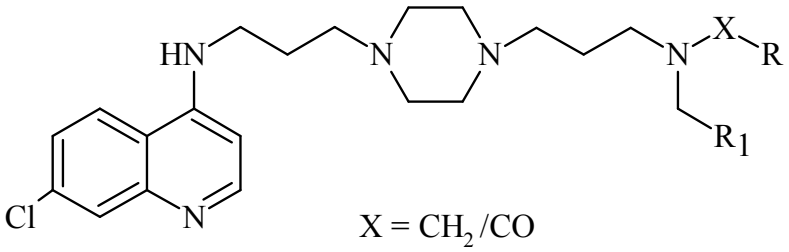
Journal of Enzyme Inhibition and Medicinal Chemistry
2009, 24,1,94-104

Graphical abstract

Topological Descriptors in Modeling the Antimalarial Activity: N^1-(7-Chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine as Prototype.

Shreekant Deshpande, V Raja Solomon, Setu B. Katti, Yenamandra S. Prabhakar*
Medicinal and Process Chemistry Division
Central Drug Research Institute, Lucknow 226 001, India



The QSAR of antimalarial activity of N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl) piperazine derivatives with different topological descriptors suggested that molecular modifications enhancing the radial centric information, topological distance between N and Cl atoms and vertex degree equality coupled with closely placed vertices, compact structural scaffolds will improve the activity.

Journal of Enzyme Inhibition and Medicinal Chemistry
2009, 24,1,94-104

Topological Descriptors in Modeling the Antimalarial Activity: N^1-(7-Chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine as Prototype.

Shreekant Deshpande, V Raja Solomon, Setu B. Katti, Yenamandra S. Prabhakar*
Medicinal and Process Chemistry Division
Central Drug Research Institute, Lucknow 226 001, India

Abstract
The antimalarial activity of *N*^{1}-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl) piperazine analogues is correlated to several descriptors from constitutional, topological, 2D autocorrelation, functional and empirical classes. Among them, the correlations with topological descriptors have favored the features of increased radial centric information, topological distance between N and Cl atoms and vertex degree equality coupled with closely placed vertices and compact structural scaffolds. The 2D autocorrelation descriptors suggested that one, two, six and eight lag spatial autocorrelations weighted by electronagativity, polarity and mass carry the information corresponding to the activity. The study proposes that while the hydrophobicity of the compound is favorable for the activity, the H-donor centers are not so preferred for the same. The PLS analysis of the identified descriptors also showed that the topological radial centric information, number of double bonds and hydrophobicity are the most influential ones for determining the antimalarial activity of these analogues.

Keywords: *N*^{1}-(7-Chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine derivatives, CP-MLR, Antimalarial activity, Topological descriptors, DRAGON software.
Abbreviation: CP-MLR - combinatorial protocol in multiple linear regression

* Corresponding author: phone:+91-522-2612411; Fax:+91-522-2623405; email: yenpra@yahoo.com

1. Introduction

Chloroquine (CQ) and other aminoquinolines (AQ) are the frontline chemotherapeutic agents for malaria because of their therapeutic efficacy and low cost [1]. These compounds enter the food vacuole of the parasite and inhibit its growth by forming complex with hematin thereby preventing the formation of hemozoin [2]. However, the emergence and proliferation of multidrug-resistant strains of *Plasmodium* species necessitated the development of alternative antimalarial agents [3]. The mechanistic investigations underlying the drug resistance have indicated that the resistance is a consequence of decreased accumulation of the drug in the food vacuole of the parasite owing to the enhanced efflux and reduced uptake [4]. To overcome the drug efflux mechanism of the parasite, Vennerstrom *et al* have investigated a series of bisquinolines **(Fig. 1)** against CQ-sensitive and CQ resistant strains [5]. Ryckebusch *et al* have further modified the bisquinolines by replacing one of the 7-chloro-4-aminoquinoline moieties with various amines and amides and prepared N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine derivatives **(Fig. 1)** for the evaluation of antimalarial activity against the chloroquine-resistant *P. falciparum* FcB1 strain [6]. Moreover, the studies with CQ/ AQ analogues have suggested that resistance mechanism does not involve any change in the target of this class of drugs but involves only compound specific resistance [7]. In this scenario, the structure activity relationship studies on 4-aminoquinoline class have suggested that 7-chloro-4-aminoquinoline nucleus as obligatory for the antimalarial activity, particularly for maintaining the pk_a of quinoline nitrogen, inhibition of hematin formation and for the accumulation of the drug at the target site [8]. Our synthetic explorations with modified 4-aminoquinolines also suggested the role of lipophilicity and pKa of the side chain on the antimalarial activity [9]. An early quantitative structure-activity relationship (QSAR) study of tebuquine (compounds related to amodiaquine) and other similar analogues discovered a correlation of the antimalarial activity with the decreasing size and electron donation property of the 4-anilino substituents of these molecules [10]. A recent QSAR on the antimalarial activity of 7-chloro-4-(3',5'-disubstituted anilino)quinolines proposed that the 4-anilino moiety of these compounds map diverse domains in the activity space and also prefer electron rich substituent groups in this substructure space for better antimalarial activity [11]. Topological features of the compounds are critical components in drug-target (receptor) interactions [12]. Therefore, discovering of these features is important for modifying the activity of the compounds [13]. Keeping this in view, the diverse chemical space of N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine derivatives **(Fig. 1)** [6] has been delineated in terms of topological descriptors to rationalize the antimalarial activity profile of these analogues.

2. Computation

Dataset. Several N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine analogues are reported in the recent literature along with their in vitro antimalarial activity (IC_{50}, inhibitory concentration in moles per liter against the chloroquine-resistant *P. falciparum* FcB1 strain) [6]. Among these analogues, one compound, cyclopropane carboxylic acid(3-[4-[3-(7-chloro-quinolin-4ylamino)-propyl]-piperazin-1-yl}-propyl)-cyclopropyl methyl amide, is embedded with a highly reactive cyclopropane carbonyl moiety. Due to this reason it has been kept out from the analogues selected for the QSAR study. The selected compounds are divided into two sets, as N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amide derivatives (briefly referred as amides; **Table 1**) and N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amine derivatives (briefly referred as amines; **Table 2**). For the purpose of QSAR study the

antimalarial activity of the compounds has been transformed into logarithm of reciprocal inhibitory concentration and expressed as $-\log IC_{50}$. The structure databases of these compounds were generated in ChemDraw [14]. In DRAGON software [15] these structure databases have resulted in 474 and 483 molecular descriptors for the compounds of **Tables 1** and **2**, respectively. They represent 0D- to 2D- characteristics of the molecular structures. According to the formalism of quantification of embedded structural information, all these indices belong to ten descriptor classes [15].

Model Development. The QSAR model generation and validation was carried out using CP-MLR (combinatorial protocol in multiple linear regression) [16] in conjunction with a three-stage descriptor classification protocol [11]. CP-MLR [16] is a 'filter' based variable selection procedure for the model identification and development in QSAR and QSPR studies [11, 16-18]. This involves a combinatorial strategy with appropriately placed 'filters' interfaced with MLR. It results in the extraction of diverse structure-activity models, each having unique combination of descriptors from the datasets under study. In this, the filters set the thresholds for the descriptors in terms of inter-parameter correlation cutoff limits in subset regressions (filter-1), t-values of the regression coefficients (filter-2), internal explanatory power (filter-3; square-root of adjusted multiple correlation coefficient of regression equation, r-bar) and the external consistency (filter-4; Q^2 i.e. cross-validated R^2 from the leave-one-out procedure).

The three-stage descriptor classification protocol [11] is implemented with the 3-descriptor combinations (baseline models) as they are the simplest ones obtained to explain the activity. In the 1st stage of classification protocol, the baseline models from the individual descriptor classes of the dataset were used to sort the descriptor classes into four categories. They are primary contributors (category I: a descriptor class forms model with its constituent descriptors), collective contributors (category II: a descriptor class unable to form model with its constituent descriptors, but forms model(s) in combination with a descriptor from another such descriptor class), secondary contributors (category III: a descriptor class forms model(s) only in combination with category I) and non-contributors (category IV: a descriptor class unable to form model(s) in any manner like category I, II or III; category IV are omitted from the study). The sorted descriptor classes were collated in the 2nd stage to identify all the 3-descriptor models across the categories. In the last stage, the individual descriptors of all 3-descriptor models were pooled to discover the higher models for the activity.

Throughout this study, for the filters-1, 2, and 4 of CP-MLR the thresholds were assigned as 0.3, 2.0, and $0.3 \leq Q^2 \leq 1.0$, respectively. The filter-3 was assigned an initial value of 0.71. In order to collect the descriptors with higher information content, the threshold of filter-3 was successively incremented with increasing number of descriptors (per equation) by considering the r-bar value of the preceding optimum model as the new threshold for next generation. With these provisions CP-MLR was used in conjunction with the three-stage descriptor classification protocol [11].

All the identified models were reassessed for the chance correlations, if any, by repeated randomization of the biological response [18-19]. For this each identified model was subjected to one hundred simulation runs with scrambled activity. The emerging regression equations with correlation coefficients better than or equal to the one corresponding to unscrambled response data were counted to express the percent chance correlation of the model under examination.

Additionally, the proposed models were verified by creating two divergent test sets, one emanating from the cluster analysis of the bit-packed version of the MACCS fingerprints (FP BIT MACCS) [20] of the compounds and the other from the random selection procedure, with each containing about one-third of the total compounds under analysis. As the total number of descriptors involved in this study is very large, only the descriptors significant to the models are addressed in the discussion. The complete descriptor file of each structure database is provided as supplementary material to the article.

3. Results and Discussion

In multi-descriptor class environment, exploring for models along the descriptor class provides scope to understand the phenomenon under investigation in relation to the concepts embedded in them. With this view, N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine analogues [6] **(Tables 1 and 2)** are analyzed using different 0D to 2D-descriptor classes from DRAGON software [15]. For these compounds all the descriptor classes have evolved as collective contributors (category II) to explain the antimalarial activity. At the end of a search for four parameter models, the following two are selected to explain the activity of compounds of **Table 1** (N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amide derivatives) from a pool of equations.

$$-\log IC50 = -2.654 - 0.952(0.123)nDB + 2.989(0.438)ICR + 0.012(0.003)T(N..Cl) - 0.236(0.063)nHDon$$

$$n = 33 \quad r = 0.900 \quad Q^2 = 0.707 \quad s = 0.254 \quad F = 29.97 \qquad (1)$$

$$-\log IC50 = 1.137 - 0.869(0.139)nDB + 3.311(0.503)ICR - 2.160(1.004)GATS1e - 3.062(1.293)GATS2p$$

$$n = 33 \quad r = 0.864 \quad Q^2 = 0.639 \quad s = 0.293 \quad F = 20.57 \qquad (2)$$

In this and all other regression equations, n is the number of compounds, r is the correlation coefficient, Q^2 is cross-validated R^2 from leave-one-out (LOO) procedure, s is the standard error of the estimate and F is the F-ratio between the variances of calculated and observed activities. The values given in the parentheses are the standard errors of the regression coefficients. In the randomization study, none of the identified models has shown any chance correlation. These are further validated through two test sets corresponding to the cluster analysis of the bit-packed version of the MACCS fingerprints (FP BIT MACCS) [20] of the compounds and the random selection procedure, with each one containing twelve out of thirty-three compounds of **Table 1**. The test sets predictions are in agreement with their experimental values **(Table 1; Fig. 2)**.

The Equations 1 and 2 have collectively shared six descriptors They are from four different descriptor classes namely, constitutional (nDB), topological (ICR and T(N..Cl)), 2D-autocorrelations (GATS1e and GATS2p), and functional groups (nHDon) [16]. In these equations, nDB represents the number of double bonds. Its negative regression coefficient suggests in favor of minimum number of double bonds in the compounds for improved activity. The topological parameter ICR is radial centric information index. It represents the mean information content derived from atom eccentricities. Its regression coefficient suggests in favor of increased ICR in the molecules for better activity. These two descriptors are common to equations 1 and 2. The other topological descriptor T(N..Cl) in equation 1 represents the sum of

topological distances between N and Cl atoms in the molecules. Its positive regression coefficient is in favor of increased separation between these two atoms. The functional group descriptor nHDon (equation 1) represents the number of donor atoms for H-bonds (with N and O). Its regression coefficient suggests in favor of minimum donor atoms for H-bonds for better activity. The 2D autocorrelation descriptors (equation 2), GATS1e is Geary autocorrelation of lag 1 weighted by atomic Sanderson electro negativities (e) and GATS2p is Geary autocorrelation of lag 2 weighted by atomic polarizabilities, (p). These two descriptors represent influence of spatial autocorrelations of specific (1 and 2) path lengths (lags) of molecular graphs weighted by atom physicochemical properties ('e' and 'p') on the activity. While GATS1e suggests in favor of decreasing one-centered fragments with electronegativity for the activity, GATS2p suggests in favor of decreasing polarity along two-centered fragments for the same.

The six descriptors of equations 1 and 2 are furthermore evaluated using the PLS (partial least squares) analysis to come out with a 'single window' structure-activity model for the activity. The PLS cross-validation procedure [21] has suggested that three components as optimum for these descriptors to explain variance in the activity. Equation 3 is MLR like PLS equation resulted from the three components of seven descriptors **(equations 1 and 2)**.

$$-\log IC_{50} = -1.132 - 0.939(0.343)nDB + 3.105(0.320)ICR + 0.010(0.135)T(N..Cl) - 0.977(0.049)GATS1e - 1.203(0.046)GATS2p - 0.153(0.106)nHDon$$

$$n = 33,\quad k = 6,\quad r = 0.902,\quad Q^2 = 0.763,\quad s = 0.247,\quad F = 42.33 \qquad (3)$$

In the PLS model, the values given in the parentheses following the regression coefficients are the fraction contributions of corresponding descriptors in explaining the variance in the activity of the compounds. In the statistics of PLS equations, k is the number of explanatory descriptors used in the PLS analysis. The remaining statistical terms (n, r, Q^2, s and F) associated with the equations represent the same information as defined earlier. All the six descriptors have influenced the activity between 4-34 per cent. Among them, nDB and ICR have shown maximum influence (each one more than 30%) on the activity. **Figure 3** shows a plot of the fraction contribution of normalized regression coefficients of these seven descriptors to the activity.

For the compounds listed in **Table 2** (N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amine derivatives), at the end of a search of all the descriptor classes the following two models are selected to explain the activity of these analogues.

$$-\log IC50 = -18.903 - 0.000013(0.000001)GMTIV + 26.870(7.696)MATS8m + 10.479(1.936)MATS6e - 0.536(0.116)Hy$$

$$n = 44\quad r = 0.854\quad Q^2 = 0.655\quad s = 0.264\quad F = 26.32 \qquad (4)$$

$$-\log IC50 = -25.672 - 0.019(0.002)S0K + 1.344(0.398)ICR + 1.710(0.435)IVDE + 28.020(7.646)MATS8m - 0.748(0.120)Hy$$

$$n = 44\quad r = 0.868\quad Q^2 = 0.659\quad s = 0.255\quad F = 23.23 \qquad (5)$$

Similar to the case of compounds of **Table 1**, equations 4 and 5 are also validated through two test sets with each one containing fourteen out of forty four compounds of **Table 2** and their

predictions are in agreement with their experimental values **(Table 2; Fig. 4)**. Collectively, seven descriptors have taken part in these two equations. They are from topological (GMTIV, S0K, ICR and IVDE), 2D-autocorrelations (MATS6e and MATS8m), and Empirical (Hy) descriptor classes.

Among the topological descriptors, ICR **(equation 5)** has shown its significance in modeling the activity of amine analogues (**Table 2**) as well. The other topological descriptors identified for these analogues are GMTIV (Gutman Molecular Topological Index by valence vertex degrees) **(equation 4)**, S0K (Kier symmetry index) and IVDE (mean information content on the vertex degree equality) **(equation 5)**. GMTIV is the product summation of valence vertex degree of all of atoms for all topological distances. Its negative regression coefficient suggests in favor of closely placed vertices and compact structural scaffolds for improved activity. The Kier symmetry index, S0K is for the zero order paths (atoms) and accounts for the molecular symmetry in terms of atom topological uniqueness. In equation 5 it is associated with a negative regression coefficient and points for structural features with reduced topological equivalence for better activity. IVDE is based on the partition of vertices according to vertex degree equality. It is a measure of the lack of structural homogeneity or the diversity of a molecule and related to Shannon's entropy measure. Its positive regression coefficient in equation 5 is in favor of increasing vertex degree equality for the activity. The 2D autocorrelation descriptors MATS6e and MATS8m (**equations 4** and **5**) are Moran autocorrelations of lags 6 and 8 / weighted by atomic Sanderson electronegativities (e) and atomic mass (m) respectively. Their positive regression coefficient suggests in favor of increased autocorrelation contents of six- and eight-member structural graphs weighted by ‘e’ and ‘m’ for the activity. In these equations Hy represents hydrophilic factor of the molecules. The negative coefficient of Hy suggests in favor of hydrophobicity in the compounds for improved activity. MATS8m and Hy are common descriptors to equations 4 and 5. Equation 6 is MLR like PLS equation for the activity from the three components of seven descriptors of equations 4 and 5.

$$\begin{aligned}-\log IC50 = {} & -18.964 - 0.000007(0.190)GMTIV - 0.008(0.192)S0K + 0.602(0.062)ICR \\ & + 0.555(0.049)IVDE + 24.677(0.128)MATS8m + 8.202(0.157)MATS6e \\ & - 0.654(0.221)Hy \\ n = 44, \quad k = 7, \quad r = 0.876, \quad & Q^2 = 0.715, \quad s = 0.241, \quad F = 44.03 \qquad (6)\end{aligned}$$

In this equation, Hy is one descriptor having the maximum influence (~22%) on the activity. It suggests in favor of hydrophobic compounds for better activity. In these compounds, ICR has accounted for about 6 per cent variance in the activity, whereas it accounted for about 32 per cent variance in the activity in case of amide analogues (**equation 3**). The remaining five descriptors have influenced the activity between 5 to 19 per cent. The fraction contributions of these seven descriptors to the activity are shown in **Figure 5**.

4. Conclusions

In conclusion, several descriptors from constitutional, topological, 2D autocorrelation, functional and empirical classes have contributed in modeling the antimalarial activity of N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine analogues **(Tables 1 and 2)**. In the models, the topological descriptors recommended for increased radial centric information (ICR), topological distance between N and Cl atoms (T(N..Cl)), closely placed vertices and compact structural

scaffolds (GMTIV), reduced topological equivalence (S0K) and increasing vertex degree equality (IVDE) for better the activity in the compounds. From 2D autocorrelation class, the one, two six and eight lag spatial autocorrelation descriptors weighted by electronagativity, polarity and mass are found to carry information corresponding to the activity. In amide analogues **(Table 1)**, the increasing unsaturation (nDB) is found to be unfavorable for the activity. While the hydrophobicity of the compound is favorable for the activity, the presence of H-donor groups in the molecule is not so preferred for the same. The PLS analysis carried out on the identified descriptors of these analogues (**Tables 1** and **2**) suggested that nDB, ICR and Hy are the most influential ones in modeling the antimalarial activity of these analogues.

5. Supplementary data. Complete dataset of molecular descriptors of structure databases corresponding Tables 1 and 2; PLS loadings, weights and sensitivity of independent and dependent descriptors of PLS models.

6. Acknowledgement
SD and VRS thank ICMR and CSIR, New Delhi, India, respectively, for the financial support in the form of Senior Research Fellowships. CDRI Communication No. 7231

7. References

[1] (a) R. G. Ridley, Nature. 415, (2002) 686-693. (b) P. M. O' Neill, P. G. Bray, S. R. Hawley, S. A. Ward, B. K. Park, Pharmacol. Ther. 77, (1998) 29-58.

[2] (a) A. V. Pandey, H. Bisht, V. K. Babbarwal, J. Srivastava, K. C. Pandey, V. S. Chauhan, Biochem. J. 355, (2001) 333-338. (b) A. C. Chou, R. Chevli, C. D. Fitch, Biochemistry. 19, (1980) 1543-1549. (c) T. J. Egan, H. M. Marques, Coord. Chem. Rev. 190-192, (1999**)** 493-517. (d) A. Dorn, R. Stoffel, H. Matile, A. Bubendorf, R. G. Ridley, Nature. 374 (1995, 269-271. (e) D. J. Sullivan, I. Y. Gluzman, D. G. Russell, D. E. Goldberg, Proc. Natl. Acad. Sci. U.S.A. 93, (1996) 11865-11870.

[3] J. Wiesner, R. Ortmann, H. Jomaa, M. Schlitzer, Angew. Chem. Int. Ed. 42, (2003) 5274-5293.

[4] (a) P.G. Bray, R.E. Howells, S.A. Ward, Biochem.Pharmacol. 43, (1992) 1219-1227. (b) P.G. Bray, R.E. Howells, G.Y. Ritchie, S.A. Ward, Biochem. Pharmacol. 44, (1992) 1317-1324.

[5] (a) J.L. Vennerstrom, W.Y. Ellis, A.L. Ager, S.L. Andersen, L. Gerena, W.K. Milhous, J. Med. Chem., 35, (1992) 2129-2134. (b) J.L. Vennerstrom, A.L. Ager, A. Dorn, S.L. Andersen, L. Gerena, R.G. Ridley, W.K. Milhous, J. Med. Chem., 41, (1998) 4360-4364.

[6] (a) A. Ryckebusch, R. Deprez-Poulain, L. Maes, M.A. Debreu-Fontaine, E. Mouray, P. Grellier, C. Sergheraert, J. Med. Chem., 46, (2003) 542–557. (b) A. Ryckebusch, M.A.; Debreu-Fontaine, E.; Mouray, P.; Grellier, C.;Sergheraert, P. Melnyk, Bioorg. Med. Chem. Lett., 15, (2005) 297–302.

[7] (a) R. G. Ridley, H. Hofheinz, H. Matile, C. Jaquet, A. Dorn, R. Masciadri, S. Jolidon, W. F. Richter, A. Guenzi, M. A. Girometta, H. Urwyler, W. Huber, S. Thiathong, W. Peters, Antimicrobial. Chemother. 40, (1996) 1846-1854. (b) D. De, F. M. Krogstad, L. D. Byers, D. J. Krogstad, J. Med. Chem. 41, (1998) 4918-4926.

[8] (a) T. J. Egan, R. Hunter, C. H. Kaschula, H. M. Marques, A. Misplon, J. C. Walden, J. Med. Chem. 43, (2000) 283-291. (b) C. H. Kaschula, T. J. Egan, R. Hunter, N. Basilico, S. Parapani, D. Tarameli, E. Pasini, D. Monti, J. Med. Chem. 45, (2002) 3531-3539.

[9] (a) V. R. Solomon, W. Haq, K, Srivastava, S. K. Puri, S. B. Katti, J. Med. Chem. 50, (2007) 394-398. (b) V. R Solomon, S. K. Puri, K, Srivastava, S. B. Katti, Bioorg. Med. Chem. 13, (2005) 2157–2165.

[10] L. M. Werbel, P. D. Cook, E. F. Elslager, J. H. Hung, J. L. Johnson, S. J. Kesten, D. J. McNamara, D. F. Ortwine, D. F. Worth, J. Med. Chem. 29, (1986) 924-939.

[11] M. K. Gupta, Y. S. Prabhakar, J. Chem. Inf. Model. 46, (2006) 93-102.

[12] (a) E, Estrada E. Uriarte, Curr Med Chem. 8(13), (2001) 1573-88. (b) JR. Votano, Curr Opin Drug Discov Devel. 8(1), (2005) 32-37. (c) R. Gozalbes, JP, Doucet, F. Derouin, Curr Drug Targets Infect Disord. 2(1), (2002) 93-102.

[13] (a) M.C. Bagchi, D. Mills, S. C. Basak, J. Mol. Model. 13, (2007) 111–120. (b) M.P. González, C. Terán, M. Teijeira, Bioorg. Med. Chem. Lett. 16 (5), (2006) 1291-1296. (c) M. K. Gupta, R. Sagar, A. K. Shaw, Y. S. Prabhakar, Bioorg. Med. Chem. 13 (2005) 343–351

[14] ChemDraw Ultra 6.0 and Chem 3D Ultra, Cambridge Soft Corporation, Cambridge, U.S.A.

[15] DRAGON software version 5.0-2005. By R. Todeschini, V. Consonni, A. Mauri, M. Pavan, Milano, Italy. http://disat.unimib.it/chm/Dragon.htm.

[16] Y. S. Prabhakar, QSAR Comb. Sci. 22, (2003) 583-595.

[17] Y. S. Prabhakar, Internet Electron. J. Mol. Des. *3*, (2004) 150-162. http://www.biochempress.com.

[18] Y. S. Prabhakar, V. R. Solomon, R. K. Rawal, M. K. Gupta, S. B. Katti, QSAR Comb. Sc. 23, (2004) 234-244.

[19] S. S. So, M. Karplus, J. Med. Chem. 40, (1997) 4347-4359.

[20] MOE: The Molecular Operating Environment from Chemical Computing Group Inc., 1255 University Street, Suite 1600, Montreal, Quebec, Canada H3B 3X3. (b) R. D. Brown, Y. C. Martin, *J*. Chem. Inf. Comput. Sci. 36, (1996) 572-584.

[21] S. Wold, Technometrics. 20, (1978) 397–405.

Figures captions.

Figure 1. General structures of (**a**) bisquinoline (**b**) N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amide derivatives and (**c**) N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amine derivatives as antimalarial agents against the chloroquine-resistant *P. falciparum* FcB1.

Figure 2. Plots of training (▲) and test (o) sets predicted activities versus observed activity corresponding to equations 1(**a**, **b**) and 2 (**c**, **d**). Number of compounds in training set is 21 and in test set is 12.

Figure 3. Plot of fraction contribution of MLR-like PLS coefficients (normalized) of the 6 descriptors from equations 1 and 2 to the activity. The serial numbers 1 to 6 on the horizontal axis refer to the descriptors nDB, ICR, T(n..Cl), GATS1e, GATS2p, and nHDon, respectively.

Figure 4. Plots of training (▲) and test (o) sets predicted activities versus observed activity corresponding to equations 4 (**a**, **b**) and 5 (**c**, **d**). Number of compounds in training set is 30 and in test set is 14.

Figure 5. Plot of fraction contribution of MLR-like PLS coefficients (normalized) of the 7 descriptors from equations 4 and 5 to the activity. The serial numbers 1 to 7 on the horizontal axis refer to the descriptors GMTIV, S0K, ICR, IVDE, MATS8m, MATS6e and Hy, respectively.

Table 1. Observed and modeled in vitro antimalarial activity of N^1-(7-chloro-4-quinolyl)1,4-bis(3-aminoprpyl)piperazine- amide derivatives (**Fig 1b**) against the FcB1R strain of *P. falciparum*

Comp. No	R	-logIC_{50}									
		Obs.[a]	Eq.1[b]	Set 1[c]	Set 2[c]	Eq.2	Set 1	Set 2	Eq.3	Set 1	Set 2
1	4-quinolinyl	7.70	7.32	7.41	*7.29*	7.08	7.11	*7.09*	7.32	7.35	*6.99*
2	1-naphthyl	6.81	7.12	7.18	7.04	7.18	7.25	7.17	7.10	7.13	6.98
3	Phenyl	7.09	7.21	*7.22*	*7.11*	7.15	*7.17*	*7.10*	7.16	*7.06*	*7.05*
4	3-thiophenyl	6.95	6.93	6.99	6.80	6.90	6.95	6.84	6.90	6.92	6.81
5	3-phenoxyphenyl	7.68	7.96	7.90	7.98	7.82	7.80	7.77	7.83	7.78	7.51
6	4-chlorophenyl	7.78	8.44	*8.38*	8.57	7.30	*7.42*	7.13	7.90	*7.84*	7.77
7	4-methoxyphenyl	7.70	7.56	7.54	*7.53*	7.66	7.73	*7.61*	7.60	7.69	*7.59*
8	4-fluorophenyl	7.31	7.11	7.14	7.01	7.31	7.26	7.27	7.20	7.32	7.21
9	4-nitrophenyl	7.91	7.88	8.06	7.83	8.34	7.94	8.37	8.03	8.07	8.02
10	4-hydroxyphenyl	6.38	6.92	*6.93*	6.88	6.95	*6.96*	6.94	6.87	*7.09*	6.72
11	Benzyl	7.17	6.96	6.98	*6.93*	7.10	7.10	*7.06*	7.01	6.98	*6.96*
12	phenethyl	7.09	7.43	7.42	7.42	7.54	7.52	7.52	7.40	7.38	7.31
13	cyclohexyl	6.67	6.97	7.01	6.93	6.90	6.90	6.89	6.90	6.85	6.86
14	cyclopropyl	6.54	6.66	*6.71*	6.59	6.64	*6.63*	6.61	6.62	*6.85*	6.67
15	Hexyl	7.46	7.43	7.39	*7.41*	7.31	7.22	*7.31*	7.37	7.21	*7.21*

16	Propyl	7.13	6.60	6.63	6.49	6.61	6.58	6.54	6.62	6.66	6.62
17	Ethyl	6.56	6.64	*6.69*	6.56	6.72	*6.74*	6.69	6.64	*6.73*	6.59
18	methyl	6.40	6.31	6.42	*6.23*	6.45	6.54	*6.40*	6.35	6.47	*6.28*
19	*tert*-butyl	6.79	6.61	6.67	*6.56*	6.89	6.98	*6.83*	6.71	6.64	*6.59*
20	isopropyl	6.83	6.64	*6.71*	*6.58*	6.77	*6.81*	*6.73*	6.68	*6.74*	*6.63*
21	*N*-Boc-Gly	6.41	6.71	*6.53*	6.86	6.71	*6.47*	6.87	6.66	*6.79*	6.63
22	*N*-Boc-L-Pro	6.37	6.57	6.48	*6.60*	6.59	6.35	*6.64*	6.56	6.37	*6.27*
23	*N*-Boc-(R)-1,2,3,4-tetrahydro-3-isoquinolinyl	6.41	6.44	*6.34*	6.53	6.62	*6.41*	6.80	6.50	*6.53*	6.28
24	*N*-Boc-(S)-1,2,3,4-tetrahydro-3-isoquinolinyl	6.40	6.45	6.28	*6.50*	6.62	6.42	*6.66*	6.50	6.53	*6.60*
25	Gly	6.50	6.38	6.64	*6.41*	6.54	6.64	*6.50*	6.46	6.70	*6.87*
26	L-Pro	6.44	6.75	6.99	6.77	6.57	6.62	6.54	6.66	6.77	6.80
27	(R)-1,2,3,4-tetrahydro-3-isoquinolinyl	7.72	7.23	7.33	7.18	7.37	7.39	7.29	7.37	7.43	7.47
28	(S)-1,2,3,4-tetrahydro-3-isoquinolinyl	7.56	7.27	7.38	7.25	7.39	7.42	7.33	7.37	7.43	7.47
29	1,2,3,4-tetrahydro-3-isoquinolinyl	7.42	7.30	7.28	7.26	7.48	7.46	7.43	7.36	7.40	7.34
30[d]	1-propanol carbonyl	6.76	6.77	*6.81*	6.70	6.59	*6.61*	6.53	6.68	*6.64*	6.35
31[d]	1-propanoic acid carbonyl	5.86	5.86	*5.68*	5.90	5.92	*5.67*	6.10	5.84	*6.49*	5.96
32[d]	1-propanamide carbonyl	ND[e]	5.90			5.90			5.90		
33[d]	N-Boc ethyl amine carbonyl	7.47	6.82	*6.83*	6.78	6.43	*6.41*	6.41	6.85	*7.25*	7.40
34[d]	ethyl amine carbonyl	6.40	6.59	*6.74*	*6.52*	6.35	*6.39*	*6.32*	6.49	*6.56*	*6.54*
	Training r^2			0.82	0.78		0.80	0.70		0.81	0.84
	Test r^2			0.73	0.83		0.66	0.81		0.57	0.74

[a], Observed activity; ref. [6a] for compounds 1-29 and ref. [6b] for compounds 30-34.
[b], For all equations 'set 1' corresponds to MACCS cluster an 'set 2' corresponds to random selection; unless otherwise stated, all predicted activities are from LOO cross-validation.
[c], Test set compounds are identified in italics and their activities are predicted with the corresponding model developed using the remaining compounds as the training group.
[d], For Compounds 30-34, amide 'H' is replaced by methylene cyclopropane.
[e], ND - not determined, see ref. [6b].

Table 2. Observed and modeled in vitro antimalarial activity of N^1-(7-chloro-4-quinolyl)1,4-bis(3-aminoprpyl)piperazine-amine derivatives (**Fig 1c**) against the FcB1R strain of *P. falciparum*

Comp. No[a]	R	-logIC$_{50}$									
		Obs[b]	Eq.4[c]	Set 1[d]	Set 2[d]	Eq.5	Set 1	Set 2	Eq.6	Set 1	Set 2
35	4-quinolinyl	7.13	7.39	7.42	*7.38*	7.44	7.43	*7.45*	7.31	7.34	*7.35*
36	1-naphthyl	7.58	7.36	*7.38*	7.35	7.45	*7.42*	7.48	7.34	*7.35*	7.36
37	Phenyl	7.78	7.57	7.58	7.52	7.78	7.80	7.73	7.62	7.58	7.58
38	3-thiophenyl	7.94	7.61	7.62	*7.58*	7.56	7.55	*7.53*	7.56	7.61	*7.51*
39	3-phenoxyphenyl	7.70	7.12	*7.18*	7.14	7.54	*7.54*	7.58	7.28	*7.11*	7.36
40	4-chlorophenyl	7.78	7.97	8.06	*7.91*	7.86	7.84	*7.85*	7.91	7.96	*7.91*
41	4-methoxyphenyl	7.71	7.96	8.05	7.93	7.99	7.99	8.01	7.98	7.93	7.98
42	4-fluorophenyl	7.73	7.90	*7.95*	7.87	7.86	*7.83*	7.86	7.88	*7.94*	7.88
43	4-nitrophenyl	8.23	7.78	7.82	*7.81*	7.98	7.92	*8.04*	7.91	7.87	*7.95*
44	4-hydroxyphenyl	7.26	7.56	*7.61*	7.50	7.36	*7.30*	7.33	7.43	*7.55*	7.40
45	Benzyl	7.73	7.42	7.41	7.37	7.58	7.56	7.53	7.46	7.41	7.42
46	Phenethyl	7.59	7.49	*7.51*	*7.45*	7.67	*7.65*	*7.61*	7.57	*7.43*	*7.53*
47	Cyclohexyl	7.37	7.57	7.60	7.51	7.59	7.65	7.48	7.56	7.51	7.46
48	Cyclopropyl	7.99	7.70	7.72	7.61	7.48	7.48	7.33	7.65	7.69	7.54
49	Hexyl	8.02	7.60	*7.64*	*7.54*	7.76	*7.81*	*7.70*	7.72	*7.55*	*7.63*
50	Propyl	7.61	7.69	7.72	7.59	7.76	7.76	7.65	7.70	7.71	7.61
51	Ethyl	7.44	7.75	*7.76*	7.64	7.82	*7.78*	7.71	7.75	*7.79*	7.65
52	Methyl	7.24	7.67	7.72	7.56	7.75	7.77	7.64	7.64	7.74	7.52
53	*tert*-butyl	8.01	7.85	7.89	7.75	8.11	8.06	8.04	7.93	7.98	7.88
54	Isopropyl	7.73	7.79	7.83	7.68	7.97	7.95	7.88	7.84	7.87	7.76
55	4-quinolinyl	6.85	6.93	*6.84*	7.12	6.65	*6.55*	6.90	6.80	*6.85*	6.94
56	1-naphthyl	6.67	6.84	6.71	7.04	6.87	6.70	7.14	6.74	6.74	6.88
57	Phenyl	7.26	7.58	7.55	*7.59*	7.71	7.71	*7.71*	7.56	7.53	*7.60*
58	3-thiophenyl	7.20	7.45	7.39	7.46	7.36	7.30	7.38	7.34	7.33	7.34
59	3-phenoxyphenyl	6.58	6.67	6.53	*6.85*	6.86	6.69	*7.02*	6.73	6.61	*6.96*
60	4-chlorophenyl	ND[e]	7.90			7.59			7.82		
61	4-methoxyphenyl	8.02	7.73	7.69	7.71	7.66	7.49	7.75	7.83	7.77	7.92
62	4-fluorophenyl	ND[e]	7.74			7.68			7.76		
63	4-nitrophenyl	7.34	7.53	7.54	*7.60*	7.38	7.28	*7.52*	7.53	7.55	*7.66*
64	4-hydroxyphenyl	7.34	7.06	*7.20*	6.99	6.61	*6.67*	6.66	7.01	*7.17*	7.06
65	Benzyl	7.58	7.28	7.19	7.33	7.51	7.46	7.56	7.35	7.28	7.40
66	Phenethyl	7.40	7.44	*7.37*	7.48	7.36	*7.29*	7.41	7.47	*7.33*	7.50
67	Cyclohexyl	7.80	7.75	7.71	*7.77*	7.64	7.69	*7.63*	7.70	7.68	*7.67*
68	Cyclopropyl	8.06	8.56	*8.52*	8.55	8.10	*8.17*	8.03	8.36	*8.49*	8.29
69	Hexyl	7.42	7.75	7.71	7.71	7.88	7.98	7.89	7.80	7.58	7.75
70	propyl	8.13	8.09	8.09	8.05	8.28	8.31	8.23	8.14	8.14	8.11
71	Ethyl	8.66	8.31	*8.36*	*8.27*	8.28	*8.33*	8.66	8.37	*8.40*	*8.31*
72	methyl	8.30	7.93	7.84	7.75	7.94	7.79	7.78	8.05	8.06	7.96
73	*tert*-butyl	ND[e]	9.56			9.59			9.51		

74	isopropyl	9.05	8.80	8.93	8.70	8.86	8.97	8.83	8.91	9.07	8.93
75	hydrogen	7.99	7.72	7.75	*7.65*	7.56	7.57	*7.49*	7.68	7.75	*7.57*
76	1-pentane nitrile	8.00	8.23	8.25	8.19	7.78	7.78	7.75	8.11	8.11	8.07
77	N-Boc propyl amine carbonyl	7.90	7.89	7.89	*7.89*	7.97	7.88	*8.01*	7.99	7.84	*8.04*
78	propyl amine carbonyl	6.82	7.40	*7.38*	7.35	7.18	*7.08*	7.12	7.16	*7.32*	7.09
79	pyrrolidinyl	7.71	7.88	7.85	7.80	7.84	7.86	7.73	7.86	7.89	7.75
80	piperidinyl	7.94	7.87	7.83	*7.80*	7.79	7.79	*7.72*	7.87	7.86	*7.76*
81	azepaninyl	7.87	7.87	*7.84*	7.78	7.71	*7.72*	7.57	7.85	*7.83*	7.74
	Training r^2			0.79	0.73		0.76	0.77		0.79	0.79
	Test r^2			0.57	0.67		0.69	0.62		0.53	0.65

[a], for compounds 55-74, R = R_1; for compounds 75-81, R_1 is cyclopropyl methyl.
[b], observed activity; ref. [6a] for compounds 35-74 and ref. [6b] for compounds 75-81.
[c –d], see the corresponding footnote 'b' and 'c' in Table 1.
[e], ND - not determined, see ref.[6a].

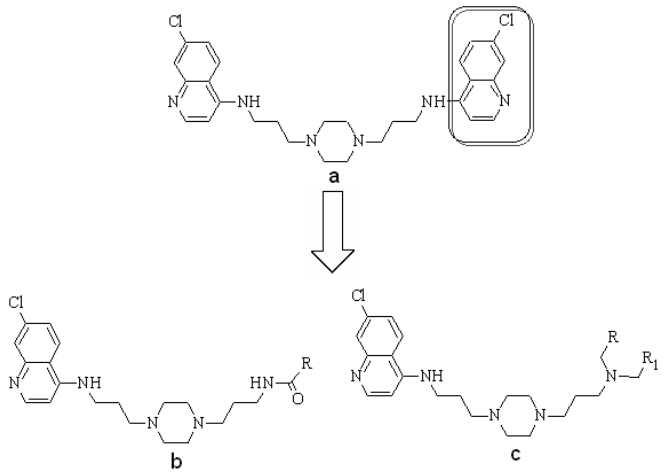


Figure 1. General structures of (**a**) bisquinoline (**b**) N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amide derivatives and (**c**) N^1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine amine derivatives as antimalarial agents against the chloroquine-resistant *P. falciparum* FcB1.

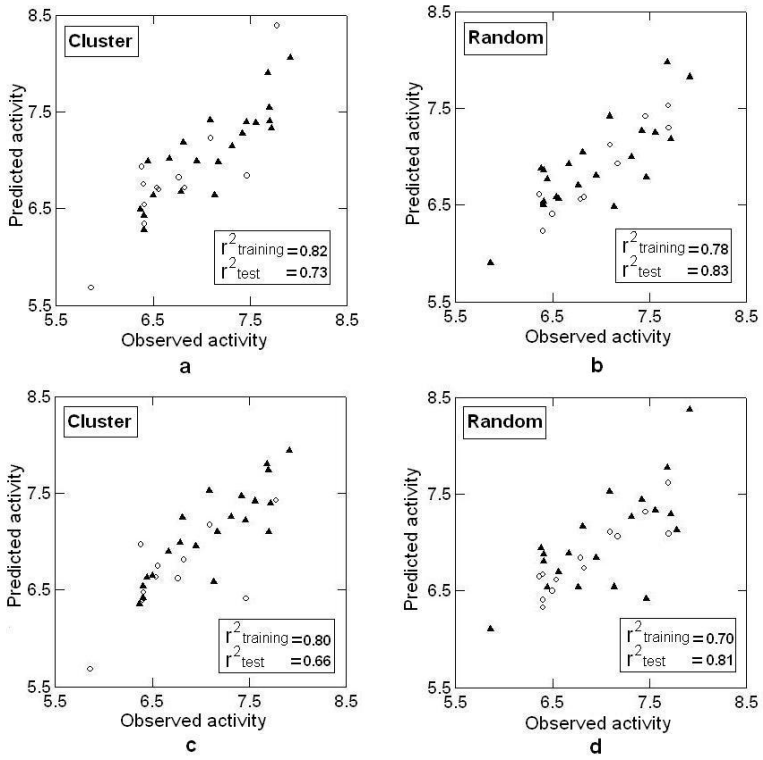


Figure 2. Plots of training (▲) and test (o) sets predicted activities versus observed activity corresponding to equations 1(**a**, **b**) and 2 (**c**, **d**). Number of compounds in training set is 21 and in test set is 12.

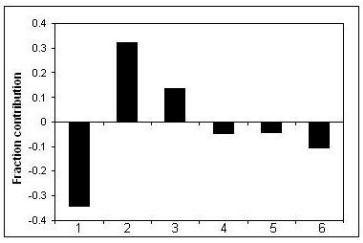


Figure 3. Plot of fraction contribution of MLR-like PLS coefficients (normalized) of the 6 descriptors from equations 1 and 2 to the activity. The serial numbers 1 to 6 on the horizontal axis refer to the descriptors nDB, ICR, T(n..Cl), GATS1e, GATS2p, and nHDon, respectively.

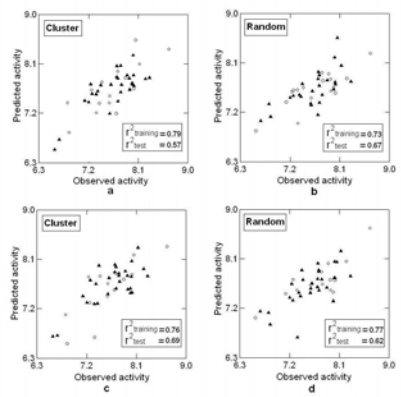


Figure 4. Plots of training (▲) and test (o) sets predicted activities versus observed activity corresponding to equations 4 (**a**, **b**) and 5 (**c**, **d**). Number of compounds in training set is 30 and in test set is 14.

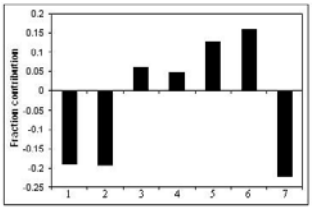


Figure 5. Plot of fraction contribution of MLR-like PLS coefficients (normalized) of the 7 descriptors from equations 4 and 5 to the activity. The serial numbers 1 to 7 on the horizontal axis refer to the descriptors GMTIV, S0K, ICR, IVDE, MATS8m, MATS6e and Hy, respectively.