Bioorganic & Medicinal Chemistry Letters 20 (2010) 4021–4026

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Exploring structural requirements of 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines as antiamoebic agents using comparative QSAR modelling

Nilanjan Adhikari, Milan Kumar Maiti, Tarun Jha *

Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, PO Box 17020, Jadavpur University, Kolkata 700 032, India

ARTICLE INFO

Article history:
Received 21 September 2009
Revised 6 January 2010
Accepted 26 May 2010
Available online 1 June 2010

Keywords:
Amoebiasis
QSAR modelling
PCRA
Stepwise regression
FA-MLR
PLS

ABSTRACT

Amoebiasis is a potentially lethal disease and causes 70,000 deaths per year. To find structural requirements for more active antiamoebic agents than metronidazole, comparative QSAR modelling was done on thirty 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines. The best model was obtained by using PLS technique with R^2_A and R^2_{CV} value of 88.50% and 82.90%, respectively. Amoebicidal activity may increase when Wang–Ford charges at atom numbers 6 and 12 have large positive values. Number of six-membered ring and sum of Kier–Hall electrotopological states may also increase amoebicidal activity when these have large positive values. Increasing value of rotatable bond fraction, approximate surface area and mean atomic polarizability scaled on carbon atom may be detrimental for antiamoebic activity. Decrease in values of electrostatic potential charges at atom numbers 1 and 12 may be conducive for activity. Electrophilic attacks may be favourable at these positions.

Amoebiasis is a common infection of human G. I. tract. It is a potentially lethal disease. 45 million people are carrier of *Entamoeba histolytica* in their intestinal tract. One-tenth of them gradually suffer from invasive amoebiasis. This includes liver abscesses, problems in lungs, spleen, brain and kidney. It causes 70,000 deaths per year.[1] These protozoa can be differentiated into at least 18 zymodems. Zymodems are populations of organisms differing from similar populations by one or more isoenzymes. Pathogenic strains of *E. histolytica* are derived from particular zymodems. These zymodems belongs to seven potentially pathogenic zymodems.[1] Due to a chitin wall, *E. histolytica* is resistant to gastric acid. The quadrinucleated cysts of *E. histolytica* release trophozoites under anaerobic conditions without harming the host.[2] Trophozoites adhere to colonic epithelial cells by the membrane lectin. It has similarity to the host adherence proteins. Trophozoites then lyse the host cell and invade through submucosa by releasing a factor. This factor inhibits IFN_γ activated macrophages. These processes result in dysentery. These parasites may invade the liver and can cause liver abscesses and amoebic granuloma.[3] Amoebic proteins may be associated with the tissue invasion by: (a) cystein proteinases that can breakdown extracellular matrix; (b) lectin, a parasite surface protein, can bind to carbohydrates on the colonic epithelial cells; (c) amebapore, a channel-forming protein, creates hole in the plasma membrane of the host cell and lyses these.[2] Metronidazole is the best drug for treating amoebiasis. It blocks ferridoxin-dependent pyruvate oxidoreductase which is present only in those organisms but absent in human beings.[2] Again, it is noticeable that it kills trophozoites without infecting cysts. Hence, it is effective against invasive amoebiasis in the intestine and the liver. It is less effective against these organisms in lumen or gut.[3] Thus, drugs for better antiamoebic activity is still in search in the field of medicinal chemistry. To find the structural requirements for more active amoebicidal agents, comparative QSAR study was performed on some 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines. This is done as a part of our composite programme of rational drug design, discovery and development.[4] The general structure of these compounds with arbitrary numbering is shown in Figure 1.

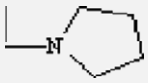
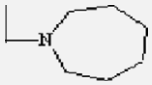
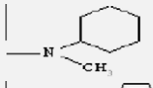
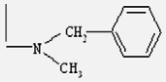
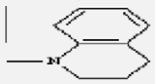
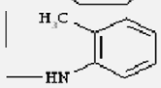
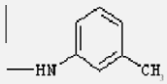
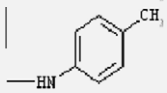
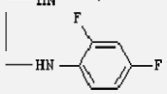
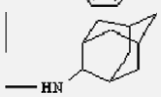
In vitro antiamoebic activities (IC_{50}) of thirty 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines against (*HM1:IMSS)* strain of *E. histolytica* were collected.[5] Here, negative logarithms of inhibitory activity of these compounds (pIC_{50}) were used to develop QSAR models to obtain linear relationship. pIC_{50} values of these compounds are shown in Table 1.

Atomic charge is the difference between the charge on the core and the electron density of the atom. Wang–Ford charges (*Q*) and electrostatic potential charges (EP) were calculated by using CHEM 3D PRO package.[6] According to Austin Model 1 (AM1) method[7], energy minimizations of these structures were done under Molecular Orbital Package (MOPAC) module. It is done by using restricted Hartree–Fock closed shell (RHF) wave function. Wang–Ford charges (*Q*) and electrostatic potential charges (EP) were calculated from these energy minimised geometry. Different quantum chemical descriptors like molecular refractivity (MR), approximate

* Corresponding author. Tel.: +91 33 24146666x2495 (o); fax: +91 33 24146927.
E-mail address: tjupharm@yahoo.com (T. Jha).

0960-894X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2010.05.098

Table 1
In vitro antiamoebic activities of 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines

Compd[a]	X	R	IC_{50} (μM)	pIC_{50} (M)
1[b]	H		3.730	5.428
2	Br		2.820	5.550
3	Cl		2.310	5.636
4	H		4.390	5.358
5	Br		1.090	5.963
6	Cl		0.890	6.051
7	H		5.940	5.226
8	Br		5.250	5.280
9[b]	Cl		3.710	5.431
10	H		7.310	5.136
11[b]	Br		4.440	5.353
12	Cl		2.910	5.536
13	H		6.190	5.208
14[b]	Br		2.780	5.556
15	Cl		1.770	5.752
16[b]	H		4.780	5.321
17	Br		3.820	5.418
18	Cl		1.680	5.775
19	H		5.010	5.300
20	Br		3.350	5.475
21	Cl		2.810	5.551
22	H		9.560	5.020
23	Br		5.330	5.273
24[b]	Cl		2.400	5.620
25	H		1.760	5.754
26	Br		0.670	6.174
27	Cl		0.510	6.292
28	H		1.790	5.747
29	Br		0.580	6.237
30	Cl		0.470	6.328

[a] Compound number.
[b] Compounds designed for test set.

surface area (SAA), grid surface area (GSA), polarizability (Pol) and molecular volume (*V*) were calculated by using the software HYPERCHEM release 7.0 pro package.[8] By Hyperchem software, energy minimizations of 3D structures were done separately by using molecular mechanical (MM+) force field without cut off for non-bonded interactions, solvation and constrains. Energy minimized structures were subjected to geometrical optimisation by semiempirical Austin Model 1 (AM1) method[7] using Polak–Rebiere algorithm with a RMS gradient of 0.1 kcal/Å mol. By using Dragon software[9], different constitutional descriptors[10] were calculated. Statistical qualities of equations were justified by parameters like correlation coefficient (*R*), adjusted R^2 (R^2_A), variance ratio (*F*) at specified degrees of freedom, probability factor related to *F*-ratio (*p*) and standard error of estimate (SEE). Leave-one-out (LOO) cross validation method[11] was applied to validate QSAR models. Predicted residual sum of square (PRESS), total sum of squares (SSY), cross-validated R^2 (R^2_{CV}), standard deviation error of prediction (SDEP) and standard error of PRESS (S_{PRESS}) were considered for validation of predictive powers of these models. On the basis of developed QSAR models on the training set, activities of the test set compounds were predicted. R^2_{pred} values for the test set compounds were calculated.

Principle component regression analysis (PCRA).[12] By principle component method, six factor scores having factor loading greater than 0.70 were extracted with VARIMAX rotation. These factor scores were used as independent parameters for developing QSAR equations. As factor scores contain information for the different descriptors, the chance of loss of information is less. Using forward selection method, the following equation was developed:

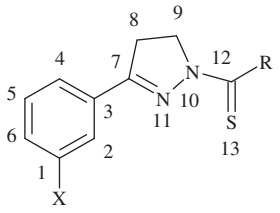


Figure 1. General structure of 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines with arbitrary numbering.

$$pIC_{50} = 5.592(\pm0.030) + 0.195(\pm0.031)f_4 + 0.169(\pm0.031)f_1 - 0.149(\pm0.031)f_6 + 0.098(\pm0.031)f_3 + 0.068(\pm0.031)f_2 \quad (1)$$

$n = 30;\ R = 0.906;\ R^2 = 0.821;\ R^2_A = 0.783;\ F(5, 24) = 21.947;$
$p < 0.00000;\ SEE = 0.165;\ SSY = 3.638;\ PRESS = 0.974;$
$R^2_{CV} = 0.732;\ SDEP = 0.180;\ S_{PRESS} = 0.201.$

where *n* is the number of data points. Eq. 1 explains 78.30% and predicts 73.20% variances of antiamoebic activity. Eq. 1 shows importance of factors 4, 1, 6, 3 and 2. Factor 4 is highly loaded with Q_{10}, Q_{12}, EP_{10}, EP_{12} and EP_{13}. It shows importance of these descriptors. Factor 1 is highly loaded with Q_1, Q_2, Q_4, Q_5, Q_6, EP_1, EP_2, EP_3, EP_4, EP_5, and EP_6. Q_1, Q_2, Q_4, Q_5, Q_6, Q_7, Q_8, Q_9, Q_{10}, Q_{11}, Q_{12} and Q_{13} stand for Wang–Ford charges at atom numbers 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13, respectively. EP_1, EP_2, EP_3, EP_4, EP_5, EP_6, EP_7, EP_8, EP_9, EP_{10}, EP_{11}, EP_{12} and EP_{13} represent electrostatic potential charges of atom numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13 respectively. Factor 6 shows importance of SAA, RBF, Mp and *n*Bnz. Mp and RBF represent the mean atomic polarizability scaled on carbon atom and rotatable bond fraction, respectively. *n*Bnz denotes the number of benzene-like rings. Factor 3 is highly loaded with the sum of Kier–Hall electrotopological states (Ss), the mean atomic Sanderson electronegativity scaled on carbon atom (Me), the mean electrotopological state of a molecule (Ms) and the number of halogen atoms in a molecule (*n*X). It shows importance of these descriptors. Factor 2 shows importance of GSA, V, Pol, MR, Sv, Se, Sp and *n*R06. Sv, Se and Sp encode the sum of atomic van der Waals volume, the sum of atomic Sanderson electronegativity and the sum of atomic polarizabilities, respectively. All are scaled on carbon atom. *n*R06 stands for the number of six-membered ring. Observed versus LOO-predicted activities of Eq. 1 is graphically represented in Figure 2.

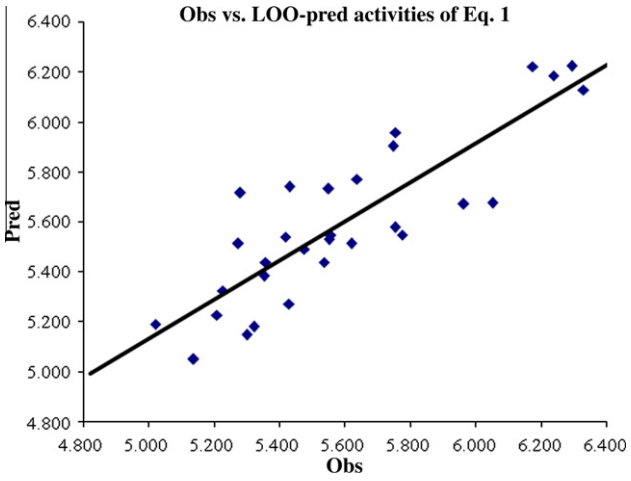


Figure 2. Observed versus LOO-predicted activities of Eq. 1.

Stepwise regression.[13] Attempts were made to develop QSAR models for antiamoebic activity of these compounds. By forward stepwise selection method on the basis of *F* value (*F* = 3.0 for inclusion; *F* = 2.9 for exclusion), Eq. 2 was derived.

$$\begin{aligned} pIC_{50} &= 4.822(\pm 0.225) + 0.329(\pm 0.041)nX \\ &\quad + 0.131(\pm 0.029)nR06 - 7.090(\pm 1.438)RBF \\ &\quad + 1.419(\pm 0.381)Q_{12} \end{aligned} \quad (2)$$

$n = 24;\ R = 0.923;\ R^2 = 0.851;\ R^2_A = 0.820;\ F(4, 19) = 27.174;$
$p < 0.00000;\ SEE = 0.164;\ SSY = 3.423;\ PRESS = 0.767;$
$R^2_{CV} = 0.776;\ SDEP = 0.179;\ S_{PRESS} = 0.201.$

95% confidence intervals of regression coefficients are shown in parentheses. Eq. 2 explains 82.00% and predicts 77.60% variances of antiamoebic activity. Here, positive coefficient of Wang–Ford charge of the atom number 12 (Q_{12}) indicates that antiamoebic activity is increased when the charge has a large positive value at C_{12} (Fig. 1). The charge of C_{12} largely depends on types of substituted aminoaryl and heterocyclic amine groups. Eq. 2 also suggests that the positive coefficient of total number of halogen atoms (*n*X) may contribute positively to antiamoebic activity. Hence, halogen substitution at position X (Fig. 1) and substitituted aminoaryl groups at position R (Fig. 1) may be beneficial for antiamoebic activity. Positive coefficient of number of six-membered ring (*n*R06) indicates that increase in total number of six-membered ring may be beneficial for antiamoebic activity. Hence, six-membered ring substitution at position R along with the benzene ring (Fig. 1) may be conducive for antiamoebic activity. Increase in the value of rotatable bond fraction (RBF)[10] of these compounds may be detrimental to antiamoebic activity. Observed versus LOO-predicted activities of Eq. 2 is shown in Figure 3.

Factor analysis-multiple linear regression (FA-MLR).[12,13] For designing the test set and the training set, *k*-means cluster analysis (*k*-MCA)[14] was performed. *k*-MCA splits these compounds in 3 clusters with 15, 5 and 10 members. Selections of the training set and the test set were carried out randomly using compounds belonging to each cluster. From each cluster, 25% members were considered for designing the test set. Remaining compounds were treated as the training set. Factor analysis was performed on the training set. It is a data-preprocessing step to select descriptors for QSAR equations. QSAR models were developed depending on the training set. It was observed that 6 factors can explain the data matrix to the extent of 97.95%. Antiamoebic activity (pIC_{50}) is highly loaded with factor 3 (highly loaded with Ss, Me, Ms and *n*X), factor 1 (highly loaded with Q_1, Q_2, Q_4, Q_5, Q_6, EP_1, EP_2, EP_3, EP_4, EP_5 and EP_6) and factor 6 (highly loaded with SAA, Mp, RBF and *n*Bnz). pIC_{50} is moderately loaded with factor 2 (highly loaded with GSA, V, Pol, MR, Sv, Se, Sp and *n*R06) and factor 5 (moderately loaded with Q_7, Q_8, Q_9, Q_{11}, EP_7, EP_8, EP_9 and EP_{11}). pIC_{50} is poorly loaded with factor 4 (highly loaded with Q_{10}, Q_{12}, EP_{10}, EP_{12} and EP_{13}). Different combinations of parameters having factor loading of more than 0.70 were subjected to multiple linear regression after removing intercorrelated parameters. The first model obtained according to FA-MLR technique is:

$$\begin{aligned} pIC_{50} &= 7.365(\pm 0.545) - 1.674(\pm 0.206)EP_1 \\ &\quad - 0.009(\pm 0.001)SAA + 0.038(\pm 0.006)Ss \\ &\quad - 0.577(\pm 0.251)EP_{12} \end{aligned} \quad (3)$$

$n = 24;\ R = 0.931;\ R^2 = 0.866;\ R^2_A = 0.838;\ F(4, 19) = 30.822;$
$p < 0.00000;\ SEE = 0.155;\ SSY = 3.423;\ PRESS = 0.677;$
$R^2_{CV} = 0.802;\ SDEP = 0.168;\ S_{PRESS} = 0.189.$

Eq. 3 explains 83.80% and predicts 80.20% of variances of activity. Increasing values of electrostatic potential charges at atom numbers 1 and 12 (EP_1 and EP_{12}) may decrease antiamoebic activity. Electrophilic attacks are likely to occur where the coefficients of electrostatic potential charges are negative. Eq. 3 suggests that electrophilic substitutions may be favourable at C_1 and C_{12} (Fig. 1). The higher negative electrostatic potential at C_1 and C_{12} may be conducive as far as antiamoebic activity is concerned. Here, EP charge of C_1 may largely depend on halogen substitution at position X (Fig. 1). EP charge of C_{12} mainly depends on substitituted heterocyclic amine and aminoaryl groups at position R (Fig. 1). The negative coefficient of SAA implies that higher value of the whole molecular surface area may correspond to lower antiamoebic activity. Increasing value of sum of Kier–Hall electrotopological states (Ss) may be favourable for antiamoebic activity. Observed versus LOO-predicted activities of Eq. 3 is graphically shown in Figure 4.

Another model was developed by replacing the variables 'EP_1' and 'EP_{12}' of Eq. 3 with 'Q_6' and 'Q_{12}' and is shown below:

$$\begin{aligned} pIC_{50} &= 7.231(\pm 0.572) + 3.673(\pm 0.474)Q_6 \\ &\quad - 0.009(\pm 0.001)SAA + 0.040(\pm 0.006)Ss \\ &\quad + 0.789(\pm 0.349)Q_{12} \end{aligned} \quad (4)$$

$n = 24;\ R = 0.934;\ R^2 = 0.873;\ R^2_A = 0.846;\ F(4, 19) = 32.560;$
$p < 0.00000;\ SEE = 0.151;\ SSY = 3.423;\ PRESS = 0.655;$
$R^2_{CV} = 0.809;\ SDEP = 0.165; S_{PRESS} = 0.186.$

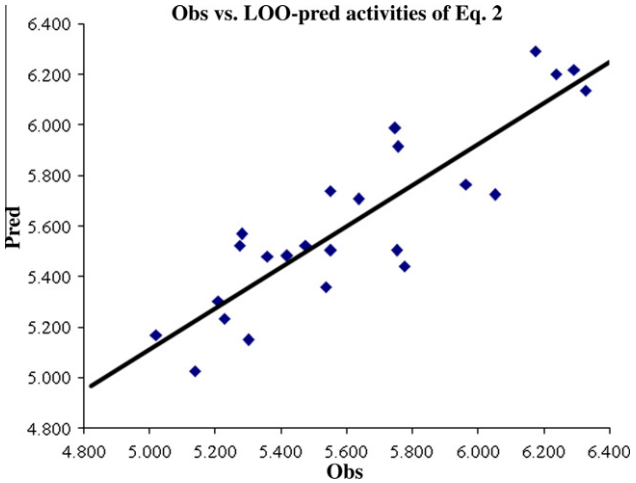


Figure 3. Observed versus LOO-predicted activities of Eq. 2.

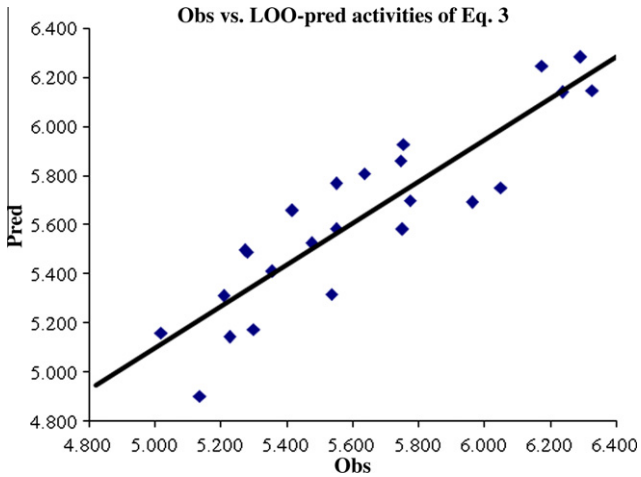


Figure 4. Observed versus LOO-predicted activities of Eq. 3.

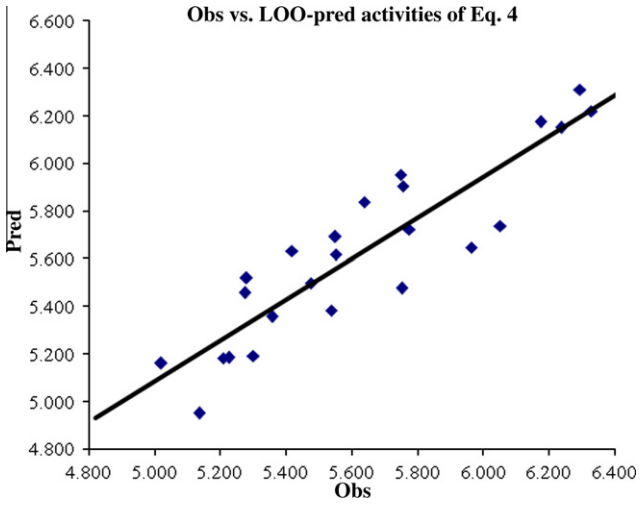


Figure 5. Observed versus LOO-predicted activities of Eq. 4.

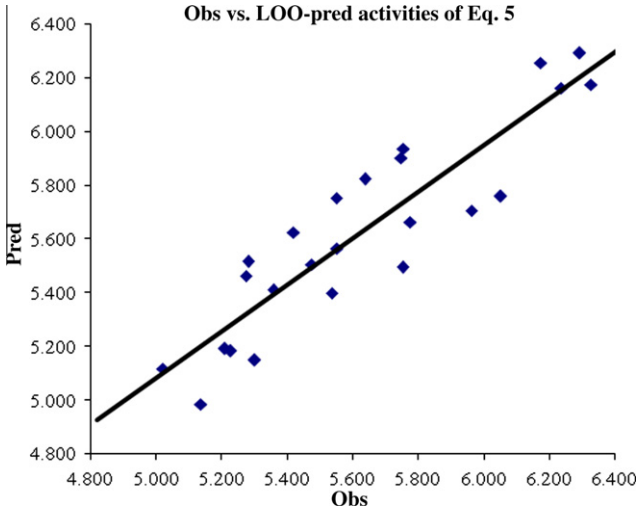


Figure 6. Observed versus LOO-predicted activities of Eq. 5.

Eq. 4 explains 84.60% and predicts 80.90% variances of biological activity. Here, positive coefficients of Q_{12} and Ss as well as the negative coefficient of SAA may be conducive to antiamoebic activity of these compounds. Also, Eq. 4 emphasises about Wang-Ford charge at atom number 6 (Q_6) which may play a pivotal role. Regarding antiamoebic activity, the increasing value of Q_6 may be beneficial as evidenced by the negative coefficient of this parameter in Eq. 4. Charge of C_6 may mainly depend on the atom present in X position (Fig. 1). Halogen atom at X may increase the positive charge of C_6 and thus, may increase the antiamoebic activity. Graphical representation of Observed versus LOO-predicted activities of Eq. 4 is shown in Figure 5.

Another tetravariate model was developed by replacing 'EP_{12}' of Eq. 3, with 'Q_{12}' and is shown in Eq. 5:

$$\begin{aligned} pIC_{50} = {} & 6.492(\pm 0.517) - 1.592(\pm 0.196)EP_1 \\ & - 0.009(\pm 0.001)SAA + 0.043(\pm 0.006)Ss \\ & + 0.965(\pm 0.332)Q_{12} \end{aligned} \quad (5)$$

$n = 24;\ R = 0.939;\ R^2 = 0.882;\ R^2_A = 0.857; F(4, 19) = 35.463;$
$p < 0.00000; \text{SEE} = 0.146;\ \text{SSY} = 3.423;\ \text{PRESS} = 0.612;$
$R^2_{CV} = 0.821;\ \text{SDEP} = 0.160;\ S_{PRESS} = 0.179.$

Eq. 5 explains 85.70% and predicts 82.10% variances of antiamoebic activity. Here, likewise in Eq. 3, the negative coefficient of EP_1, SAA and the positive coefficient of Ss may be conducive to antiamoebic activity. Eq. 5 also suggests that the positive coefficient of Q_{12} may correspond to higher antiamoebic activity. Observed versus LOO-predicted activities of Eq. 5 is shown in Figure 6.

The correlation matrix among the response variable and descriptors used in Eqs. 2–5 is shown in Table 2.

All coefficients of parameters and intercepts in all equations are of 95% confidence intervals as supported by their *t*- and *p*- values. *t*- and *p*- values of Eqs. 2–5 are shown Table 3.

Predicted values of the test set compounds were calculated on the basis of Eqs. 2–5. Observed and predicted values of the test set compounds are listed in Table 4. Significant R^2_{pred} values for the test set were obtained and are shown below:

$R^2_{pred} = 0.677$ for Eq. 2
$R^2_{pred} = 0.880$ for Eq. 3
$R^2_{pred} = 0.809$ for Eq. 4
$R^2_{pred} = 0.845$ for Eq. 5

Partial least square (PLS).[13] The number of optimum components or latent variables was found to be six by cross-validation method to obtain the equation. Based on standardised regression coefficient, variables were selected for the equation. The equation is shown below:

$$\begin{aligned} pIC_{50} = {} & 5.735 + 3.270\ Q_6 + 1.472\ Q_{12} - 0.005\ SAA \\ & + 0.056\ Ss - 5.461\ RBF - 0.112\ nX - 1.070\ Mp \end{aligned} \quad (6)$$

$n = 30;\ R = 0.954;\ R^2 = 0.909;\ R^2_A = 0.885;\ F(6, 23) = 38.48;$
$p < 0.00000;\ \text{SEE} = 0.330;\ \text{PRESS} = 0.621;\ R^2_{CV} = 0.829;$
$SDEP = 0.144;\ S_{PRESS} = 0.164.$

Eq. 6 explains 88.50% and predicts 82.90% variances of the biological activity. Here, positive coefficients of Q_6 and Q_{12} imply that increasing values of these parameters may be beneficial for anti-

Table 2
Correlation matrix of response variable and descriptors used to develop QSAR models

	pIC_{50}	Q_6	Q_{12}	EP_1	EP_{12}	SAA	Ss	RBF	*n*X	*n*R06
pIC_{50}	1.00	0.58	0.14	−0.54	0.16	−0.19	0.51	−0.41	0.67	0.34
Q_6		1.00	0.21	−0.99	0.18	0.41	0.16	−0.06	0.56	0.03
Q_{12}			1.00	−0.17	0.87	0.12	−0.22	0.36	0.01	−0.14
EP_1				1.00	−0.11	−0.41	−0.11	0.11	−0.55	0.03
EP_{12}					1.00	0.29	0.14	0.52	0.23	−0.17
SAA						1.00	0.27	0.74	0.30	−0.33
Ss							1.00	0.20	0.79	0.22
RBF								1.00	−0.01	−0.14
*n*X									1.00	−0.11
*n*R06										1.00

Table 3
t- and *p*-values of Eqs. 2–5

Intercept/parameters	*t*-Value	*p*-Value
Eq. 2		
Intercept	21.472	0.00000
nX	7.953	0.00000
nR06	4.459	0.00027
RBF	−4.931	0.00009
Q_{12}	3.722	0.00145
Eq. 3		
Intercept	13.505	0.00000
EP_1	−8.122	0.00000
SAA	−7.264	0.00000
Ss	6.700	0.00000
EP_{12}	2.297	0.03314
Eq. 4		
Intercept	12.633	0.00000
Q_6	7.749	0.00000
SAA	−7.228	0.00000
Ss	6.951	0.00000
Q_{12}	2.261	0.03566
Eq. 5		
Intercept	12.551	0.00000
EP_1	−8.136	0.00000
SAA	−7.642	0.00000
Ss	7.797	0.00000
Q_{12}	2.906	0.00906

Table 4
Observed and predicted values of test set compounds

Compound	Observed	Predicted value			
		Eq. 2	Eq. 3	Eq. 4	Eq. 5
1	5.428	5.381	5.383	5.353	5.382
9	5.431	5.562	5.519	5.581	5.558
11	5.353	5.389	5.311	5.315	5.366
14	5.556	5.652	5.645	5.577	5.606
16	5.321	5.135	5.323	5.333	5.285
24	5.620	5.491	5.517	5.543	5.490

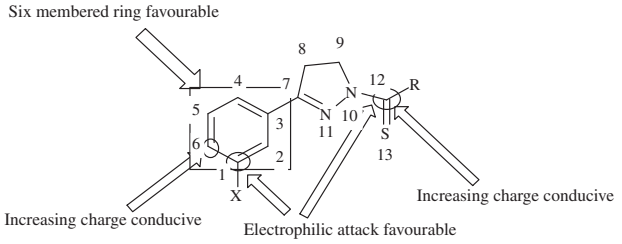


Figure 7. Structural requirements of 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines for antiamoebic activity.

amoebic activity. The charge of C_6 may largely depend on type of halogen substituent (X) (Fig. 1). The charge of C_{12} may mainly depend on type of substituted aminoaryl and heterocyclic amine groups (R) (Fig. 1). Aminoaryl groups may be favourable over heterocyclic amine groups for antiamoebic activity. Again, chlorosubstitution may be preferable regarding antiamoebic activity rather than bromosubstitution at C_1 (Fig. 1). Negative coefficients of approximate surface area (SAA), rotatable bond fraction (RBF),[10] total number of halogen atoms (*n*X) and the mean atomic polarizability scaled on carbon atom (Mp)[10] are indicating that higher values of these parameters may be detrimental to antiamoebic activity. Increase in the value of the sum of Kier–Hall electrotopological states (Ss)[10] may be conducive to antiamoebic activity.

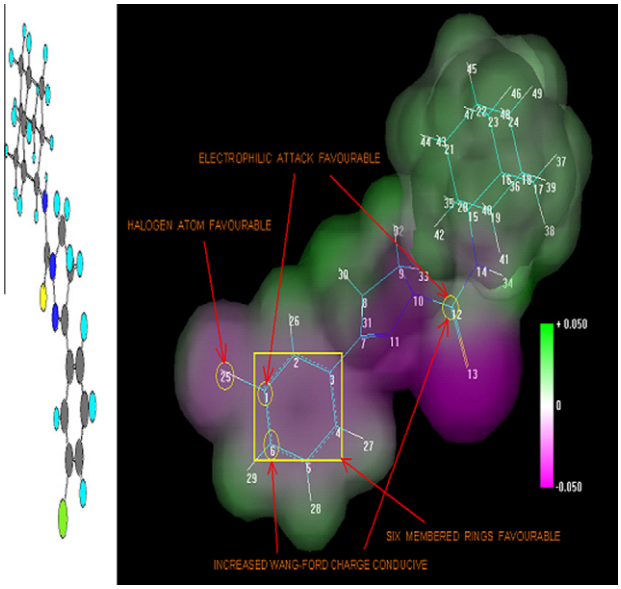


Figure 8. Energy minimized structure and 3D isosurface electrostatic potential map for the best active compound (compound **30**).

This comparative QSAR study has explored structural requirements of 1-N-substituted thiocarbamoyl-3-phenyl-2-pyrazolines for antiamoebic activity. Quality of models obtained from PCRA, stepwise regression, FA-MLR and PLS is of comparable range (explained variance ranging from 78.30% to 88.50% while predicted variance ranging from 73.20% to 82.90%). The best model was obtained from PLS technique with explained and predicted variance of 88.50% and 82.90% respectively. However, the rational selection of descriptors through factor analysis could generate statistically valid models in case of FA-MLR. It is evident from these QSAR models that the increase in the total number of halogen atoms may be conducive to amoebicidal activity. The study signifies that increasing values of Wang–Ford charges at atom numbers 6 and 12 may be beneficial for antiamoebic activity. It also emphasizes that total number of six-membered ring of these compounds may be favourable for antiamoebic activity. The study also reveals that increase in the value of rotatable bond fraction may be detrimental to antiamoebic activity. It also indicates that increase in the value of the sum of Kier–Hall electrotopological states may be conducive to amoebicidal activity. Decreased values of approximate surface area of these molecules and the mean atomic polarizability scaled on carbon atom may be helpful for antiamoebic activity. Atoms and substituents important for amoebicidal activity are shown in Figure 7.

The result was supported by the energy minimized geometry as well as the 3D isosurface electrostatic potential map of the most active compound (compound **30**) obtained during the energy minimizations by AM1 calculations. It is shown in Figure 8.

Acknowledgements

The authors are thankful to All India Council for Technical Education (AICTE), New Delhi for awarding a research project. Two of the authors (N.A. and M.K.M.) are grateful to the University Grants Commission (UGC), New Delhi for awarding Post Graduate Fellowships.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2010.05.098.

References and notes

1. Park, K. *Preventive & Social Medicine*; Banarsidas Bhanot: Jabalpur, 1997. Chapter 5.
2. Kumar, V.; Abbas, A. K.; Fausto N. In *Robbins and Cotran Pathologic basis of disease*, 7th ed., Elsevier, pp 839–840, Chapter 17.
3. Rang, H. P.; Dale, M. M.; Ritter, J. M.; Moore, P. K. *Pharmacology*; Churchill Livingstone: Edinburgh, 2003. Chapter 48.
4. (a) Adhikari, N.; Maiti, M. K.; Jha, T. *Eur. J. Med. Chem* **2010**, *45*, 1119; (b) Jha, T.; Chakrabortty, P.; Adhikari, N.; Halder, A. K.; Maiti, M. K. *Internet Electron J. Mol. Des.* **2009**, *8*, 1. http://www.biochempress.com.; (c) Halder, A. K.; Adhikari, N.; Jha, T. *Chem. Biol. Drug Des.* **2010**, *75*, 204; (d) Jha, T.; Samanta, S.; Basu, S.; Halder, A. K.; Adhikari, N.; Maiti, M. K. *Internet Electron J. Mol. Des.* **2008**, *7*, 234. http://www.biochempress.com.; (e) Halder, A. K.; Adhikari, N.; Jha, T. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1737; (f) Alam, Sk. M.; Samanta, S.; Halder, A. K.; Jha, T. *Eur. J. Med. Chem.* **2009**, *44*, 359; (g) Samanta, S.; Alam, Sk. M.; Panda, P.; Jha, T. *Eur. J. Med. Chem.* **2009**, *44*, 70; (h) Panda, P.; Samanta, S.; Alam, Sk. M.; Basu, S.; Jha, T. *Internet Electron J. Mol. Des.* **2007**, *6*, 280. http://www.biochempress.com.; (i) Samanta, S.; Panda, P.; Alam, Sk. M.; Jha, T. *Internet Electron J. Mol. Des.* **2007**, *6*, 183. http://www.biochempress.com.; (j) Basu, A.; Gayen, S.; Samanta, S.; Panda, P.; Srikanth, K.; Jha, T. *Can. J. Chem.* **2006**, *84*, 458; (k) Debnath, B.; Gayen, S.; Samanta, S.; Basu, A.; Ghosh, B.; Jha, T. *Indian J. Chem., Sect. A* **2006**, *45*, 93; (l) Samanta, S.; Debnath, B.; Basu, A.; Gayen, S.; Srikanth, K.; Jha, T. *Eur. J. Med. Chem.* **2006**, *41*, 1190; (m) Samanta, S.; Gayen, S.; Ghosh, B.; Panda, P.; Srikanth, K.; Jha, T. *Int. J. Appl. Chem.* **2006**, *2*, 169; (n) Samanta, S.; Alam, Sk. M.; Panda, P.; Jha, T. *Internet Electron J. Mol. Des.* **2006**, *5*, 503. http://www.biochempress.com.; (o) Gayen, S.; Debnath, B.; Samanta, S.; Ghosh, B.; Basu, A.; Jha, T. *Internet Electron J. Mol. Des.* **2005**, *4*, 556. http://www.biochempress.com.; (p) Samanta, S.; Debnath, B.; Gayen, S.; Ghosh, B.; Basu, A.; Jha, T. *IL Farmaco* **2005**, *10*, 818; (q) Debnath, B.; Samanta, S.; Gayen, S.; Basu, A.; Ghosh, B.; Jha, T. *Internet Electron J. Mol. Des.* **2005**, *4*, 393. http://www.biochempress.com.; (r) Gayen, S.; Debnath, B.; Basu, A.; Samanta, S.; Ghosh, B.; Naskar, S. K.; Jha, T. *Internet Electron J. Mol. Des.* **2005**, *4*, 210. http://www.biochempress.com.; (s) Debnath, B.; Gayen, S.; Basu, A.; Ghosh, B.; Srikanth, K.; Jha, T. *Bioorg. Med. Chem.* **2004**, *12*, 6137; (t) Debnath, B.; Gayen, S.; Basu, A.; Srikanth, K.; Jha, T. *J. Mol. Model.* **2004**, *10*, 328; (u) Gayen, S.; Debnath, B.; Jha, T. *Internet Electron J. Mol. Des.* **2004**, *3*, 771. http://www.biochempress.com.; (v) Samanta, S.; Srikanth, K.; Banerjee, S.; Debnath, B.; Gayen, S.; Jha, T. *Bioorg. Med. Chem.* **2004**, *12*, 1413; (w) Gayen, S.; Debnath, B.; Samanta, S.; Jha, T. *Bioorg. Med. Chem.* **2004**, *12*, 1493; (x) Debnath, B.; Gayen, S.; Naskar, S. K.; Roy, K.; Jha, T. *Drug Des. Discovery* **2003**, *18*, 81; (y) Debnath, B.; Gayen, S.; Bhattacharya, S.; Samanta, S.; Jha, T. *Bioorg. Med. Chem.* **2003**, *11*, 5493; (z) Jha, T.; Debnath, B.; Samanta, S.; De, A. U. *Internet Electron J. Mol. Des.* **2003**, *2*, 539. http://www.biochempress.com.; (aa) Debnath, B.; Gayen, S.; Naskar, S. K.; Roy, K.; Jha, T. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2837; (ab) Debnath, B.; Gayen, S.; Roy, K.; Jha, T. *Bioorg. Med. Chem.* **2003**, *11*, 1615; (ac) Debnath, B.; Vishnoi, S. P.; Sa, B.; Jha, T. *Internet Electron J. Mol. Des.* **2003**, *1*, 128. http://www.biochempress.com.; (ad) Srikanth, K.; Debnath, B.; Jha, T. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 899; (ae) Debnath, B.; Srikanth, K.; Banerjee, S.; Jha, T. *Internet Electron J. Mol. Des.* **2002**, *1*, 488. http://www.biochempress.com.; (af) Srikanth, K.; Kumar, C. A.; Ghosh, B.; Jha, T. *Bioorg. Med. Chem.* **2002**, *10*, 2119; (ag) Srikanth, K.; Debnath, B.; Jha, T. *Bioorg. Med. Chem.* **2002**, *10*, 1841; (ah) Srikanth, K.; Kumar, C. A.; Goswami, D.; De, A. U.; Jha, T. *Indian J. Biochem. Biophys.* **2001**, *38*, 120.
5. Abid, M.; Bhat, A. R.; Athar, F.; Azam, A. Eur. *J. Med. Chem.* **2009**, *44*, 417.
6. CHEM 3D PRO Version 5.0 and CHEM DRAW ULTRA Version 5.0 are software programs developed by Cambridge Soft Corporation, U. S. A.
7. Leach, A. R. *Molecular Modeling Principles and Applications*; Pearson: London, 2001. Chapter 4.
8. Hyperchem Professional Release 7.0, Hypercube Inc., Gainesville, Florida.
9. DRAGON web version 2.1 developed by Milano Chemometrics and QSAR Research Group, Dipartminto di Scienze dell' Ambiente e del Territorio Universitàdegli Studi di Milano, Bicocca.
10. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, 2000.
11. Tetko, I. V.; Tanchuk, V. Y.; Villa, A. E. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1407.
12. Franke, R. *The Theoretical Drug Design Methods*; Elsevier: Amsterdam, 1984.
13. (a) Leonard, J. T.; Roy, K. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4467; (b) Leonard, J. T.; Roy, K. *Bioorg. Med. Chem.* **2006**, *14*, 1039.
14. Tropsha, A.. In *Burger's Medicinal Chemistry and Drug Discovery*; Abraham, D. J., Ed.; John Wiley and sons: Inc., New Jersey, 2003; Vol. 1, pp 49–75.