Research article Available online www.ijrpsonline.com ISSN: 2249–3522

International Journal of Research in Pharmacy and Science

Antimalarial Activity: A QSAR Modeling of NF54 Strain of *Plasmodium falciparum* by Physicochemical Descriptor Calculation

Singh Vinayaditya*, Argal Ameeta, Mishra Vikash,
Raghuvanshi Ramsneh, Agnihotri Savita

Rajeev Gandhi College of Pharmacy, Salaiya, Bhopal, Madhya Pradesh, India

ABSTRACT

Multiple linear regression analysis and other statistical analysis were carried out on all the compounds of training set. Descriptors were selected for the model based on their correlation coefficient and those descriptors having interred correlation coefficient below 0.5 were considered. Various models were obtained after performing multiple linear regression analysis. The dataset are based on the NF54 strains of *P. falciparum* consisting of 69 organic compounds. The size of the final training set therefore became 53 compounds and test set 16 compounds. Descriptors were selected for the model based on their correlation coefficient and that descriptor having intercorrelation coefficient below 0.5 were considered. Various models were obtained after performing multiple linear regression (MLR) analysis. Model predictive power was judged based on various statistical parameters like correlation coefficient, regression coefficient (r^2), fisher statistical value (F), and standard error. The initial regression analysis was performed on all the 37 molecules which resulted in regression model. The best QSAR model has the characters of large F, low P value, r^2 and q^2 values close to 1, as well as $P< 0.001$.

KEY WORDS: *P. falciparum,* QSAR, Multiple linear regression analysis, Correlation coefficient, Regression coefficient

***Corresponding author:-**

Vinayaditya Singh
Rajeev Gandhi College of Pharmacy,
Salaiya, Bhopal,
Madhya Pradesh, India.
Phone: +91-9926670790
E Mail: mail4vinayaditya@gmail.com

INTRODUCTION

During its intraerythrocytic life cycle, a single *P. falciparum* parasite undergoes multiple morphological and physiological changes and multiplies to produce up to 36 new daughter parasites in ~48 hours. Large-scale genomic and proteomic analyses revealed a coordinated program of gene and protein expression during parasite intraerythrocytic life cycle[1-6]. Proteasomes are multicatalytic protease complexes whose principle task is the selective degradation of proteins within the cell. Although a fully intact proteasome has not been isolated from *P. falciparum*, the sequencing of this organism revealed a complete set of ORFs encoding homologs of eukaryotic subunits of the proteasome[7-9].

MATERIAL AND METHOD

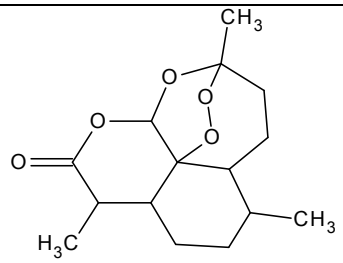
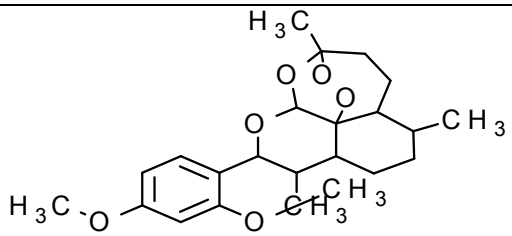
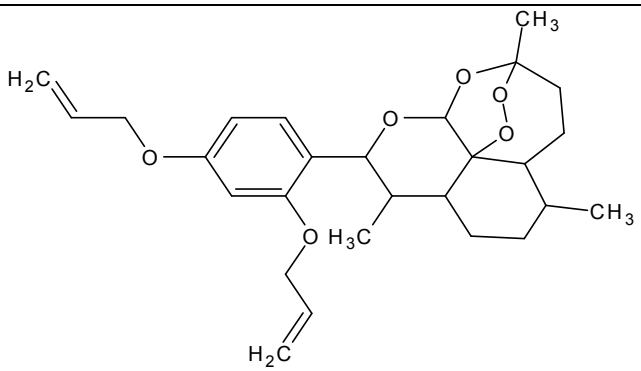
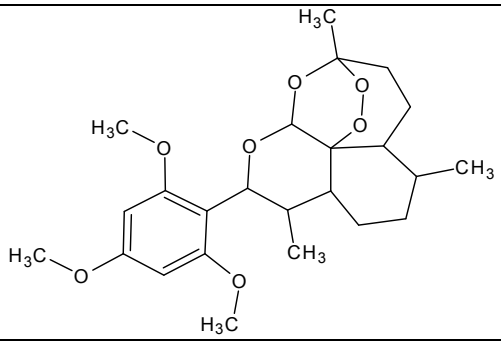
For performing the QSAR study personal computer is used. A variety of commercial and non-commercial packages are available for PC- based system to super computer systems. For the molecular structure generation and calculation of descriptors Chemsketch 12.0 is used and VALSTAT is used for the statistical analysis.

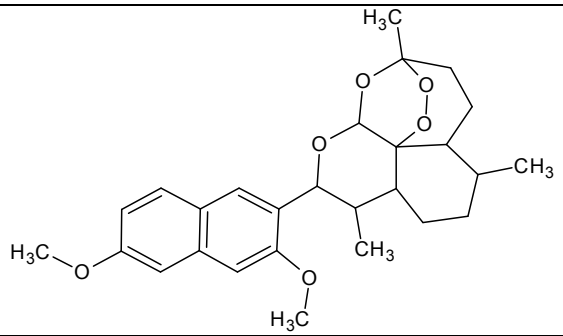
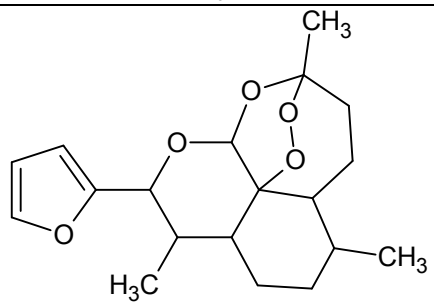
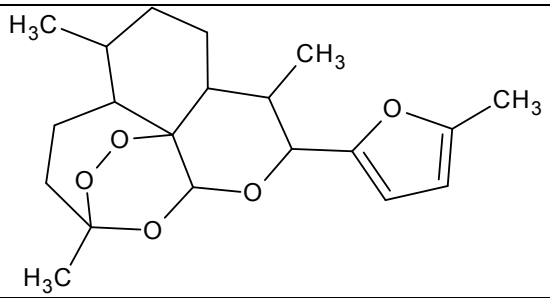
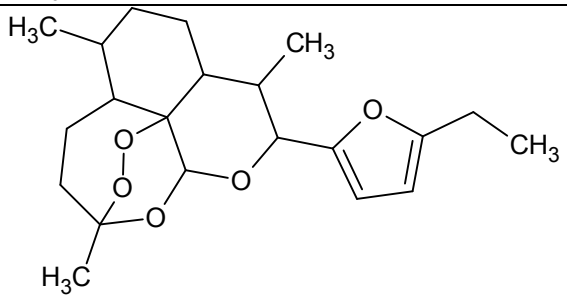
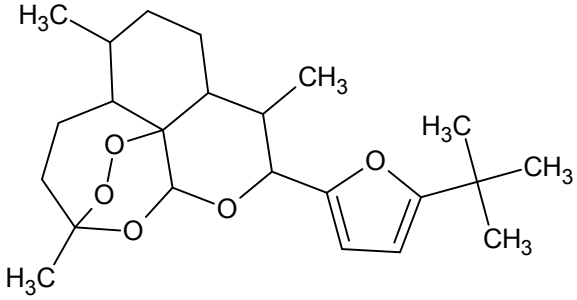
Data Set

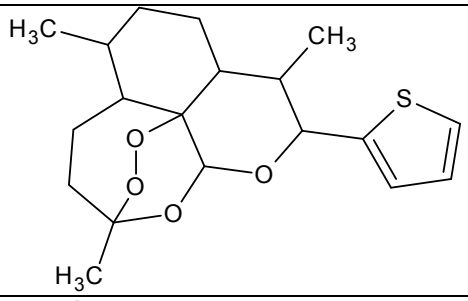
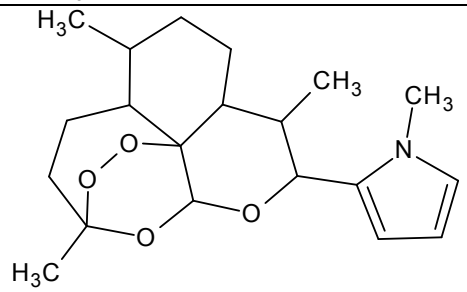
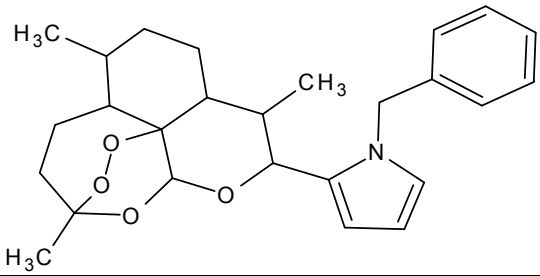
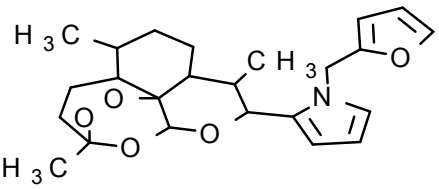
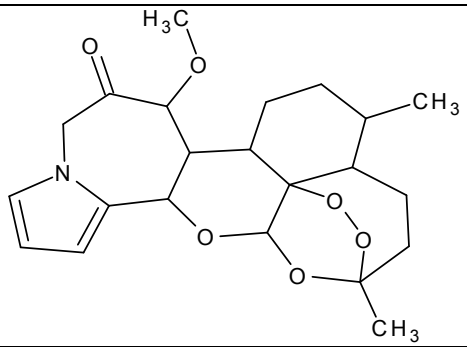
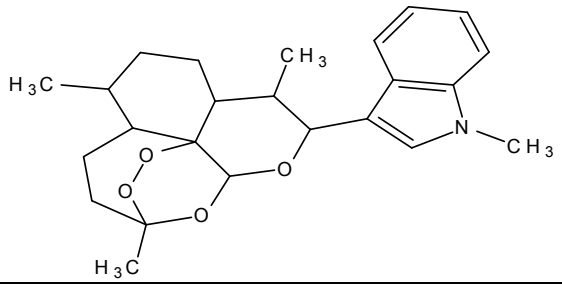
In QSAR analysis, it is imperative that the biological data be both accurate and precise to Develope a meaningful model. The overall performance of the current method used for QSAR study is critically depends on the selection of compounds for series used to build the classifier model. The most critical aspect of the construction of the series is to warrant a great molecular diversity in this data set. The biological data used in this study are the enzyme inhibition activity (as IC50) of a series of Artemisinin Analogs & 1,3,5-triazine substituted polyamines along with cyclic per oxy ketals are taken as antimalarial agents. The dataset are based on the NF54 strains of P. falciparum consisting of 69 organic compounds[10] The structural and biological activity of these compounds was listed in table 4.1. The biological activity data (ic50 in µm) was converted into negative logarithmic dose (pIC) for QSAR analysis. The data set of Artemisinin Analogs & 1,3,5-triazine substituted polyamines as antimalarial agents consist of 69 compounds. For the validation of QSAR models, statistical external validation was used and the molecules were rationally divided into training and test set. The test set should represent a balanced number of both active and inactive compounds for uniform sampling of data. Therefore, the structure and activity diversity in both sets is maintained for QSAR models development. The test set molecules captured structural features of training set molecules, thus their activities could be well predicted the size of the training set was aimed to be about two third of the whole set. Before the final

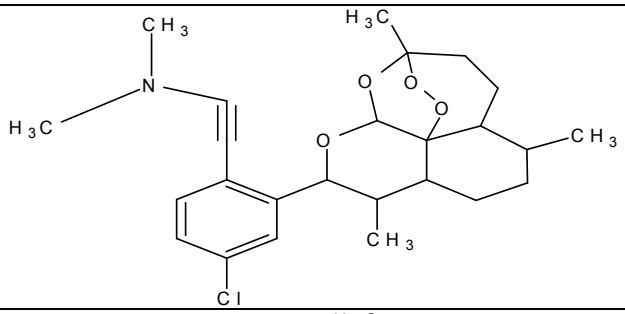
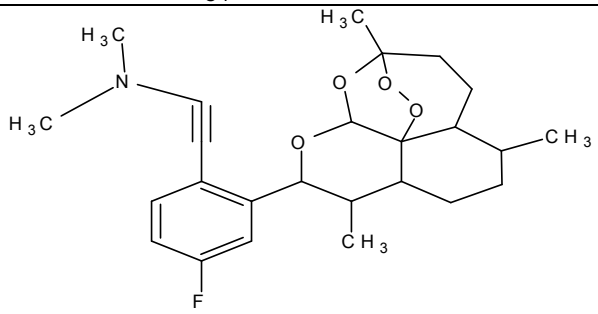
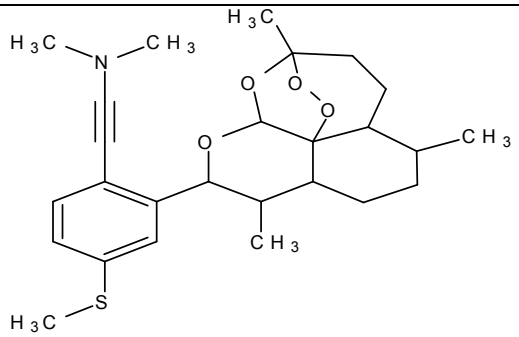
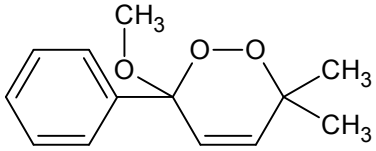
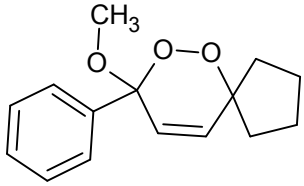
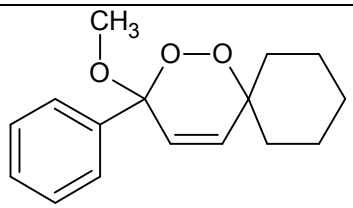
model development outliers detected in the training set were moved to the test set. The size of final training set therefore became 53 compounds.

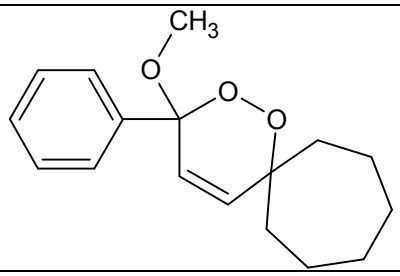
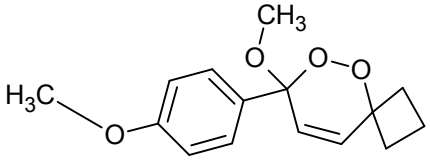
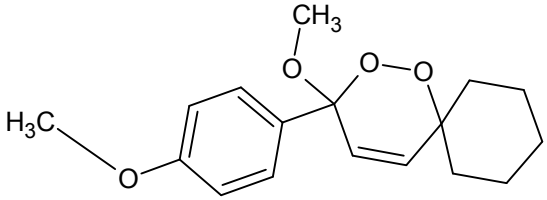
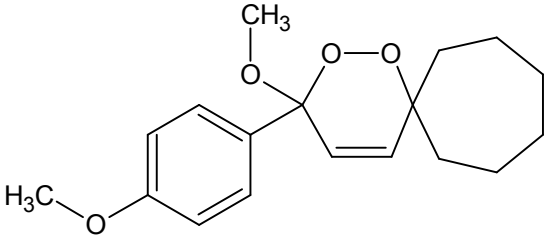
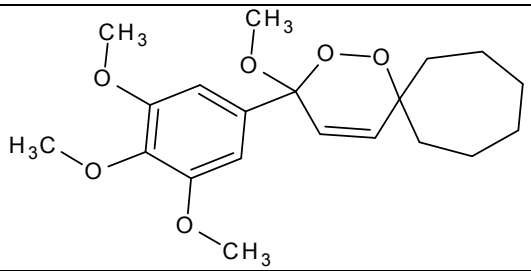
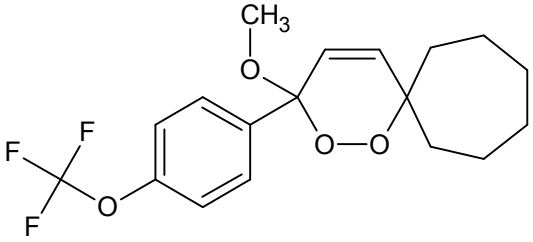
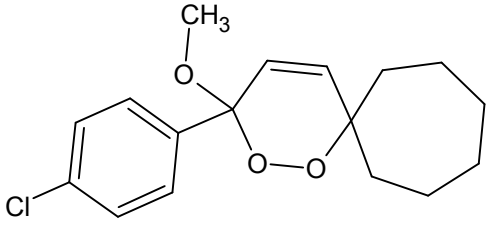
Table 3: Chemical structures and activity of compounds

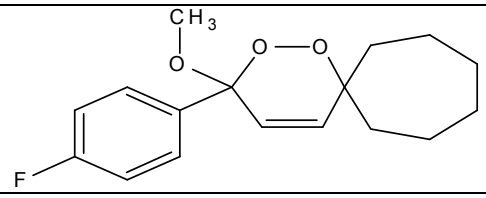
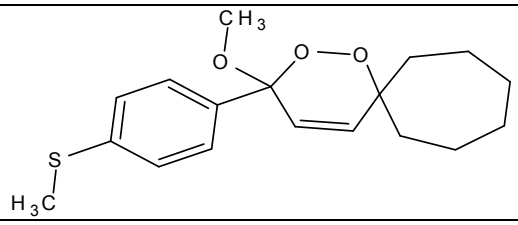
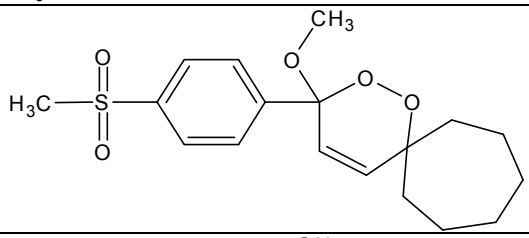
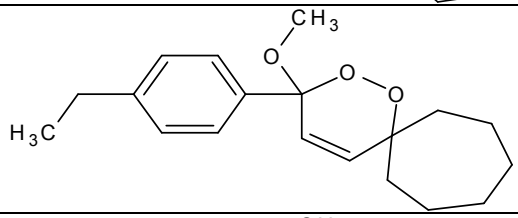
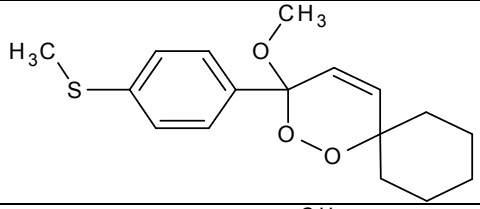
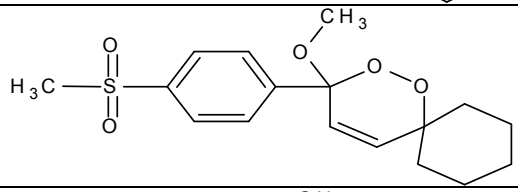
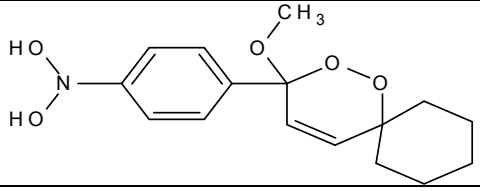
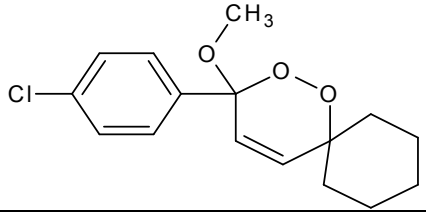
Comp. no.	Structure	Obs. B.A.
1		1.00
2		0.62
3		0.82
4		0.89

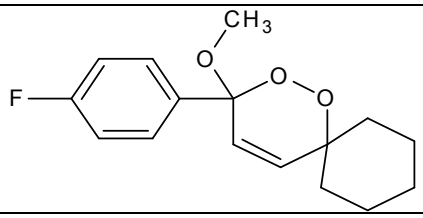
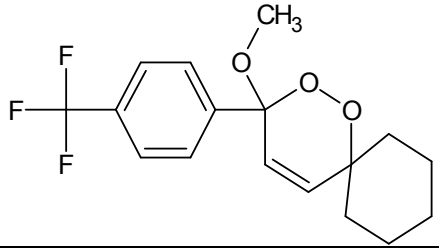
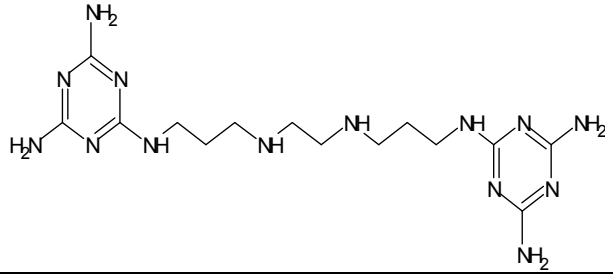
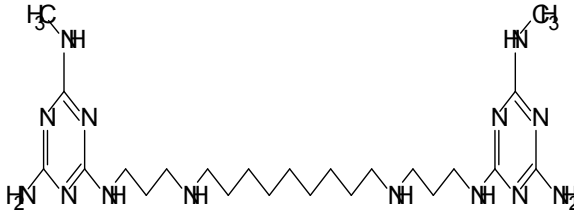
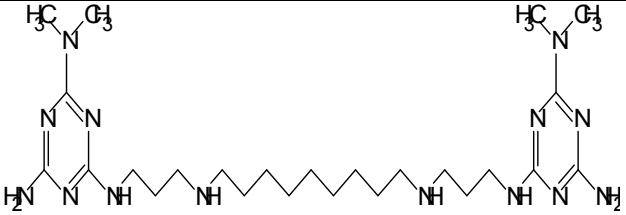
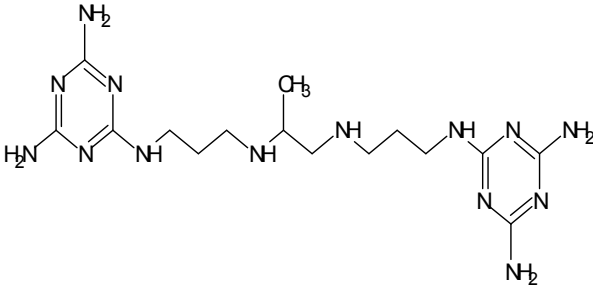
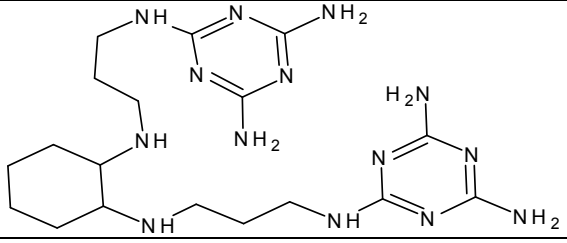
5		0.95
6		0.15
7		0.72
8		0.93
9		1.00

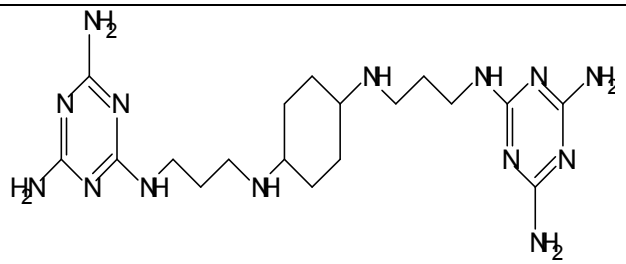
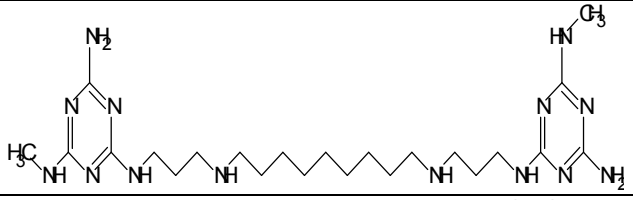
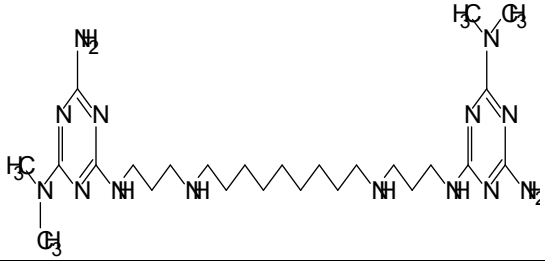
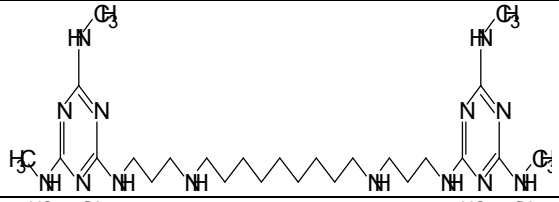
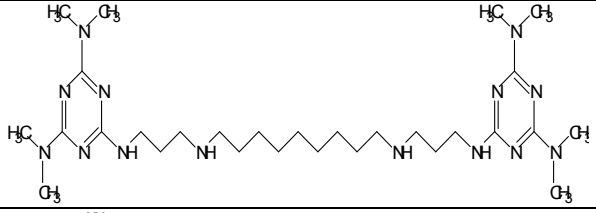
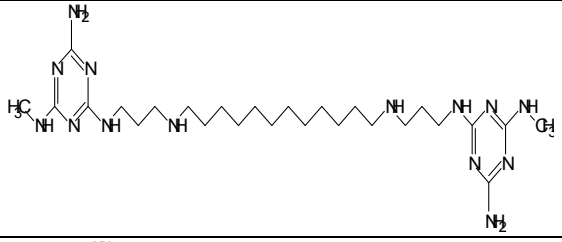
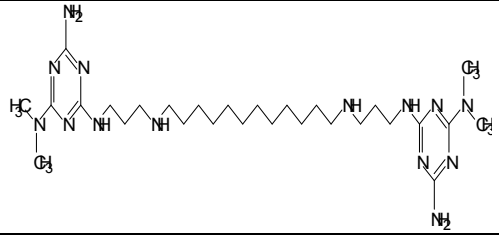
10		0.71
11		0.66
12		1.20
13		0.97
14		0.96
15		0.60

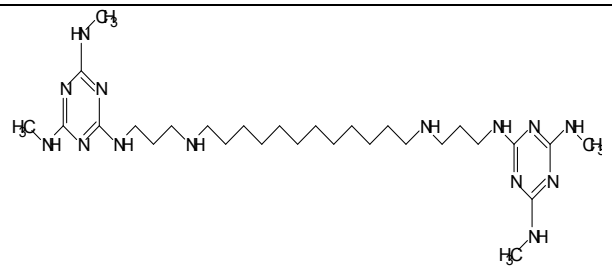
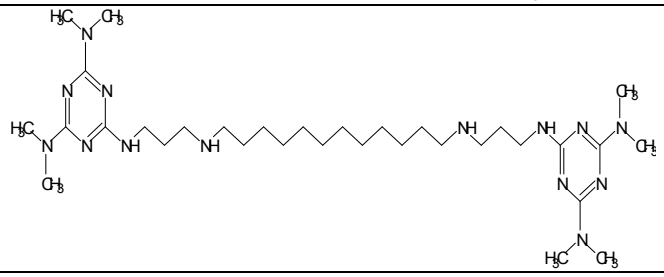
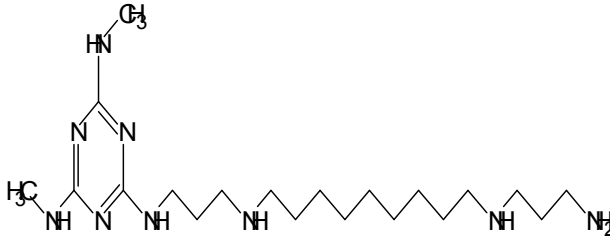
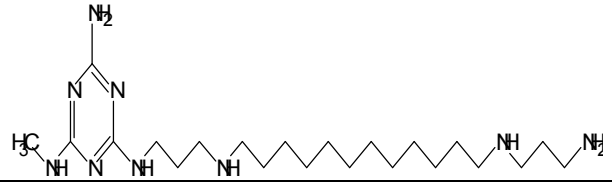
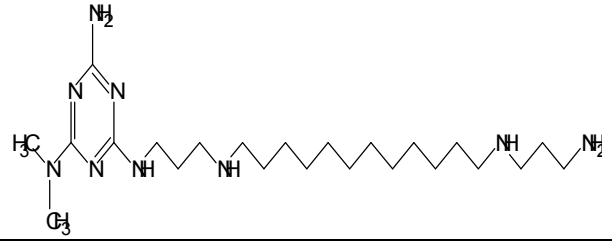
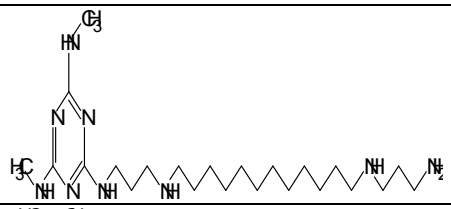
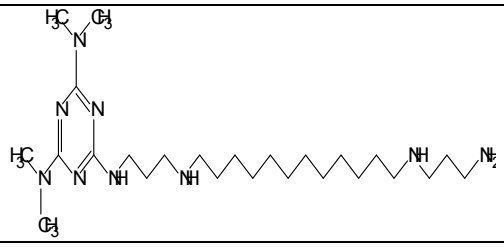
16		1.04
17		0.92
18		0.92
19		3.04
20		2.28
21		2.45

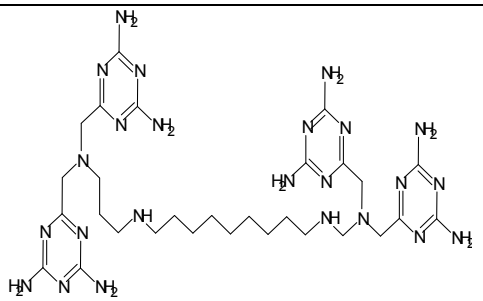
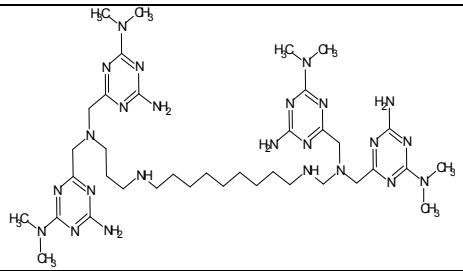
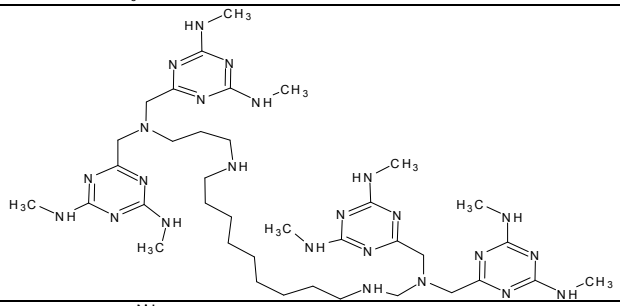
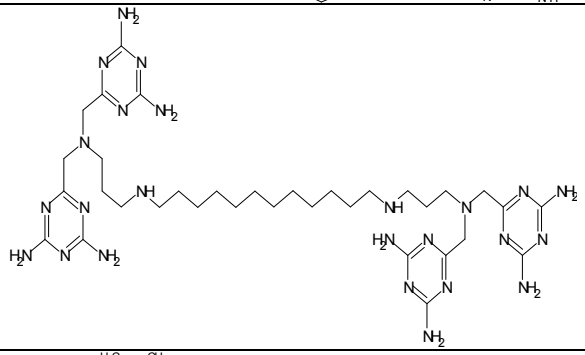
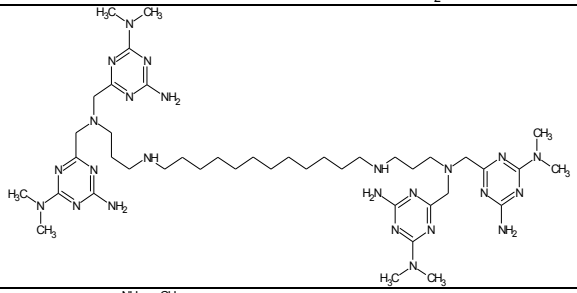
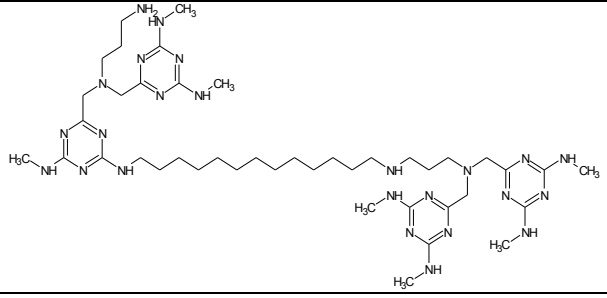
22		2.34
23		2.20
24		2.26
25		2.32
26		2.08
27		1.79
28		1.76

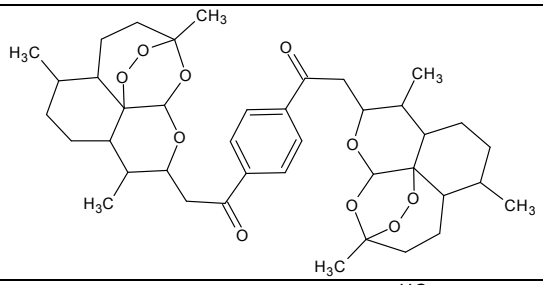
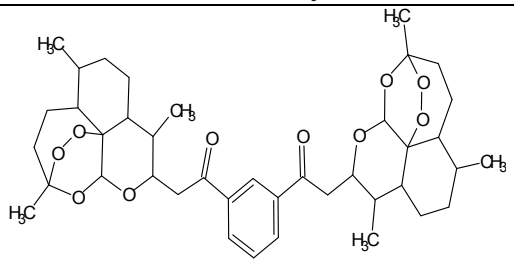
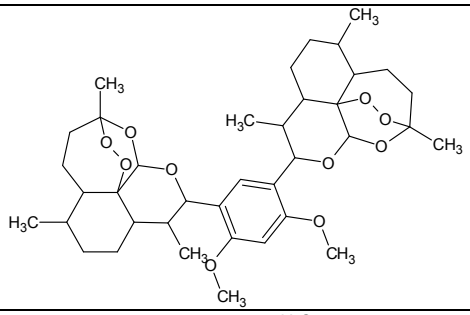
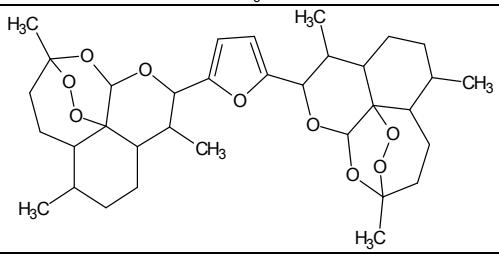
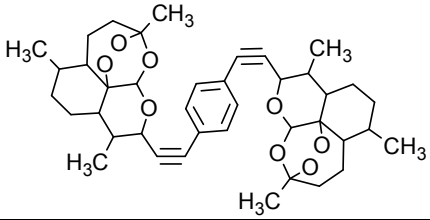
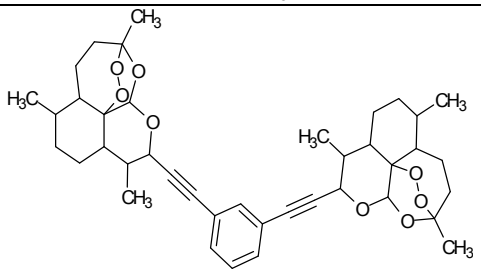
29		1.93
30		1.89
31		1.49
32		2.26
33		2.20
34		1.75
35		1.66
36		2.00

37		2.30
38		2.15
39		4.25
40		3.64
41		2.79
42		4.26
43		4.23

44		4.23
45		3.60
46		3.29
47		3.17
48		2.76
49		2.91
50		2.69

51		2.67
52		2.24
53		2.76
54		3.31
55		2.79
56		3.24
57		2.92

58		3.88
59		2.21
60		1.86
61		3.19
62		1.84
63		1.65

64		0.28
65		0.28
66		0.11
67		0.51
68		1.48
69		1.56

Molecular Structure Generation

All structure of Artemisinin Analogs & 1,3,5-triazine substituted polyamines as antimalarial agents compounds were constructed using ACD/LABs product version 12.01 supplied by Advance chemistry development Inc. All 2D (2-Dimentional) structure is converted into MOL structures in Chemsketch 12.0 2009.

Energy Minimization

The molecular mechanics (MM2) method was applied to search for lower energy conformations for each molecule. The energy minimized molecules were re-optimizing using molecular orbital package (MOPAC). To avoid the local stable conformations of the compounds, geometry optimization was run many times with different starting points of each molecule, and the conformation with the lowest energy was considered for calculation of the molecular descriptors. This was done by software Chemoffice 6.0.

Descriptors Calculation

The physicochemical properties of each molecule are calculated for QSAR analysis. Different classes of properties like Molar volume, Index of Refraction, Parachor etc. were calculated by the Chemsketch 12.0 for QSAR analysis and reported in table 4.2. Thermodynamic parameters describe free energy change during drug receptor complex formation. Spatial parameters are the quantified steric features of drug molecules required for its complimentary fit with receptor. Electronic parameters describe weak non-covalent bonding between drug molecules and receptor.

Statistical Analysis

All the data set (69 molecules) were divided into two sets. First one training set having 53 molecules for generation of QSAR models and second test set having 16 molecules for validation of generated QSAR models. VALSTAT software was used to generate QSAR models by multiple linear regression analysis. The inter-correlation between the parameters was less than 0.5 which show inter-pair correlations among the selected descriptors are very low. Acceptability of the regression model was judged by examining the different statistical parameters i.e. number of samples in regression (n), regression coefficient (r), squared regression coefficient (r2), adjusted squared regression coefficient (r2adj), F-test (Fischer's value) for statistical significance, standard error of estimate (std), cross-validated squared correlation coefficient (q2), boot strapped squared correlation coefficient (bsr2),

friedman lack of fit measure (LOF), quality factor (QF), Probable Error of correlation (PE), Kubinyi function (FIT), Akaike's Information Criterion (AIC) and correlation matrix to show mutual correlation among the parameters.

Model Development and Validation

Internal and external validation was performed to validate the QSAR model. For external validation, the activity of each compound in test set was computed. With the help of observed activity and calculated activity cross-validation coefficient q2 was calculated.

Cross validation was performed using leave-one-out method. For multiple linear regression analysis biological activity (-logIC50) values was used as dependent variables and calculated parameters (descriptors) used as independent variables.

Table 4: Values of different type of descriptors calculated for training set

S.N.	OBS.	F.W.	M.R.	M.V.	P.C.	I.R.	S.T.	DEN	POL
1	1.00	282.3321	70.31	226.4	581.7	1.533	43.5	1.24	27.87
2	0.62	404.4965	107.61	332.7	865.6	1.56	45.7	1.21	42.66
3	0.82	456.5711	125.68	387.8	1004.2	1.561	44.9	1.17	49.82
4	0.89	434.5225	113.98	354.4	924.2	1.556	46.2	1.22	45.18
5	0.95	454.5552	125.17	365.5	970.5	1.6	49.6	1.24	49.62
6	0.15	334.4067	87.37	271.7	703.6	1.556	44.9	1.23	34.63
7	0.72	348.4333	91.99	287.4	741.9	1.552	44.3	1.21	36.46
8	0.93	362.4599	96.62	303.6	781.9	1.549	43.9	1.19	38.3
9	1.00	390.513	105.89	336.3	858.4	1.542	42.4	1.16	41.97
10	0.71	350.4723	93.03	279.5	733.2	1.579	47.3	1.25	36.88
11	0.66	347.4485	91.8	255.9	663	1.636	45	1.35	36.39
12	1.20	423.5445	117.1	324.6	847.4	1.64	46.4	1.3	46.42
13	0.97	413.5066	109.27	298.9	789.3	1.651	48.5	1.38	43.31
14	0.96	403.4687	100.9	264.1	717.8	1.689	54.5	1.52	40
15	0.60	397.5072	107.69	291.6	762.7	1.66	46.7	1.36	42.69
16	1.04	445.9788	120.12	353.8	947.6	1.594	51.4	1.26	47.62
17	0.92	429.5242	115.41	347.5	917.8	1.578	48.6	1.23	45.75
18	0.92	457.6254	127.93	372.5	1001.2	1.602	52.1	1.22	50.71
19	3.04	220.2643	62.26	199.2	495.7	1.537	38.3	1.1	24.68
20	2.28	246.3016	69.49	213	546.4	1.565	43.2	1.15	27.54

21	2.45	260.3282	74.12	229.3	586.5	1.559	42.7	1.13	29.38
22	2.34	274.3547	78.75	245.6	626.6	1.554	42.3	1.11	31.22
23	2.20	262.3010	71.22	218.7	565.0	1.564	44.4	1.19	28.23
24	2.26	290.3541	80.48	251.2	645.1	1.553	43.4	1.15	31.90
25	2.32	304.3807	85.11	267.5	685.2	1.549	43.0	1.13	33.74
26	2.08	364.4327	97.84	311.1	802.4	1.541	44.2	1.17	38.79
27	1.79	358.3521	85.48	282.8	708.9	1.516	39.4	1.26	33.88
28	1.76	308.7998	83.58	256.5	663.7	1.565	44.8	1.20	33.13
29	1.93	292.3452	78.86	250.1	633.9	1.543	41.2	1.16	31.26
30	1.89	320.4463	91.39	275.2	717.2	1.578	46.1	1.16	36.23
31	1.49	352.4451	92.19	280.6	744.5	1.570	49.5	1.25	36.55
32	2.26	302.4079	88.00	277.6	704.9	1.546	41.5	1.08	34.88
33	2.20	306.4197	86.76	259.0	677.2	1.584	46.7	1.18	34.39
34	1.75	338.4185	87.56	264.4	704.4	1.576	50.3	1.27	34.71
35	1.66	307.3416	81.07	230.0	645.0	1.622	61.7	1.33	32.13
36	2.00	294.7732	78.95	240.2	623.6	1.571	45.3	1.22	31.29
37	2.30	278.3186	74.23	233.9	593.8	1.547	41.5	1.18	29.42
38	2.15	328.3261	79.11	260.6	648.5	1.519	38.3	1.25	31.36
39	4.25	392.4659	110.15	274.3	863.0	1.735	97.8	1.430	43.66
40	3.64	518.7051	153.33	424.2	1207.6	1.642	65.6	1.222	60.78
41	2.79	546.7583	162.73	461.3	1293.8	1.623	61.8	1.185	64.51
42	4.26	406.4925	114.74	291.2	900.1	1.717	91.2	1.395	45.48
43	4.23	446.5563	120.81	328.0	1022.8	1.657	94.5	1.36	47.89
44	4.23	446.5563	120.81	328.0	1022.8	1.657	94.5	1.36	47.89
45	3.60	518.7051	153.33	424.2	1207.6	1.642	65.6	1.222	60.78
46	3.29	546.7583	162.73	461.3	1293.8	1.623	61.8	1.185	64.51
47	3.17	546.7583	164.07	458.6	1273.7	1.634	59.4	1.192	65.04
48	2.76	602.8646	182.89	532.7	1446.2	1.602	54.3	1.131	72.5
49	2.91	560.7848	167.23	473.7	1326.9	1.623	61.5	1.183	66.29
50	2.69	588.838	176.63	510.8	1413.2	1.607	58.5	1.152	70.02
51	2.67	588.838	177.97	508.1	1393	1.617	56.4	1.158	70.55
52	2.24	644.9443	196.79	582.2	1565.6	1.591	52.2	1.107	78.01
53	2.76	409.6157	124.94	379.8	1008.3	1.571	49.6	1.078	49.53
54	3.31	437.6688	133.46	412.2	1094.6	1.56	49.7	1.061	52.9
55	2.79	451.6954	138.16	430.7	1137.7	1.554	48.6	1.048	54.77

56	3.24	451.6954	138.83	429.3	1127.6	1.559	47.5	1.051	55.04
57	2.92	479.7486	148.24	466.4	1213.9	1.548	45.8	1.028	58.76
58	3.88	736.8839	204.37	521	1640.8	1.713	98.3	1.414	81.02
59	2.21	849.0966	244.68	663.8	1945.5	1.658	73.7	1.279	97
60	1.86	849.0966	247.36	658.4	1905.2	1.674	70.1	1.289	98.06
61	3.19	807.0168	227.53	603.5	1839.7	1.677	86.3	1.337	90.2
62	1.84	919.2295	267.84	746.3	2144.4	1.636	68.1	1.231	106.18
63	1.65	919.2295	270.46	735.8	2105.2	1.656	67	1.249	107.22
64	0.28	694.8507	181.78	540.9	1462.1	1.586	53.3	1.28	72.06
65	0.28	694.8507	181.78	540.9	1462.1	1.586	53.3	1.28	72.06
66	0.11	670.8293	175.75	526.4	1408.6	1.582	51.2	1.27	69.67
67	0.51	600.7395	155.5	466	1246.7	1.581	51.1	1.28	61.64
68	1.48	658.8202	177.57	514.1	1406.3	1.607	55.9	1.28	70.39
69	1.56	658.8202	177.57	514.1	1406.3	1.607	55.9	1.28	70.39

RESULT AND DISCUSSION:

Multiple linear regression analysis and other statistical analysis were carried out on all the compounds of training set. Descriptors were selected for the model based on their correlation coefficient and those descriptors having interred correlation coefficient below 0.5 were considered. Various models were obtained after performing multiple linear regression analysis.

Model predicted power was judged based on various statistical parameters like:

r = Multiple correlation coefficient,

r2 = Explained variance (squared multiple r),

s = Standard error of estimate,

F = Variance ratio at specific degree of freedom (Fischer's F-test for significance)

The r2 static was a measure of the extent to which the total variation of the dependent variable was explained by the regression. A high r2 value suggests that the regression model explain the variation in the dependent variable well.

The QSAR model having higher r2 and F- ratio among the several models was tested by external validated procedure. The output of cross validation are q2 = cross validated r2;The initial regression analysis was performed on all the 53 molecules of training set which resulted in regression model. The best model has the characters of large F, low P- value, r2 and q2 values close to 1, as well as p<0.001. Four statistically significant QSAR models have been developed by using multiple linear regression analysis

Table 5: Pearson correlation matrix between selected descriptor

	OBS 1/C								
OBS 1/C	1	FW							
FW	0.0388	1	MR						
MR	0.1271	0.9887	1	MV					
MV	0.0730	0.9767	0.9888	1	PC				
PC	0.1492	0.9860	0.9972	0.9912	1	IR			
IR	0.3557	0.5154	0.5152	0.3912	0.4836	1	ST		
ST	0.5853	0.5242	0.5243	0.4234	0.5328	0.8417	1	DEN	
DEN	-0.1220	0.2174	0.1210	0.0132	0.0932	0.7054	0.5553	1	POL
POL	0.1271	0.9887	1	0.9888	0.9972	0.5152	0.5243	0.1210	1

MODEL NO.1:

BA = [1.78131(± 4.24809)] +FW [-0.00313618(± 0.000493079)] +IR [4.78785(± 3.29487)] +ST [0.0779656(± 0.00917458)] +DENSITY [-8.31492(± 1.07985)]

Fraction contribution of FW is =-0.19226

Fraction contribution of IR is = 0.0875819

Fraction contribution of ST is = 0.431246

Fraction contribution of DENSITY is =-0.288912

n=53,r=0.892899,r^2=0.797268,r^2adj=0.780374,variance=0.281797,std=0.530846, QF=1.68203, PE=0.0185649, F=47.1916, FIT=2.73574, LOF=19.6256, AIC=0.340505

Model 1 explains only 79.7% variance in the antimalarial activity. It shows that descriptor formula weight (FW) and density contribute negatively; whereas Surface tension (ST) & Index of refraction (IR) contribute positively towards antimalarial activity.

MODEL NO.2:

BA = [-6.25906(± 3.65945)] +PC [-0.00147488(± 0.000174118)] +IR [12.6478(± 2.92158)] +ST [0.0759864(± 0.00740405)] +DENSITY [-11.903(± 0.983352)]

Fraction contribution of PC is= -0.16584

Fraction contribution of IR is = 0.176758

Fraction contribution of ST is = 0.328687

Fraction contribution of DENSITY is= -0.328715

n=51,r=0.937417,r^2=0.878751,r^2adj=0.868208, variance=0.17567,std=0.41913,QF=2.23658, PE=0.0113188, F=83.3465, FIT=5.04175, LOF=11.9151, AIC=0.203424

Model 2 explains 87.8% variance in the antimalarial activity. It shows that descriptor Index of refraction (IR) and Surface tension (ST) contribute positively, where as Parachor (PC) & Density contribute negatively towards antimalarial activity. It is good significant equation & makes a new hope for the development of new model.

MODEL NO.3:

BA = [-8.55948(± 3.71485)] +IR [14.2751(± 2.96932)] +ST [0.0712652(± 0.00729387)] +DENSITY [-11.9278(± 0.980829)] +POL [-0.0297548(± 0.00349377)]

Fraction contribution of IR is = 0.198708

Fraction contribution of ST is = 0.307043

Fraction contribution of DENSITY is= -0.328094

Fraction contribution of POL is = -0.166155

n=51,r=0.937843,r^2=0.87955,r^2adj=0.869076,variance=0.174514,std=0.417748,QF=2.245, PE=0.0112443, F=83.9751, FIT=5.07978, LOF=11.8367, AIC=0.202084

Model 3 explains 87.9% variance in the antimalarial activity. It shows that descriptor Density and Polarizability (POL) contribute negatively, whereas Index of refraction (IR) & Surface tension (ST) contribute positively towards antimalarial activity.

MODEL NO.4:

BA= [-4.57266(± 3.39459)] +MV [-0.00421025(± 0.000451976)] +IR [11.536(± 2.71954)] +ST [0.0717592(± 0.00675491)] +DENSITY [-11.5706(± 0.917241)]

Fraction contribution of MV is= -0.165395

Fraction contribution of IR is= 0.169514

Fraction contribution of ST is=0.327109

Fraction contribution of DENSITY is=-0.337982

n=50,r=0.945079,r^2=0.893175,r^2adj=0.883679, variance=0.149599,std=0.38678,QF=2.44345, PE=0.0100716, F=94.0622, FIT=5.81641, LOF=10.0118, AIC=0.169468

Model 4 explains 89.3% variance in the antimalarial activity with low standard error shows the relative good fitness of the model. It shows that descriptor molar volume (MV) and Density contribute negatively; whereas index of refraction (IR) & surface tension (ST) contribute positively towards antimalarial activity. The graph between experimental and predicted biological activity of training set compounds by using model 4 is shown in **Fig. 2.** The validation criteria for selection of the model are cross validated squared correlation coefficient (q^2). The cross validation correlation coefficient (q^2) was 0.8607 means model 4 have good predictive power. The graph between experimental BA and predicted BA of test set compounds by using model 4 is shown in **Fig. 3**.

Table 6: Predicted activities and residuals of training set from Model 4

Comp.	Obs. B.A.	Cal B.A.
1	1	0.922008
2	0.62	1.33137
5	0.95	1.57744
7	0.72	1.3322
8	0.93	1.42311
9	1	1.44194
10	0.71	1.42549
11	0.66	0.867233
14	0.96	-0.33098
15	0.6	1.08193
19	3.04	2.28496
20	2.28	2.38426
21	2.45	2.43494
22	2.34	2.52342
23	2.2	1.95482
24	2.26	2.08581
25	2.32	2.17466
26	2.08	1.49561
28	1.76	1.73017
29	1.93	1.69778
30	1.89	2.38087
32	2.26	2.59791
33	2.2	2.313
34	1.75	1.38844
37	2.3	1.57708
39	4.25	4.93862
42	4.26	4.44978
43	4.23	4.19881
44	4.23	4.19881
45	3.6	3.12049
46	3.29	2.9099
49	2.91	2.87913
50	2.69	2.68316
51	2.67	2.5829
53	2.76	3.06174
55	2.79	2.91417
56	3.24	2.81203
57	2.92	2.6869
58	3.88	3.64974
60	1.86	2.11721
61	3.19	2.92395
62	1.84	1.79379
63	1.65	1.81811
64	0.28	0.484961

65	0.28	0.572194
66	0.11	0.481097
67	0.51	0.565338
68	1.48	0.969018
69	1.56	0.963481

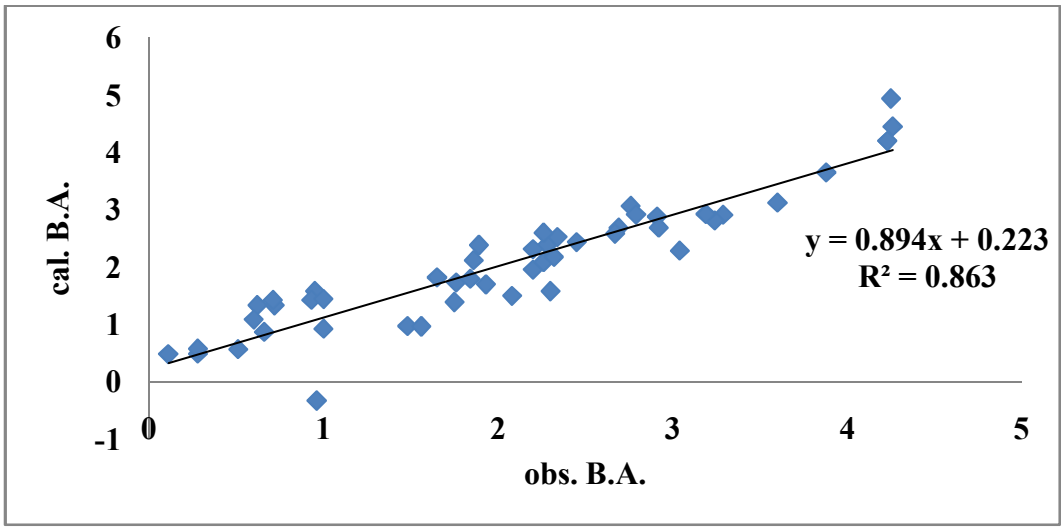


Figure 2: Experimental vs predicted biological activity (BA) of training set of Compounds by multiple linear regression model

Table 7: Predicted activities and residuals of test set from Model 4

Comp.	Obs. B.A.	Cal. B.A.
3	0.82	1.48661
4	0.89	1.08431
12	1.2	1.2675
13	0.97	0.727641
16	1.04	1.43553
17	0.92	1.42367
18	0.92	1.96214
31	1.49	1.44622
35	1.66	2.20894
40	3.64	3.15152
41	2.79	2.93156
47	3.17	2.81661
48	2.76	2.47532
52	2.24	2.26701
54	3.31	2.97799
59	2.21	2.24904

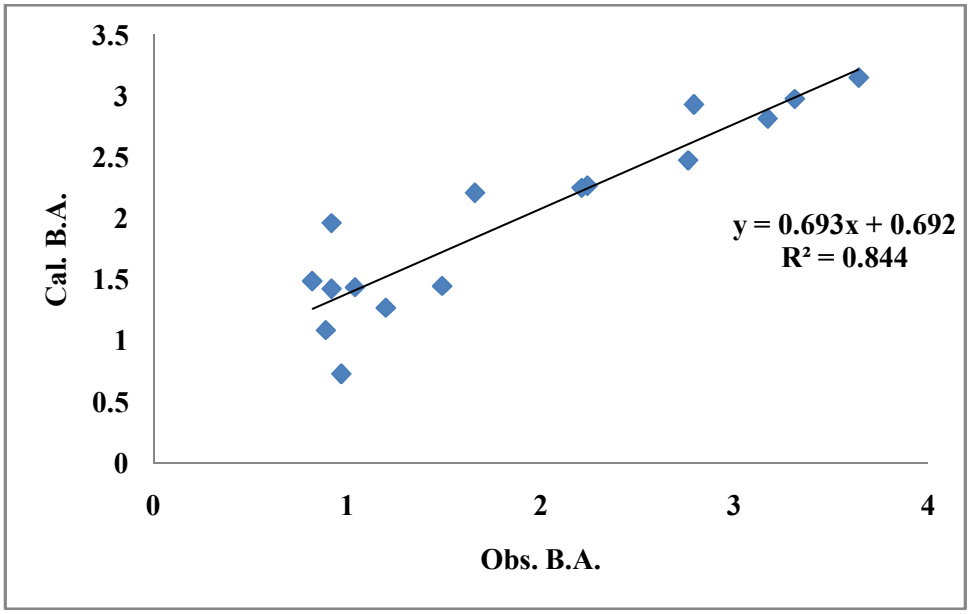


Figure 3: Experimental vs predicted biological activity (BA) of test set compounds by multiple linear regression model

The dataset are based on the NF54 strains of P. falciparum consisting of 69 organic compounds. The size of the final training set therefore became 53 compounds and test set 16 compounds. Descriptors were selected for the model based on their correlation coefficient and that descriptor having intercorrelation coefficient below 0.5 were considered. Various models were obtained after performing multiple linear regression (MLR) analysis. Model predictive power was judged based on various statistical parameters like correlation coefficient, regression coefficient (r^2), fisher statistical value (F), and standard error. The initial regression analysis was performed on all the 37 molecules which resulted in regression model. The best QSAR model has the characters of large F, low P value, r^2 and q^2 values close to 1, as well as $P< 0.001$. The best models observed in this QSAR study was found Model 4.

REFERENCE

1. Ben Mamoun C, Gluzman IY, Hott C, MacMillan SK, Amarakone AS, et al: Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite Plasmodium falciparum revealed by microarray analysis. Mol Microbiol 2001, 39(1):26-36.

2. Bozdech Z, Zhu J, Joachimiak MP, Cohen FE, Pulliam B, DeRisi JL: Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. Genome Biol 2003, 4(2):R9.

3. Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, et al:: A proteomic view of the Plasmodium falciparum life cycle. Nature 2002, 419(6906):520-526.

4. Le Roch KG, Johnson JR, Florens L, Zhou Y, Williamson KC, et al.: Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome Res 2004, 14(11):2308-2318.

5. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, et al:: Discovery of gene function by expression profiling of the malaria parasite life cycle. Science 2003, 301(5639):1503-1508.

6. Young JA, Fivelman QL, Blair PL, de la Vega P, Le Roch KG, Zhou Y, et al: The Plasmodium falciparum sexual development transcriptome: a microarray analysis using ontology-based pattern identification. Mol Biochem Parasitol 2005, 143(1):67-79.

7. Gille C, Goede A, Schloetelburg C, Preissner R, Kloetzel PM, Gobel UB, Frommel C: A comprehensive view on proteasomal sequences: implications for the evolution of the proteasome. J Mol Biol 2003, 326(5):1437-1448.

8. Li GD, Li JL, Mugthin M, Ward SA: Molecular cloning of a gene encoding a 20S proteasome beta subunit from Plasmodium falciparum. Int J Parasitol 2000, 30(6):729-733.

9. Paugam A, Bulteau AL, Dupouy-Camet J, Creuzet C, Friguet B: Characterization and role of protozoan parasite proteasomes. Trends Parasitol 2003, 19(2):55-59.

10. Katritzky A R, Kulshyn O V, Stoyanova-Slavova I. et al. Antimalarial activity: A QSAR modeling using CODESSA PRO software. Bioorg. & Medi. Chem. 2006, 14, 2333-2357.