Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2009, 1(2): 133-144
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X

QSAR Studies of 4-Anilinoquinazolines with Antimalarial Activity: A Hansch Approaches

Aroop K. Saha, Nitish Gupta*

Department of Applied Chemistry, Shri G. S. Institute of Technology and Science, Indore, India

---

Abstract

In the present work, we report quantitative structure–activity relationship (QSAR) analysis of series of 57 6-ureido-4-anilinoquinazolines derivatives with antimalarial activity by applying linear free energy related (LFER) approach of Hansch. Multi Linear Regression analysis is used to correlate the physicochemical properties and *in vitro* antimalarial activity. Two best QSAR models with 3 variables are selected on the basis of various significance statistical parameters. Among the various descriptors studied, standard Gibbs free energy (SGFE), LogP, Molecular refractivity (MR) and energy of highest occupied molecular orbital (HOMO) parameters has significance impact and show good correlation coefficients (r = 0.946 and 0.393) with antimalarial activity. The calculated F value of both models determines a confidence limit superior to 99% for this model.

Key Words: Antimalarial activity, Malaria, 6-ureido-4-anilinoquinazolines, *Plasmodium falciparum.*

Introduction

Despite years of continual effort, malaria is still one of the most infectious diseases in the world [1]. A major contributor to malarial morbidity and mortality is almost certainly the increasing resistance of malaria parasites to available drugs. Resistance is primarily seen in *P. falciparum*, the most virulent human malaria parasite [2]. The hard statistics about malaria are far from comforting, causing between 1-3 million deaths each year [3-4]. In tropical sub-Saharan Africa, the two groups at greatest risk of severe malaria morbidity and mortality are very young children

and pregnant women [5]. The spread of chloroquine-resistant *P. falciparum* strains has dashed hopes of global malaria eradication and due to the paucity of other affordable drugs, has complicated the clinical management of malaria in endemic areas. This reason has highlighted the need to identify alternative antimalarial compounds [6].On the other hand Quinazoline is an important scaffold due to variety of pharmacological properties associated with the derivatives bearing this heterocycle [7].

Malaria is caused by different species of plasmodium, namely *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae* and is transmitted to humans by female mosquitoes belonging to the genus Anopheles. Endemic maps indicate that *P. falciparum* and *P. vivax* account for 95% of malaria infections [8-9]. It is spread by sporozoa of the genus plasmodium, characterized by periodic fever, anemia and enlargement of the liver and spleen [10]. Despite continuous research efforts of decades, no effective vaccine is yet discovered for control of malaria. Treatment of malaria is becoming more difficult due to the spreading resistance of the parasite to standard antimalarial drugs, in particular to chloroquine, which had been the affordable and effective antimalarial prominent supporter for more than 50 years [11].

QSAR models are mathematical equations constructing a relationship between chemical structures and biological activities and one of the most important areas in chemometrics, give information that is useful for molecular design and medicinal chemistry [12-13]. For the development of QSAR model MLR analysis (using VALSTAT [14]) is subjected to the antimalarial activity against *P. falciparum*. MLR models revealed the relationship between the antimicrobial activity and structural descriptors. In the present paper, thermodynamic, steric and electronic descriptors calculated from ChemOffice 8.0.3 [15] were used, for the development of QSAR model. Quantitative structure activity relationship (QSAR) is a useful method for the design of bioactive compounds and the prediction of activity. The series and their antimalarial activity were based on the earlier reported work by Madapa et al. [7].

Results and Discussion

The statistical details (table 1) of the QSAR models given above express its good statistical quality. Among all possible multiple combinations of describing parameters best two tri-parametric QSAR models with the significance statistical parameters generated for the series is summarized below and the values of descriptor are given in table 4:

Model-1:

$pMIC_{50}$ = [6.5065(± 0.1652)] +SGFE [-0.0002 (± 0.0002)] +MR [-4.9407(± 1.017)] +HOMO [1.4821(± 3.5975)] ……(1)

n=41, r=0.946, r^2=0.894, variance=0.038, std=0.194, FIT=627.327, chance<0.001, r^2_{bs}=0.891

Model-2:

$pMIC_{50}$ = [6.3543(± 0.0903)] + LogP [-0.0003 (± 0.0008)] +MR [-4.5647 (± 1.0792)] +HOMO [1.4153(± 3.7699)] ……(2)

n=41, r=0.939, r^2=0.883, variance=0.042, std=0.205, F=92.805, FIT=556.828, chance<0.001, r^2_{bs}=0.052

Table 1: QSAR statistics of significant models

Eqn.No.	N Train	N Test	r	r^2	Q^2	r^2_{pred}	S_{DEP}	S_{PRESS}	F
1.	41	16	0.946	0.894	0.874	0.886	0.201	0.212	92.805
2.	41	16	0.939	0.883	0.866	0.899	0.208	0.219	104.554

N Train= number of training set, N Test= number of test set, r= coefficient of correlation, r^2= coefficient of determination, q^2= cross-validated squared correlation coefficient, r^2_{pred}= predicted coefficient of correlation, S_{DEP}= standard error of prediction, S_{PRESS}= predictive residual sum of square, F= Fisher's F-value.

The model shows that thermodynamic parameter such as MR, SGFE, LogP shows negative contribution and electronic parameter HOMO shows positive contribution to QSAR models. Suggest that a good percentage of the total variance in biological activity is accounted by the model Figure 1 and Figure 3 displays the scatter plot between observed and predicted activity for training and test set of model-1 respectively. For model-2 the scatter plot between observed and predicted activity display in Figure 2 and Figure 4 for training and test set of model-2 respectively. The F-statistic (compared to the critical value of 26.1 at the 0.01 level of significance). Furthermore, the calculated F value also determines a confidence limit superior to 99% for this model. The Inter-correlation matrix for model-1 (table 2) and model-2 (table 3) displays the correlation of various descriptors. The stability of model as judged by Loo method is fairly good (Q^2= 0.874 and 0.866) suggesting that the models will be useful for meaningful predictions.

All thermodynamic parameter is negatively contributed to models; MR indicates by the removal of bulky group will improve the biological activity, SGFE and LogP indicates the addition of liophobic substitution will increase the activity. Electronic parameter HOMO negatively contributed to the model, which suggests favorable electron-donating groups in the derivatives will increase the biological activity.

Table 2: Inter-correlation matrix for QSAR model-1

	SGFE	**MR**	**HOMO**
SGFE	1.000000		
MR	0.337394	1.000000	
HOMO	0.279279	0.489663	1.000000

Table 3: **Inter-correlation matrix for QSAR model-2**

	LogP	**MR**	**HOMO**
LogP	1.000000		
MR	0.338344	1.000000	
HOMO	0.235304	0.489663	1.000000

Table 4: Descriptors used in best QSAR models

Comp. No.	**Molar Refractivity (MR)**	**Standard Gibbs Free Energy (SGFE)**	**HOMOEnergy (HOMO)**	**LogP**
1	119.84	218.79	-8.65	6.15
2	119.60	198.44	-8.72	6.22
3	116.22	850.12	-8.46	4.09
4	120.88	596.44	-8.44	3.37
5	111.86	624.61	-8.48	5.17
6	113.84	576.80	-8.51	5.45
7	120.88	596.44	-8.53	3.37
8	114.79	955.12	-8.37	4.70
9	118.67	780.03	-8.52	5.78
10	127.35	490.23	-8.23	3.24
11	127.35	490.23	-8.50	3.24
12	120.48	1017.42	-8.23	6.52
13	119.46	701.44	-8.36	3.98
14	119.46	701.44	-8.37	3.98
15	120.88	596.44	-8.52	3.37
16	120.09	675.03	-8.46	5.17
17	124.08	140.04	-8.53	5.80
18	116.22	850.12	-8.28	4.09
19	120.09	675.03	-8.28	5.17
20	117.61	246.25	-8.87	5.93
21	120.72	1037.77	-8.18	6.45
22	115.73	373.53	-8.71	5.13
23	118.67	780.03	-8.48	5.78
24	118.10	722.84	-8.25	4.88
25	116.44	807.49	-8.51	5.57
26	116.68	827.84	-8.36	5.50
27	113.86	801.59	-8.33	5.23
28	116.44	807.49	-8.55	5.57

29	118.67	780.03	-8.38	5.78
30	108.82	802.80	-8.41	4.74
31	116.61	1192.51	-8.23	5.44
32	118.71	860.79	-8.34	6.39
33	120.48	1017.42	-8.26	6.52
34	109.97	751.89	-8.59	4.37
35	115.01	15.56	-8.87	5.82
36	114.56	934.77	-8.58	4.77
37	118.43	759.68	-8.57	5.85
38	120.33	695.38	-8.41	5.10
39	109.04	598.36	-8.47	4.90
40	114.08	597.15	-8.45	5.38
41	118.43	759.68	-8.53	5.85
42	113.63	781.24	-8.49	5.30
43	118.90	800.38	-8.32	5.71
44	124.57	616.63	-8.23	4.76
45	121.75	590.38	-8.22	4.49
46	122.68	743.91	-8.53	3.96
47	115.01	750.68	-8.61	4.86
48	113.84	576.80	-8.60	5.45
49	119.60	198.44	-8.71	6.22
50	111.64	829.05	-8.40	5.01
51	120.77	-362.80	-8.98	6.58
52	114.79	220.00	-8.86	5.66
53	115.28	696.59	-8.42	4.61
54	126.79	589.17	-8.21	4.97
55	126.55	568.82	-8.24	5.04
56	127.72	7.58	-8.56	5.41
57	120.33	695.38	-8.23	5.10

Table 5: Observed and Loo predicted Values for antimalarial activity (training set)

Comp. No.	Observed Activity ($pMIC_{50}$)	Loo Predicted Values for Model-1	Loo Predicted Values for Model-2
1	5.000	5.161	4.971
2	5.000	5.221	5.031
3	5.000	4.964	4.942
4	5.000	5.144	5.085
6	5.000	5.116	5.082

7	5.000	5.066	5.177
8	5.000	5.046	5.108
9	5.000	4.845	4.956
10	5.000	4.596	4.767
13	5.699	5.781	5.824
14	5.699	5.861	5.831
15	5.699	5.743	5.726
16	5.699	5.807	5.896
17	5.699	5.545	5.611
18	5.699	5.723	5.755
19	5.699	5.838	5.834
20	5.699	5.708	5.737
21	6.000	5.800	5.738
23	6.000	5.761	5.807
24	6.000	5.900	5.837
25	6.000	6.294	6.352
26	6.000	6.453	6.352
27	6.000	6.345	6.352
28	6.000	6.358	6.352
29	6.000	6.339	6.352
31	6.301	6.324	6.352
32	6.301	6.322	6.352
35	6.301	6.223	6.352
40	6.301	6.291	6.352
42	6.301	6.349	6.352
45	6.301	6.334	6.352
46	6.301	6.329	6.352
47	6.602	6.307	6.352
49	6.602	6.359	6.352
51	6.602	6.320	6.352
52	6.602	6.362	6.352
53	6.602	6.443	6.352
54	6.602	6.300	6.352
55	6.602	6.581	6.352
56	6.602	6.438	6.352
57	6.602	6.334	6.352

Table 6: Observed and Loo predicted Values for antimalarial activity (tset set)

Comp. No.	Observed Activity (pMIC50)	Loo predicted values for model-1	Loo predicted values for model-2
5	5.000	4.849	4.852
11	5.000	5.028	5.094
12	5.000	4.943	5.092
22	6.000	5.548	5.623
30	6.000	6.316	6.352
33	6.301	6.327	6.352
34	6.301	6.322	6.352
36	6.301	6.309	6.352
37	6.301	6.273	6.352
38	6.301	6.334	6.352
39	6.301	6.502	6.352
41	6.301	6.332	6.352
43	6.301	6.369	6.352
44	6.301	6.369	6.352
48	6.602	6.365	6.353
50	6.602	6.336	6.352

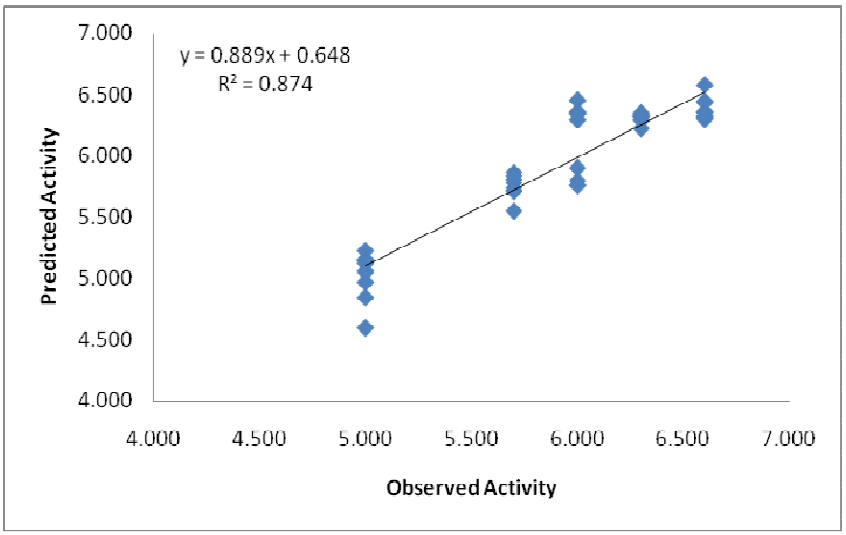


Figure 1: Scatter Plot between observed activity and predicted activity of Model 1 (training set)

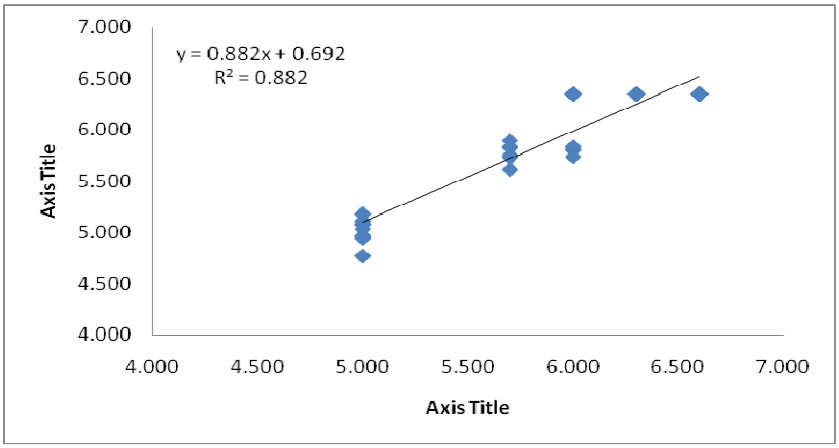


Figure 2: Scatter Plot between observed activity and predicted activity of Model 2 (training set)

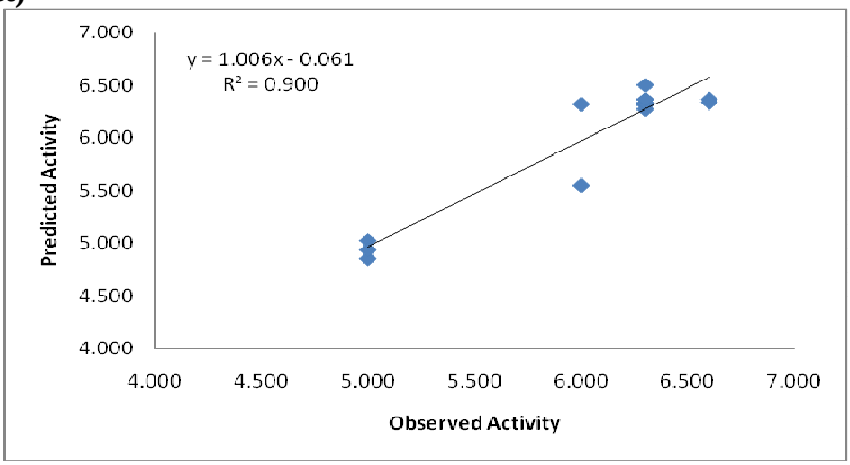


Figure 3: Scatter Plot between observed activity and predicted activity of Model 1 (test set)

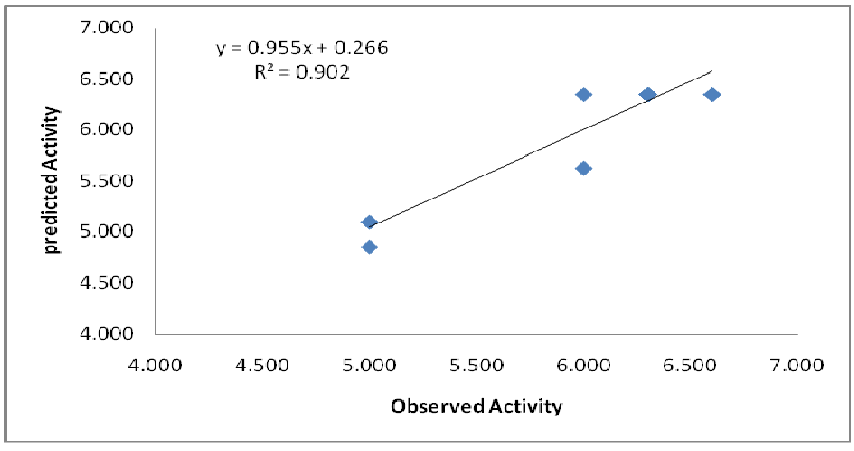


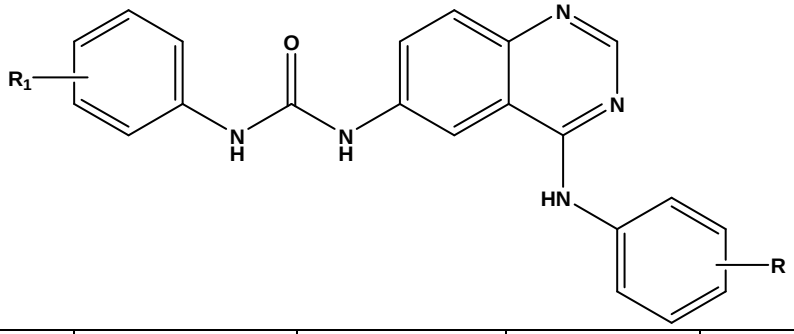
Figure 4: Scatter Plot between observed activity and predicted activity of Model 2 (test set)

Materials and Methods

Experimantal: *Data Set and Parameters*

In the present QSAR study, a series of 57 6-ureido-4-anilinoquinazoline derivatives with their in vitro antimalarial activities against Chloroquine sensitive 3D7 strain of *P. falciparum* taken from the reported work of Madapa et al. [7] (table 7). The biological activity data MIC (µg/mL) were converted to negative logarithmic dose in moles (pMIC) to obtain a linear relationship in the QSAR equation.

Table 7: Biological data and structure of the compound in the series



Comp. No.	R	R_1	MIC	pMIC
1	3-CF3	3-Cl-4-Me	10	5.000
2	3-CF3	3,5-Cl2	10	5.000
3	3-OMe	3-CN	10	5.000
4	3-OMe	3-COMe	10	5.000
5	4-F	4-Br	10	5.000
6	4-F	3,4-Cl2	10	5.000
7	4-OMe	3-COMe	10	5.000
8	4-Me	3-CN	10	5.000
9	3-Cl	3-Cl-4-Me	10	5.000
10	3,4-(OMe)2	3-COMe	10	5.000
11	3,4-(OMe)2	4-COMe	10	5.000
12	H	3,4-Cl2	10	5.000
13	4-Me	3-COMe	2	5.699
14	4-Me	4-COMe	2	5.699
15	4-OMe	4-COMe	2.0	5.699
16	3-OMe	3,4-Cl2	2	5.699
17	3-OMe	4-Br	2	5.699
18	4-OMe	3-CN	2	5.699
19	4-OMe	3,4-Cl2	2	5.699
20	3-CF3	4-Br	2.0	5.699
21	H	3-Cl-4-Me	1	6.000
22	3-CF3	3-CN	1	6.000
23	4-Cl	3-Cl-4-Me	1	6.000

24	4-OMe	4-Br	1	6.000
25	4-Cl	4-Br	1	6.000
26	4-Me	4-Br	1	6.000
27	4-Me	4-Cl	1	6.000
28	3-Cl	4-Br	1	6.000
29	4-Me	3,4-Cl2	1	6.000
30	H	4-Cl	1	6.000
31	H	3-CN	0.5	6.301
32	3-F	4-Br	0.5	6.301
33	3-F	4-Cl	0.5	6.301
34	3-F	3-CN	0.5	6.301
35	3-F	4-Cl-3-CF3	0.5	6.301
36	3-Cl	3-CN	0.5	6.301
37	3-Cl	3,4-Cl2	0.5	6.301
38	3-OMe	3-Cl-4-Me	0.5	6.301
39	4-F	4-Cl	0.5	6.301
40	4-F	3-Cl-4-Me	0.5	6.301
41	4-Cl	3,4-Cl2	0.5	6.301
42	4-Cl	4-Cl	0.5	6.301
43	4-Me	3-Cl-4-Me	0.5	6.301
44	3,4-(OMe)2	4-Br	0.5	6.301
45	3,4-(OMe)2	4-Cl	0.5	6.301
46	3,4-(OMe)2	3-CN	0.5	6.301
47	3-F	3-Cl-4-Me	0.25	6.602
48	3-F	3,4-Cl2	0.25	6.602
49	3-CF3	3,4-Cl2	0.25	6.602
50	H	4-Br	0.25	6.602
51	3-CF3	4-Cl-3-CF3	0.25	6.602
52	3-CF3	4-Cl	0.25	6.602
53	3-OMe	4-Cl	0.25	6.602
54	3,4-(OMe)2	3-Cl-4-Me	0.25	6.602
55	3,4-(OMe)2	3,4-Cl2	0.25	6.602
56	3,4-(OMe)2	4-Cl-3-CF3	0.25	6.602
57	4-OMe	3-Cl-4-Me	0.25	6.602

Cambridge software package [15] was used for the sketching of the series of molecules with the help of drawing tools of ChemDraw Ultra 8.0.3. The sketched 2D structures were transformed into 3D structures using module of the program (Chem3D Ultra 8.0). 3D structures were then subjected to energy minimization using molecular mechanics-2 (MM2) and re-optimization via MOPAC (Molecular Orbital Package) method until the RMS gradient attained a value smaller than 0.0001 kcal/mol Å. The descriptors selected for the study fell into three categories as electronic, steric and thermodynamic. The values of the molecular descriptors (table 8) were calculated using the Chem3D Ultra 8.0.3 module by using Compute properties option.

Table 8: Descriptors selected for the QSAR study

Thermodynamic descriptors	Bend energy (Eb), critical temperature (Tc), ideal gas thermal capacity (Cp), critical pressure (Pc), boiling point (BP), Henry's law constant (H), heat of formation (Hf), total energy (Et), LogP, partition coefficient (PC) and standard Gibbs free energy (SGFE).
Steric descriptors	Balaban index (Blndx), connolly accessible area (SAS), connolly molecular area (MS), connolly solvent excluded volume (CSEV), molar refractivity (MR), principal moment of inertia-X component (PMIX), principal moment of inertia-Y component (PMIY), principal moment of inertia-Z component (PMIZ), shape attribute(ShpA) and Ovality.
Electronic descriptors	Dipole (DPL), electronic energy (ElcE), highest occupied molecular orbital energy (HOMO), lowest unoccupied molecular orbital energy (LUMO), repulsion energy (NRE), VDW-1, 4-energy (E14) and Non-1, and 4-VDW energy (Ev).

Multiple Linear Regression Analysis and QSAR equation Validation

For the development of QSAR models, the compounds in the series were randomly divided into two sets of 41 and 16 compounds as training and test sets respectively. The Hansch type model used as QSAR model and for the statistical analysis the multiple linear regression analysis technique was used by applying VALSTAT [14]. In order to validate the QSAR models "Leave-one-out (Loo)" method [16] was used**.** Once a model was derived, each compound was eliminated from the remaining compounds and the eliminated compound was predicted from this model. The same procedure was repeated after elimination of another compound, until all the compounds had been eliminated once. The predictability of each model was evaluated by using cross-validated squared correlation coefficient *(Q^2)*, correlation coefficient (r), squared correlation coefficient (r^2), sequential Fischer test (F), standard deviation of prediction (S_{PRESS}), standard deviation of error of prediction (S_{DEP}) and boot-strapping square correlation coefficient (r^2_{bs}). The sum of the squared prediction errors, called the predictive residual sum of squares (press), is calculated as the sum of squares of the difference between the predicted and observed values of activity.

Hardware and Software

For computational modeling work and statistical analysis, Chem Office 2004 version 8.0.3 and VALSTAT was used respectively and run on a Windows XP based pentium IV 2.00 GHz core 2 duo processor PC (with 2 GB of memory).

Conclusion

In conclusion, we found that thermodynamic and electronic parameter of derivatives play key roles in the biological activity of the series of compounds. We built several highly predictive QSAR models having good predictive powers *(r^2_{pred}=0.886 and 0.889)*. The prediction power of the QSAR models tested by leave–one–out (Loo) method which gave a good predictive model of (q^2=0.874 and 0.866). The result of the Hansch approach suggests involvement of removal of bulky group substitution will increase the antimalarial activity. Additionally favorable electron-donating and hydrophilic group in the derivatives will increase the activity.

References

[1] J. G. Breman, *Am. J. Trop. Med. Hyg.*, **2001**, 64, 1-11.
[2] P. J. Rosenthal, *The J. of Exp. Bio.*, **2003,** 206, 3735-3744.
[3] N. J. White, F. Nosten, S. Looareesuwan, W. M. Watkins, K. Marsh, R.W. Snow, G. Kokwaro, J. Ouma, T. T. Hien, M. E. Molyneux, T. E. Taylor, C. I. Newbold, T. K. Ruebush, M. Danis, B. M. Greenwood, R. M. Anderson, P. Olliaro, *Lancet*, **1999**, 353, 1965–1967.
[4] B. M. Greenwood, D. A. Fidock, D. E. Kyle, S. H. I. Kappe, P. L. Alonso, F. H. Collins, P. E. Duffy, *J. Clin. InVest.*, **2008**, 118, 1266–1276.
[5] (WHO/CDS/MAL/2003.1093), *Geneva: World Health Organization*, **2003**.
[6] A. Mahajan, Y. Susan, N. Margo, E. J. Constance, C. Kelly, *Bioorg. & Med. Chem. Lett.*, **2007**, 17, 5683–5685.
[7] S. Madapa, Z. Tusi, A. Mishra, K. Srivastava, S. K. Pandey, R. Tripathi, S. K. Puri, S. Batra, *Bioorg. & Med. Chem.*, **2009,** 17, 222–234.
[8] S. B. Katiyar, K. Srivastava, S. K. Puri, P. M. S. Chauhana, *Bioorg. & Med. Chem.*, **2005,** 15, 4957–4960.
[9] S. K. Shrivastava, P. M. S. Chauhan, *Curr. Med. Chem.*, **2001**, 8, 1535–1542.
[10] N. J. Domingueza, E. J. Charrisa, G. Loboa, D. N. J. Domınguezb, M. M. Morenob, F. Riggionec, E. Sanchezd, J. Olsone, J. P. Rosenthale, *Eur. J. Med. Chem.*, **2001,** 36, 555–560.
[11] M. Foley, L. Tilley, *Int. J. Parasitol.*, **1997**, 27, 231.
[12] C. Hansch, A. Kurup, R. Garg, H. Gao, *Chem. Rev.*, **2001**, 101, 619-672.
[13] B. Hemmateenejad, R. Miri, M. Akhond, M. Shamsipur, *Arch. Pharm. Pharm. Med. Chem.*, **2002**, 10, 472-480.
[14] A. K. Gupta, B. M. Arockia, S. G. Kaslhedikar, *Indian J. Pharm. Sci.*, **2004**, 66(4), 396-402.
[15] CS Chem Office, Version ultra 8.0.3, Cambridge Soft Corporation, software Publishers Association, Washington D.C., 20036, 452-1600.
[16] H. Kubinyi, *Quant. Struct. Act. Relat.*, **1994**, 13, 285–294.