by Posner et

al. [12] by using few Physicochemical, informational and 2D-autocorelation descriptors which are simple to calculate but very effective in predicting biological activity. The general structure of cyclic peroxy ketals derivatives is shown in figure 1.

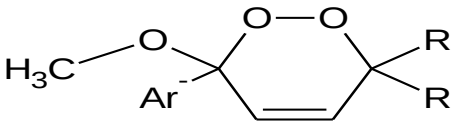

Fig 1. General Structure of peroxy ketals

2. Materials and methods

2.1 Computational chemistry

Quantitative Structure Activity Relationship (QSAR) modeling establishes a quantitative correlation between chemical structure and biological activity. The methodology used in the present study is to model the anti-malarial activities of cyclic peroxy ketals using physicochemical and informational indices. Table 1 records structural details of 20 cyclic peroxy ketals derivatives. The biological activity is also recorded in table 1. Biological activitiy values of 20 peroxy ketals are expressed as logarithm of IC_{50} (50% inhibitory concentration, in nM units). From the large pool of Physicochemical, Informational and 2D-autocorelation descriptors we have selected a few to carryout multiple regression analysis. This selection of descriptors was done by using variable selection for multiple regression analysis available with the NCSS software[13]. The calculated values of such descriptors are presented in Table 2. The intercorrelatedness among the descriptors and their correlation with the activity values logIC_{50} is presented in Table 3. The regression parameters as well as the quality of different models containing one to several correlating parameters are summarized in Table 4. Using the best four-parametric model, we have estimated and compared the values of activity. Such a comparison is demonstrated in Table 5. Finally, all the proposed models are validated by cross-validation method (Table 6)[14]. The presence/absence of co- linearity, if any, was examined by Ridge regression parameters. (Table 7, Figs. 3 and 4).

2.2 Molecular parameters used

DRAGON software [15] has been used for calculation of all Physicochemical, Informational and 2D-autocorelation descriptors. In fact before this study, topological parameters have been very successfully used by our research group in modeling different activities of drug molecules[16-19] . The details of parameters which are used in present study-

(i) Informational - theoretic topological indices

Information - theoretic topological indices are calculated by the application of information-theoretic concepts on chemical graphs [20-23, 25-26]. An appropriate set A of n elements is derived from a graph G depending upon certain structural characteristics. On the basis of an equivalence relation defined on A, the set A is partitioned into disjoint subsets A_i of order n_i (i= 1, 2,……h; $\sum_i n_i = n$). A probability distribution is then assigned to the set of equivalence classes: A_1, A_2, ……..A_h and p_1, p_2,…….. P_h

Where $pi = n_i/n$ is the probability that a randomly selected element of A will occur in the i^{th} subset.

The mean information content of an element of A is defined by Shannon's relation [24].

$$IC = -\sum_{i=1}^{k} P_i \log p_i \qquad (1)$$

The logarithm is taken at base 2 for measuring the information content in bits. The total information content of the set A is then n times IC.

Mean informationa content index

$$^{k}IC = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n} \qquad (2)$$

n_i - number of atoms in the i^{th} class

n - The total number of atoms in the molecule

k - Number of atomic layers in the coordination sphere around a given atom that are accounted for

Bonding information content index (^{k}BIC)

$$^{k}BIC = \frac{^{k}IC}{\log_2 q} \quad \text{where} \quad ^{k}IC = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n} \qquad (3)$$

n_i - number of atoms in the i^{th} class

n - The total number of atoms in the molecule

k - Number of atomic layers in the coordination sphere around a given atom that are accounted for

q - Number of edges in the molecular graph

(ii) 2D-autocorelation descriptors

Another interesting set of molecular descriptors implemented in DRAGON , and widely used in molecular modeling , are 2D-autocorelation [27-28]. These descriptors have their origin in the autocorrelation of the topological structure calculation of Broto - Moreau (ATS) , of Moran (MATS) , and of Geary(GATS). The computational of these descriptors involves summing different autocorrelation functions corresponding to the different fragment lengths , thereby leading to different autocorrelation vectors according to the lengths of the structural fragments.

Moran's Indices:

$$MATS_w\ l = \frac{N}{2L} \frac{\sum_{ij} \partial_{ij} (w_i - \bar{w})(w_j - \bar{w})}{\sum_i (w_i - \bar{w})} \qquad (4)$$

where *ATS l w* , *MATS l w* , and *GATS l w* are Broto- Moreau's autocorrelation coefficient, Moran's index, and Geary's coefficient at spatial lag *l*, respectively; where *i w* and *j w* are the values of any atomic property of atom *i* and *j* respectively; *w* is the average value of property; *L* is the number of nonzero values in the sum, *N* is the number of atoms in the molecule, and $\delta\ (l, d_{ij})$ is a Dirac-delta function defined as

$$\partial(l, d_{ij}) = \begin{cases} 1 \text{ if } d_{ij} = l \\ 0 \text{ if } d_{ij} \neq l \end{cases} \qquad (5)$$

where *dij* is the topological distance or spatial lag between atoms *i* and *j*.

(iii) Surface Tension (γ) – Surface tension is the Physicochemical parameter which is calculated by the following formula.

$$\gamma = \left(\frac{P_r}{MV}\right)^4 \qquad (6)$$

ChemSketch calculates the surface tension from calculated Molar Volume and calculated Parachor [29].

3. Results and Discussions

The data presented in Table 4 indicates that statistically allowed model start pouring using two or more parameters as correlating descriptors. We observed that in all these higher parametric models ST is

invariably present as one of the correlating descriptors. By examination of Table 4 we also observed that both R^2 and R^2_A go on increasing with each addition of descriptor in the regression analysis. This indicates that addition of descriptor in each case is favorable for the exhibition of the activity.

One-variable model

$\log IC_{50}$ = -0.0824(±0.0218) ST +5.7173

N=20, R^2 =0.4435, R^2 A= 0.4126, Se= 0.01248, F= 14.345, Q= 5.3362 (7)

Here, and here after N is the number of compound , Se is the standard error of estimation, R2 is the square of correlation coefficient , R2 Adj is the adjusted R2 , F is the Fisher's ratio, and Q is the Pogliani's quality factor which is the ratio of R/ Se.(Pogliani,1994,1996)[30-31]

Two -variable model

$\log IC_{50}$ = 7.6099(±2.3238) ^{5}BIC -0.0729(±0.0178) ST -0.1211

N=20, R^2 =0.6588, R^2 A= 0.6186, Se= 0.1006, F= 16.409, Q= 5.1907 (8)

Three variable model

$\log IC_{50}$ = 12.1159.(±1.2894) ^{5}BIC -1.7379(±0.2344) ^{1}IC -0.0562(±0.0090) ST -0.2701

N=20, R^2 = 0.9231, R^2 A= 0.9087, Se= 0.0492, F=64.001, Q= 19.5281 (9)

Four Variable model

$\log IC_{50}$=10.3058(±0.1.2209)^{5}BIC-1.7430(±0.1922)^{1}IC+1.1180(±0.3770)MATS4v-0.0551(±0.0074)ST +1.0592

N=20, R^2 =0.9515, R^2 A= 0.9386, Se= 0.0404, F= 73.586, Q= 24.1448 (10)

Table-1 Structural details of the compounds with their experimental activity log IC_{50} values.

Comp.No.	Ar	R,R	log IC_{50}
1	Ph	Me,Me	3.041
2	Ph	Cyclopentyl	2.279
3	Ph	Cyclohexyl	2.447
4	Ph	Cycloheptyl	2.342
5	4-MeOPh	Cyclobutyl	2.204
6	4-MeOPh	Cyclohexyl	2.255
7	4-MeOPh	Cycloheptyl	2.322
8	3,4,5-$(MeO)_3$Ph	Cycloheptyl	2.079
9	4-CF_3OPh	Cycloheptyl	1.785
10	4-ClPh	Cycloheptyl	1.763
11	4-FPh	Cycloheptyl	1.929
12	4-MeSPh	Cycloheptyl	1.892
13	4-$MeS(O)_2$Ph	Cycloheptyl	1.491
14	4-EtPh	Cycloheptyl	2.255
15	4-MeSPh	Cyclohexyl	2.204
16	4-$MeS(O)_2$Ph	Cyclohexyl	1.748
17	4-O_2NPh	Cyclohexyl	1.663
18	4-ClPh	Cyclohexyl	2.000
19	4-FPh	Cyclohexyl	2.301
20	4-F_3CPh	Cyclohexyl	2.146

Table 2. Calculated Values of Physicochemical, Informational and 2D-autocorelation Parameters.

S.No.	ST	D	^{0}IC	^{0}BIC	^{1}IC	^{5}BIC	MATS1v	MATS2v	MATS3v	MATS4v
1	38.300	1.100	1.348	0.259	2.261	0.778	-0.050	0.085	-0.190	0.043
2	43.700	1.120	1.247	0.212	2.108	0.705	0.022	-0.060	-0.069	0.068
3	43.100	1.090	1.224	0.203	2.035	0.713	0.021	-0.076	-0.019	-0.018
4	42.600	1.070	1.204	0.196	1.971	0.689	0.020	-0.089	0.035	-0.112
5	45.100	1.170	1.311	0.225	2.289	0.750	-0.010	-0.027	0.030	-0.088
6	43.600	1.110	1.260	0.206	2.124	0.723	-0.008	-0.018	-0.014	-0.045
7	43.100	1.090	1.240	0.199	2.055	0.700	-0.007	-0.036	0.039	-0.131
8	44.100	1.120	1.293	0.203	2.100	0.701	-0.054	0.055	0.054	-0.139
9	40.600	1.180	1.463	0.235	2.354	0.700	-0.007	-0.058	0.042	-0.125
10	44.300	1.130	1.307	0.212	2.138	0.689	0.018	-0.070	0.034	-0.118
11	41.800	1.110	1.307	0.212	2.138	0.689	0.020	-0.087	0.035	-0.114
12	45.300	1.110	1.288	0.207	2.131	0.700	0.022	-0.052	0.037	-0.134
13	47.800	1.170	1.369	0.217	2.257	0.701	0.021	-0.055	0.038	-0.134
14	42.100	1.050	1.186	0.189	2.019	0.706	0.019	-0.057	0.044	-0.133
15	46.000	1.130	1.313	0.215	2.204	0.723	0.023	-0.036	-0.014	-0.051
16	48.700	1.200	1.397	0.226	2.335	0.723	0.022	-0.039	-0.012	-0.051
17	49.100	1.190	1.420	0.233	2.309	0.716	0.028	-0.065	-0.024	-0.023
18	45.000	1.160	1.334	0.222	2.218	0.713	0.019	-0.056	-0.018	-0.025
19	42.200	1.130	1.334	0.222	2.218	0.713	0.021	-0.073	-0.019	-0.019
20	39.900	1.190	1.453	0.239	2.359	0.717	0.020	-0.046	-0.035	-0.029

As the data set contains only 20 compounds no higher parametric correlation is permitted. Therefore, the four-parametric model obtained above is the best model for estimating $\log IC_{50}$ activity of proposed set of compounds.

Table 3. Correlation matrix

	$\log IC_{50}$	ST	D	^{0}IC	^{0}BIC	^{1}IC	^{5}BIC	MATS1v	MATS2v	MATS3v	MATS4v
$\log IC_{50}$	1.000										
ST	-0.666	1.000									
D	-0.586	0.427	1.000								
^{0}IC	-0.435	0.101	0.886	1.000							
^{0}BIC	0.092	-0.177	0.635	0.805	1.000						
^{1}IC	-0.333	0.151	0.904	0.944	0.851	1.000					
^{5}BIC	0.567	-0.163	0.163	0.211	0.655	0.435	1.000				
MATS1v	-0.457	0.401	0.145	-0.003	-0.239	-0.022	-0.460	1.000			
MATS2v	0.470	-0.236	-0.039	0.092	0.381	0.163	0.657	-0.892	1.000		
MATS3v	-0.617	0.314	-0.040	-0.183	-0.665	-0.279	-0.742	0.232	-0.471	1.000	
MATS4v	0.477	-0.132	0.147	0.130	0.506	0.241	0.560	0.043	0.170	-0.875	1.000

Table 4. Regression parameters and quality of correlation

Model N°	Parameters	A_i=(1……6)	B	Se	R^2	$R^{2\,A}$	F	Q=R/Se
1.	d	-4.7456(±1.5450)	7.4746	0.1355	0.3439	0.3074	9.435	4.3279
2.	ST	-0.0824(±0.0218)	5.7173	0.1248	0.4435	0.4126	14.345	5.3362
3.	^{0}IC	-1.9183(±0.9349)	4.6297	0.1507	0.1896	0.1445	4.210	2.8894
4.	^{1}IC	-0.9804(±0.6551)	4.2457	0.1578	0.1107	0.0612	2.240	2.1085
5.	^{0}BIC	1.8874(±4.8415)	1.6985	0.1666	0.0084	0.000	0.152	0.5501
6.	^{5}BIC	9.1679(±3.1427)	-4.4244	0.1379	0.3210	0.2833	8.510	4.1085
7.	MATS1v	-6.6065(±3.0270)	2.1602	0.1488	0.2093	0.1653	4.763	3.0746
8.	MATS2v	3.7420(±1.6553)	2.2682	0.1477	0.2211	0.1779	5.111	3.1836
9.	MATS3v	-3.7821(±1.1356)	2.1024	0.1316	0.3813	0.3469	11.093	4.6922
10.	MATS4v	2.6173(±1.1367)	2.2876	0.1471	0.2275	0.1846	5.302	3.2425
11.	d ST	-2.9882(±1.4483) -0.0629(±0.0221)	8.2415	0.1149	0.5549	0.5026	10.599	6.4832
12.	^{0}IC ST	-1.6384(±0.6958) -0.0777(±0.0195)	7.6678	0.1115	0.5803	0.5310	11.755	6.8321
13.	^{1}IC ST	-0.7000(±0.5120) -0.0779(±0.0215)	7.0496	0.1219	0.4986	0.4396	8.453	5.7926
14.	^{0}BIC ST	-0.5595(±3.7894) -0.0830(±0.0227)	5.8645	0.1284	0.4442	0.3788	6.794	5.1907
15.	^{5}BIC ST	7.6099(±2.3238) -0.0729(±0.0178)	-0.1211	0.1006	0.6588	0.6186	16.409	8.0682
16.	MATS1v ST	-3.2775(±2.7392) -0.0711(±0.0235)	5.2503	0.1234	0.4867	0.4263	8.060	5.6535
17.	MATS2v ST	2.6388(±1.3362) -0.0727(±0.0208)	5.4067	0.1159	0.5473	0.4941	10.278	6.3831
18.	MATS3v ST	-2.7744(±0.9538) -0.0648(±0.0193)	4.9422	0.1050	0.6284	0.5847	14.375	7.5497
19.	MATS4v ST	2.1716(±0.8519) -0.0759(±0.0192)	5.5829	0.1093	0.5974	0.5500	12.613	7.0715
20.	^{5}BIC d ST	9.7041(±1.4650) -4.3724(±0.7995) -0.0417(±0.0122)	1.9658	0.0612	0.8811	0.8588	39.510	15.3377
21.	^{5}BIC ^{0}IC ST	9.5290(±1.3348) -2.2349(±0.3604) -0.0641(±0.0100)	1.0673	0.0562	0.8997	0.8809	47.857	16.8777
22.	^{5}BIC ^{1}IC ST	12.1159(±1.2894) -1.7379(±0.2344) -0.0562(±0.0090)	-0.2701	0.0492	0.9231	0.9087	64.001	19.5281
23.	^{0}BIC ^{5}BIC ST	-11.7084(±2.7353) 13.5225(±2.1407) -0.0779(±0.0126)	-1.5770	0.0708	0.8409	0.8111	28.194	12.9521
24.	^{5}BIC MATS1v ST	7.5625(±2.6623) -0.1043(±2.5588) -0.0726(±0.0197)	-0.0996	0.1037	0.6588	0.5948	10.298	7.827
25.	^{5}BIC MATS2v ST	7.1593(±3.1314) 0.3208(±1.5633) -0.0722(±0.0186)	0.1593	0.1035	0.6597	0.5958	10.337	7.8475
26.	^{5}BIC MATS3v ST	5.2797(±3.4566) -1.2431(±1.3597) -0.0679(±0.0187)	1.3194	0.1011	0.6757	0.6149	11.112	8.1307
27.	^{5}BIC MATS4v ST	5.8364(±2.7536) 1.0875(±0.9293) -0.0718(±0.0176)	1.1723	0.0995	0.6857	0.6267	11.633	8.3223
28.	^{5}BIC d ^{1}IC ST	13.1583(±1.6011) 2.4410(±2.2455) -2.5909(±0.8185) -0.0655(±0.0123)	-1.5082	0.0490	0.9287	0.9097	48.842	19.6672
29.	^{5}BIC ^{0}IC ^{1}IC ST	12.9805(±1.9479) 0.8512(±1.4139) -2.3533(±1.0498) -0.0537(±0.0101)	-0.7754	0.0502	0.9249	0.9049	46.179	19.1577
30.	^{0}BIC ^{5}BIC ^{1}IC ST	3.7827(±4.1451) 11.2780(±1.5885) -2.1515(±0.5108) -0.0506(±0.0109)	0.1648	0.0495	0.9271	0.9077	47.708	19.4517
31.	^{5}BIC ^{1}IC MATS1v ST	12.7605(±1.4427) -1.7721(±0.2369) 1.2242(±1.2281) -0.0594(±0.0095)	-0.5252	0.0492	0.9279	0.9086	48.231	19.5788
32.	^{5}BIC ^{1}IC MATS2v ST	12.8033(±1.7010) -1.7602(±0.2414) -0.4872(±0.7655) -0.0570(±0.0092)	-0.6978	0.0502	0.9251	0.9051	46.318	19.1598
33.	^{5}BIC ^{1}IC MATS3v ST	9.7785(±1.6303) -1.7383(±0.2135) -1.2476(±0.6033) -0.0512(±0.0085)	1.1755	0.0448	0.9401	0.9242	58.899	21.6426
34.	^{5}BIC ^{1}IC MATS4v ST	10.3058(±1.2209) -1.7430(±0.1922) 1.1180(±0.3770) -0.0551(±0.0074)	1.0592	0.0404	0.9515	0.9386	73.586	24.1448

Table 5. Observed and estimated values of $logIC_{50}$ using model No. 34 .

Comp.No.	Obs. $logIC_{50}$	Est. $logIC_{50}$	Residual
1	3.04	3.07	-0.03
2	2.28	2.32	-0.04
3	2.45	2.46	-0.02
4	2.34	2.25	0.09
5	2.20	2.21	-0.01
6	2.26	2.35	-0.10
7	2.32	2.17	0.15
8	2.08	2.04	0.04
9	1.79	1.79	-0.01
10	1.76	1.86	-0.10
11	1.93	2.00	-0.07
12	1.89	1.91	-0.02
13	1.49	1.57	-0.07
14	2.26	2.35	-0.09
15	2.20	2.08	0.13
16	1.75	1.70	0.05
17	1.66	1.68	-0.02
18	2.00	2.03	-0.03
19	2.30	2.19	0.11
20	2.15	2.11	0.04

The predictive power of this model comes out to be 0.9515, indicating that about 95% of the data is explained by this model. The estimated activity values using the best four- parametric model has been reported in Table 5, and are in good agreement with the observed ones confirming that the proposed four-parametric model is best suitable for modeling, estimating $logIC_{50}$ activity of present set of compounds. All the above models have been tested using cross validated parameters. These parameters are reported in Table 6. It is worth mentioning that PRESS is a good estimate of the real predictive power of the model. If PRESS is smaller than SSY, the model predicts better than chance and can be considered statistically significant. Table 6 shows that in this regard, all the models proposed by us are better than chance and are statistically significant. The ratio PRESS / SSY can be used to calculate the approximate confidence interval of the prediction of new compounds. To be a reasonable QSAR model, this ratio should be smaller than 0.4. The models proposed by us are found to have this ratio smaller than 0.4 and the model expressed by equation 8 ,9, 10 and 11 has the excellent predictive power. The developed models are cross-validated by leave-one-out method. The high values observed in case of eqn. 11 (R^2_{CV} = 0.9490) are indicative of their reliability in prediction of biological activity. Another cross-validated parameter related to uncertainty of prediction, the PSE, is calculated. The lowest value of PSE for model 34 (eq.11) supports its highest predictive potential (power). The highest R^2_{CV} and lowest PSE for the model 34 shows that this is the most appropriate model for modeling $logIC_{50}$ value of 20 compounds used in the present study.

We have further carried out analysis to test model-34 whether it suffers from the defect due to co- linearity. For this we have, subjected this model to Ridge analysis and calculated Ridge traces [Fig. 3 and 4]. All these results have finally demonstrated that the proposed model-34 is the most appropriate model for modeling the activity and that it is devoid of any co- linearity defect

Table 6. Cross validated parameters for different models.

Model No.	Parameters used	PRESS/SSY	R^2_{cv}	S_{PRESS}	PSE
1.	ST	1.2548	-0.2548	0.2631	0.2496
2.	^{5}BIC ST	0.5180	0.4820	0.2120	0.1954
3.	^{5}BIC ^{1}IC ST	0.0833	0.9167	0.1037	0.0928
4.	^{5}BIC ^{1}IC MATS4v ST	0.0510	0.9490	0.0851	0.0737

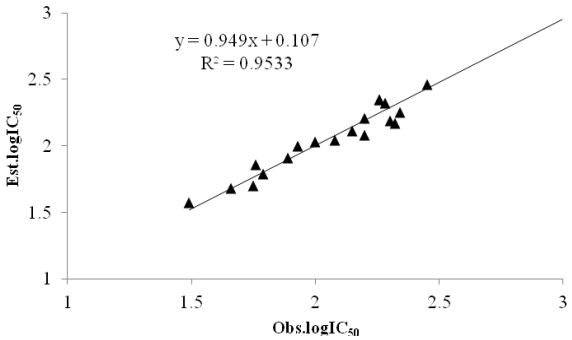

Figure 2. Correlation between Observed and estimated activity for model 34.

Table7. Ridge analysis parameters for four parametric model.

Model No.	Parameters	VIF	Tolerence	Eigenvalue	Condition no.
4	ST	1.0988	0.9101	1.8510	1.00
	^{1}IC	1.3156	0.7601	1.1269	1.64
	^{5}BIC	1.7610	0.5679	0.6461	2.86
	MATS4v	1.4603	0.6848	0.3758	4.92

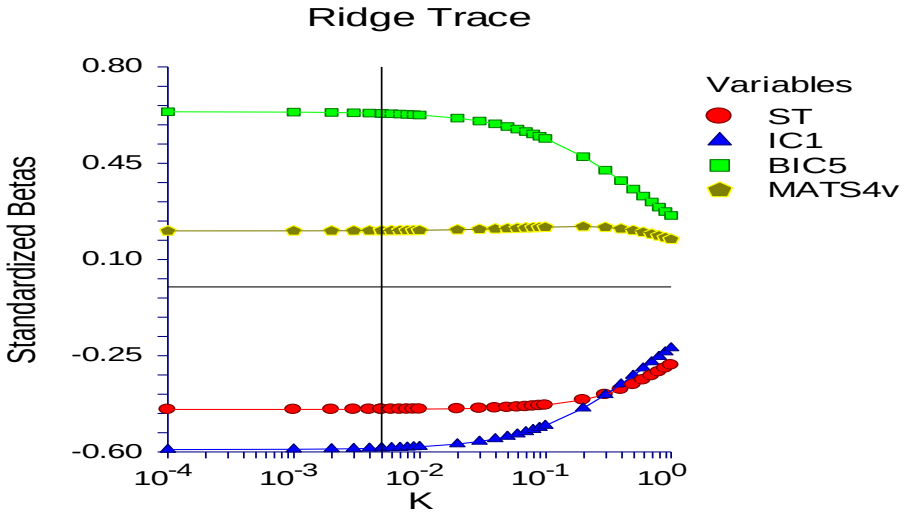

Fig.3: Ridge trace for four-variable model

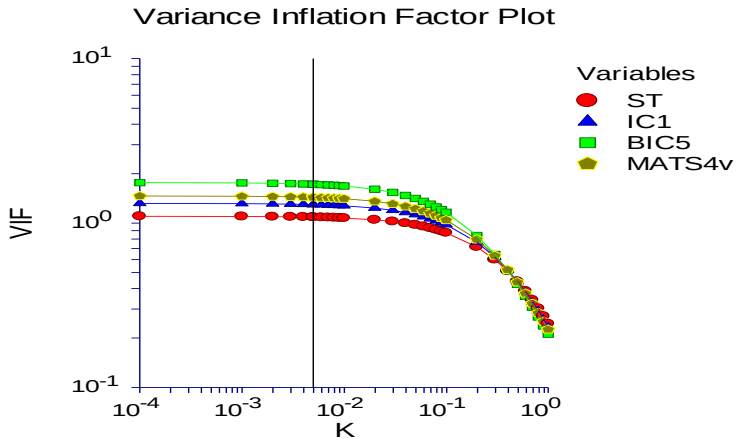

Fig.4 : VIF plot for four variable model

4. Conclusion

On the basis of above discussion we may conclude that:

1. Surface Tension (ST) plays an important role while modeling the antimalarial activity .

2. 2D- autocorrelation and informational descriptors are good for modeling the antimalarial activity of present set of compounds.

3. Higher the value of ^{5}BIC and MATS4v and lower the value of ^{1}IC and ST the better will be the antimalarial activity.

Therefore while modifying the molecular structure for better activity above points should be kept in consideration.

Acknowledgment

The authors are thankful to the National Institute of Technical Teachers Training Institute, Bhopal and management of Career College, Bhopal for their support, co-operation, and encouragement.

References

[1] D. M. Opsenica, B.A. Solaja, *J. Ser. Chem. Soc.*, 74, (2009),1155–1193.

[2] R. Arav-Boger, T.A. Shapiro, *Ann. Rev. Pharm. Toxic.* 45, (2005), 565–585.

[3] N.J. White, *J. Cl. Inv.*, 113, (2004), 1084–1092.

[4] R.G. Ridle, *Nature*, 424, (2003), 887–889.

[5] P.M. O'Neill, N.L. Searle, K.W. Kan, R.C. Storr, J.L. Maggs, S.A. Ward, K. Raynes, B.K. Park, *J. Med. Chem.*, 42, (1999), 5487–5493.

[6] J.R. Woolfrey, M.A. Avery, A.M. Doweyko, *J. Comp. Aid. Mol. Des.* 12, (1998), 165–181.

[7] N.C. Cohen, *Guidebook on Molecular Modeling in Drug Design*; Academic Press: San Diego, CA, USA, (1996).

[8] C.M.R. Sant'Anna, *Qui. Nov*, 25, (2002), 505–512.

[9] Carvalho, A.D.L. Borges, L.S.C. Bernardes, *J. Chem. Educ.*82, (2005), 588–596.

[10] C.G. Wermuth, *The Practice of Medicinal Chemistry*, 3rd ed.; Academic Press: London, UK, (2009).

[11] F.A.L. Ribeiro, M.M.C. Ferreira, *J. Mol. Str. Theo*, 663, (2003), 109–126.

[12] G.H. Posner, H. O'Dowd, P. Ploypradith, J.N. Comming, S. Xie, T.A. Shpiro, *J. Med. Chem.*, 41, (1998),2164-2167.

[13] NCSS statistical software, www.ncss.com

[14] S. Chatterjee , A.S. Hadi , B. Price, *Regression Analysis by Examples*, 3rd Ed., Wiley, New York,51, (2000).
[15] DRAGON : Software www.disat.unimib.it
[16] B. Shaik , J. Singh , S. Singh , V.K. Agrawal , P.V. Khadikar, C.T. Supuran, *Chem. Bio. Drug Des.*, 71, (2008), 244-259.
[17] B. Shaik , J. Singh, S. Singh, N. Sohani ,V.K. Agrawal , P.V. Khadikar, *Chem. Bio. Drug Des.*, 71, (2008), 230- 243.
[18] B. Louis , J. Singh , B. Shaik , V.K. Agrawal ,P.V. Khadikar , *Chem. Bio. Drug Des*, 74, (2009), 190-195.
[19] J. Singh, I . Ahmad , B. Shaik , V.K. Agrawal , P.V. Khadikar, *J. Ind. Chem. Soc*., 86, (2009),1197-1203.
[20] R. Sarkar, A.B. Roy, P.K. Sarkar , *Math. Biosci.*,39, (1978), 299-312.
[21] S.C. Basak , V.R. Magnuson, *Arzneim. Forsch*. 33, (1983),501-503.
[22] D. Bonchev , N. Trinajstic, *J. Chern. Phys*. 67, (1977), 4517-4533.
[23]A.B. Roy , S.C. Basak , D.K. Harriss ,V.R. Magnuson, *Mathematical Modelling in Science and Technology*, Pergamon Press, New York, 745, (1984).
[24] C.E. Shannon, *Bell Syst. Tech. J*., 27, (1948), 379-423.
[25] L.B. Kier , *J. Pharm. Sci.*, 69, (1980), 807-810.
[26] S.C. Basak , D.K. Harriss, V.R. Magnuson, *J. Pharm. Sc*., 73, (1984), 429-437.
[27] L. Saiz-Urra ., M.P. Gonzalez , M. Teijeira, *Bioorg. Med. Chem.*, 15, (2007), 3565-3571.
[28] J. Caballero , M. Garriga , M. Fernandez, *Bioorg. Med. Chem.*, 14, (2006), 3330-3340.
[29] ACD-Labs software ChemSketch. www.acdlabs.com
[30] L. Pogliani, *J. Phys. Chem*.**,** 100, (1996),18065-18077.
[31] L. Pogliani, *Amino Acids*.**,**6, (1994), 141-153.