RESEARCH PAPER | Pharmaceutical Science | Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

Qsar Modeling of in Vitro Vero Cell Cytotoxicity of Some Antitubercular Agents Containing Chiral Pentaamines, Bis–Cyclic Guanidines, Bis-Cyclic Thioureas and Bis-Cyclic Piperazines Moieties Using 2D Descriptors

KEYWORDS: Cytotoxicity, QSAR, 2D descriptors, FA-MLR, PLS

Sant Kumar Verma

Institute of Pharmaceutical Sciences, Guru Ghasidas Central University, Bilaspur-495 009 (India)

ABSTRACT *Cytotoxicity assay is important in screening of lead compounds in final stage of drug discovery process. In the present study QSAR analysis of in vitro vero cell cytotoxicity of 57 antitubercular compounds belonging to chiral pentaamines, bis-cyclic guanidines, bis-cyclic thioureas and bis-cyclic piperazines scaffolds was performed. Three chemometric tools (FA-MLR, Stepwise MLR, PLS) were applied on freely online available PaDEL 2D descriptor (an open source) for the model development. All the models were statistically robust both internally (Q2: 0.877-0.918) and externally (Q2F1: 0.933-0.942; Q2F2: 0.933-0.941; Q2F3: 0.866-0.883, CCC: 0.958-0.966). All the models satisfy the external validation parameters proposed by Tropsha, Roy and Nicola (2002, 2008 and 2011). All the models highlight the importance of hydrogen bonding as important property for cytotoxicity. The first theoretical model of in vitro vero cell cytotoxicity for antitubercular agents may be helpful for drug development process.*

Introduction

Among the reemerging diseases tuberculosis remains one of the major concerns in recent times for the developing countries which was announced long before in 1993 by WHO. The reemergence is due to the fact that the disease is co-infected with the human immunodeficiency virus (HIV) and there is 5-15% increase of infection annually [1, 2].

There were an estimated 8.7 million new cases of tuberculosis (TB) (13% co-infected with HIV) and 1.4 million people died from TB, including almost one million deaths among HIV-negative individuals and 430000 among people who were HIV positive in 2011. Globally, 3.7% (2.1-5.2%) of new cases and 20% (13-26%) of previously treated cases are estimated to have multi drug resistant (MDR) TB [3]. Totally drug resistant (TDR) TB is a generic term for TB strains that are resistant to a wider range of drugs than strains classified as extensively drug resistant (XDR) TB [4]. The major obstacles to the global control of this infectious disease include the difficulties to detect and cure a sufficient number of cases to interrupt transmission.

QSAR models can serve as an important tool for automated pre-virtual screening for in silico activity prediction, optimization of the lead, data mining and combinatorial library design. QSAR studies on various scaffolds were performed by different researcher groups to generate hit molecules, with improved activity, in the process of rational design of more potent antitubercular agents [5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16].

Cytotoxicity assays were performed to screen the potential ligands in pharmaceutical industry. It is the toxic properties of any substance to cells. The choice of cytotoxicity depends on the use either to choose cytotoxic agents (for cancer therapy) or least cytotoxic agent for therapeutic use. The in vitro vero cell cytotoxicity assay is being performed by different researcher groups for screening new antitubercular agents as well as to find the selectivity index (SI) which is the ratio of measured IC_{50} in vero cell to MIC [17, 18]. Compounds with MIC value less than 1 µg/mL seems to be good lead compound and interesting compounds are those with an MIC ≤ 6.25 µg/mL and an SI ≥ 10 [19]. Computational approaches in designing, screening and finding optimal interactions at active sites by molecular docking were performed by different groups considering the MIC values. Till the date there is no availability of theoretical cytotoxicity models for antitubercular agent for screening. Therefore there is a need to explore the properties associated with cytotoxicity.

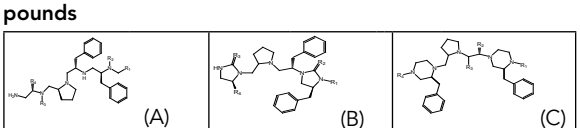
For the present study we have taken the in vitro vero cell cytotoxicity data of some chiral pentaamines and bis-heterocyclic compounds showing good antitubercular activity [20]. The study was done taking by 2D descriptors available on PaDEL software (open source) using FA-MLR, Stepwise MLR and FA-PLS techniques. This study provides significant insight on the applicability of such statistical models in identifying the features relevant for cytotoxicity.

Materials and methods

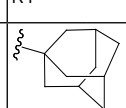
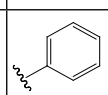
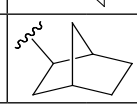
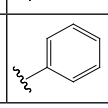
1. Data set

In the present study we have considered in vitro vero cell cytotoxicity (IC_{50} µg/ml) of 57 anti-tubercular compounds (Table 1) from chiral pentaamines, bis-cyclic guanidines, bis-cyclic thioureas and bis-cyclic piperazines [20]. The vero cell IC_{50} values in µg/ml unit of selected data set compounds were firstly converted to mole/ml unit and further converted to -logIC_{50} (known as pIC_{50}) (Table 2) to reduce the skewness of the data and used as response variable for subsequent QSAR analysis.

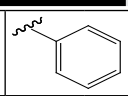
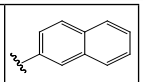
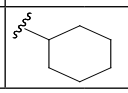
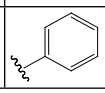
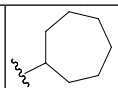
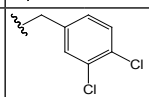
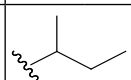
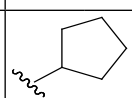
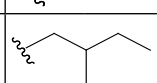
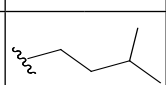
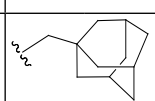
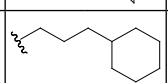
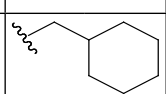
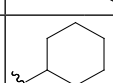
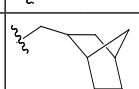
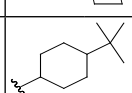
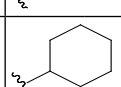
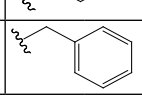
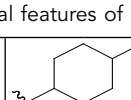
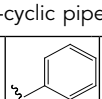
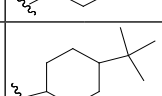
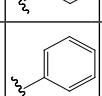
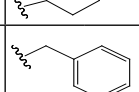
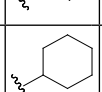
Table 1: Structural features of selected data set compounds

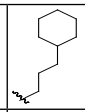
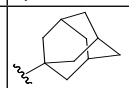


(A): Structural features of chiral pentaamines

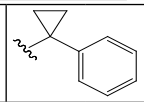
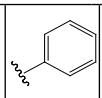
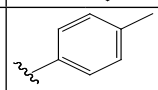
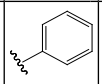
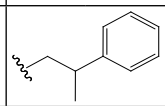
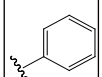
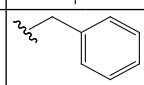
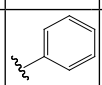
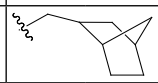
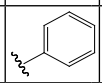
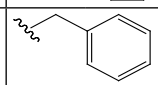
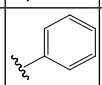
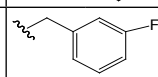
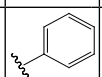
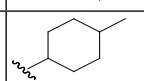
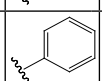
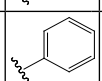
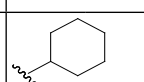
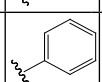
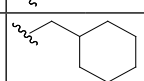
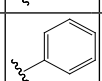
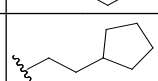
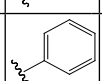
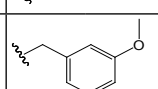
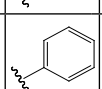
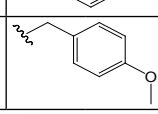
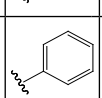
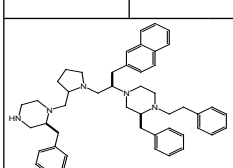
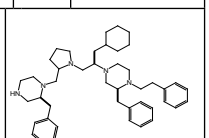
Compound No.	R1	R2	R3	R4
1		-H	-H	
2		-H	-H	

Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

3		-H	-H	
5*		-H	-H	
(B): Structural features of bis-cyclic guanidines and bis-cyclic thioureas				
6*		-NH	-NH	
7		-NH	-NH	
8*		-NH	-NH	
9		-NH	-NH	
10		-NH	-NH	
11*		-NH	-NH	
12		-NH	-NH	
13		-NH	-NH	
14		-NH	-NH	
15		-NH	-NH	
16		-NH	-NH	
17*		-NH	-NH	
18		-S	-S	
19		-S	-S	-Isopropyl
(C): Structural features of bis-cyclic piperazines				
4*			-H	-H
20			-H	-H
21*			-H	-H

22			-H	-H
23			-H	-H
24			-H	-H
25*			-H	-H
26			-H	-H
27			-H	-H
28			-H	-H
29*			-H	-H
30			-H	-H
32			-H	-H
33*			-H	-H
34		-Isopropyl	-H	-H
35		-n-butane	-H	-H
36*			-H	-H
37			-H	-H
38	-n-heptane		-H	-H
39			-H	-H
40*			-H	-H
41*		-2-methyl propane	-H	-H
42			-H	-H

Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

44			-H	-H
45			-H	-H
46*			-H	-H
47			-H	-H
48*			-H	-H
49			-H	-H
50			-H	-H
51			-H	-H
52	-2-methyl-butane		-H	-H
53*			-H	-H
54*			-H	-H
55			-H	-H
56			-H	-H
57*			-H	-H
compound no. 31 		compound no. 43* 		

*denotes test set compounds

Table 2: Observed vero cell cytotoxicity values of chiral pentaamines, bis-cyclic thioureas, bis-cyclic guanidines and bis-cyclic piperazines compounds

Compound no.	pIC_{50} # (observed)	Compound no.	pIC_{50} # (observed)
1	8.779	30	8.805
2	8.781	31	8.823
3	8.189	32	8.807
4*	8.678	33*	8.830
5*	8.764	34	8.793
6*	7.223	35	8.793
7	7.469	36*	8.848
8*	7.227	37	8.832
9	7.275	38	8.800
10	7.293	39	8.803
11*	7.310	40*	8.864
12	7.692	41*	8.798
13	7.681	42	8.834
14	7.675	43*	8.857
15	7.387	44	8.860
16	7.638	45	8.872
17*	7.664	46*	8.895
18	7.904	47	8.882
19	8.274	48*	8.889
20	8.787	49	8.907
21*	8.825	50	8.925
22	8.839	51	8.907
23	8.858	52	8.948
24	8.857	53*	8.960
25*	8.854	54*	9.028
26	8.824	55	9.197
27	8.810	56	9.331
28	8.803	57*	9.317
29*	8.782		

***denotes test set compounds, #unit: mole/ml**

2. Descriptors

Descriptors are chemical information about a molecular structure with a numeric representation. The structures of selected data set compounds were sketched and AM-1energy minimized using CS Chemoffice software package [21] and saved in .mol format which is one of the suitable input format for PaDEL and used as input structure for descriptor calculations. Only freely available online PaDEL 2D descriptors were considered for the present study [22].

Initially 727 descriptors were calculated using PaDEL software version 2.12 [22]. Then we deleted the descriptors with high intercorrelation (0.95), as well as zero and constant value descriptors. Finally pruned 213 descriptors were chosen for QSAR analysis of selected data set. Additionally we used the MICs (minimum inhibitory concentrations) of compounds converted to –logMIC as descriptor.

Model development

The creditability of QSAR analysis is basically depends on the selected data set, appropriate descriptors, the statistical methods used for analysis of data set and most importantly validation strategies adopted for developed model [23]. Initially we have divided the whole data set in to test set (with 19 compounds, comprising approximately 33% of total compounds) and training set (with 38 compounds, comprising approximately 67% of total compounds). K-means clustering on standardized descriptors matrix was used as a tool for the division of the compounds in to test and training set [24]. It is a non hierarchical clustering technique to classify objects based on characters in to K number of group (where, K is positive integer number).

Initially models were developed from the training set compounds considering the model fitting parameters like

Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

squared correlation coefficient (R2), explained variance (R2a) and F value at specific degree of freedom (df) indicating the robustness of models [24]. The predictive ability of the equation for internal validation was evaluated by a well-known method of Leave-one-out (LOO) cross-validation R2 (Q2) [25, 26, 27]. In addition predicted residual sum of squares (PRESS) was also checked. For the external validation the parameters Q2F1, Q2F2, Q2F3 [28, 29, 30] and metric of r^2_m [$r^2_{m(LOO)}$, $r^2_{m(overall)}$ and $r^2_{m(test)}$] [31, 32] and CCC [33] were calculated. Different metric for r^2_m were calculated using online available link [34].

For the modeling three chemo metric tools (FA-MLR, Stepwise MLR and FA-PLS) were used [35, 36, 37, 38]. The FA was performed using the statistical software SPSS [36, 39]. K-means clustering, stepwise regression, standardization of the variables and PLS were performed using statistical software MINITAB [40].

Results and discussions

Initial division of data set was done by using K means clustering techniques and the PCA score plot (Figure 1) shows uniform distribution of test set compounds within the domain of training set.

FA-MLR

The factor analysis of the data matrix shows 10 factors could explain 95.083% of the variance of the total data matrix (all descriptors along with the response variable). The response variable is highly loaded with factor 1 (which in turn highly loaded in SHBint5, maxHBint2, SHBint2, maxHBint5, maxHBint8, mindNH, SdNH, maxdNH, minHdN and SHBd) while it is moderately

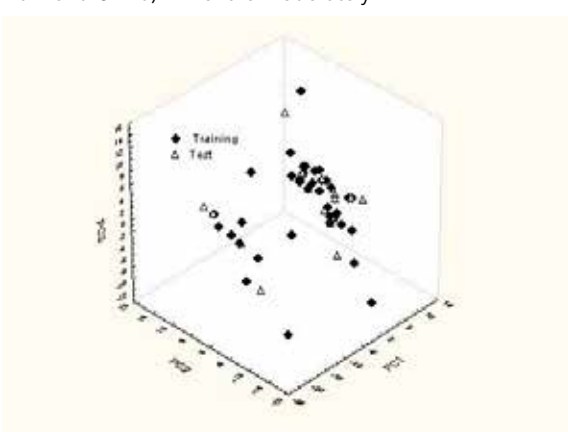


Fig.1. PCA score plot of first three components of the descriptor matrix shows distribution of training and test set compounds in 3D space

loaded with factor 2 (highly loaded in nBonds, WTPT-1, VAdjMat, Zagreb, SP-2). The best predictive equation was obtained in the combination of maxHBint2 (maximum E-state descriptor of strength for potential hydrogen bonds of path length 2) and nBonds (number of bonds excluding bonds with hydrogen) descriptors.

$$pIC_{50} = 7.486\ (\pm 0.652) - 0.597\ (\pm 0.036)\ maxHBint2 + 0.024\ (\pm 0.012)\ nBonds$$

$$n_{Training} = 38,\ R^2 = 0.890,\ R^2_a = 0.883,\ Q^2 = 0.877,\ r^2_{m(LOO)} = 0.868,\ PRESS = 1.546,\ F = 141\ (df\ 2, 35),$$

$$n_{Test} = 19,\ Q^2_{F1} = 0.942,\ Q^2_{F2} = 0.941,\ Q^2_{F3} = 0.883,\ r^2_{m(test)} = 0.868,\ r^2_{m(overall)} = 0.879 \quad (1)$$

According to standardized coefficient maxHBint2 contribute more than nBonds. The standard errors of regression coefficients are given within parenthesis. The leave-one-out predicted variance was found to be 87.7%. Two variables maxHBint2 and nBonds in equation (1) could explain 88.3% of the variance (adjusted coefficient of variation) of the vero cell toxicity. The test set compounds were externally predicted applying equation (1), the predicted R^2 (Q^2_{F1}) for the test set was found to be 0.942. The absence of high intercorrelation (r) was found among predictor variables.

Eq. 1 explains that the compounds having greater value for nBonds descriptor will show higher vero cell toxicity due to positive contribution of nBonds descriptor towards vero cell toxicity. The hydrogen bonding descriptor maxHBint2 contribute negatively towards vero cell cytotoxicity. Bis-cyclic guanidines (compound no. 7, 9, 10, 12, 13, 14, 15, 16) and bis-cyclic thioureas (compound no. 18, 19) were observed with positive value for maxHBint2 descriptor among the all training set compounds showing less cytotoxicity value. The chiral pentaamines (compound no. 2, 3) and bis-cyclic piperazines (compound no. 20, 22, 23, 24, 26, 27, 28, 30, 31, 32, 34, 35, 37, 38, 39, 42, 44, 45, 47, 49, 50, 51, 52, 55, 56) were found to be with zero value for maxHBint2 descriptor therefore these compounds show greater pIC_{50} value in comparison to bis-cyclic guanidines and bis-cyclic thioureas.

STEPWISE

$$pIC_{50} = 9.532\ (\pm 0.225) - 2.128\ (\pm 0.130)\ maxHBint8 - 0.82\ (\pm 0.210)\ SHBint9 - 2.2\ (\pm 0.996)\ minHBd$$

$$n_{Training} = 38,\ R^2 = 0.947,\ R^2_a = 0.943,\ Q^2 = 0.918,\ r^2_{m(LOO)} = 0.901,\ PRESS = 1.026,\ F = 203.188\ (df\ 3, 34),$$

$$n_{Test} = 19,\ Q^2_{F1} = 0.933,\ Q^2_{F2} = 0.933,\ Q^2_{F3} = 0.866,\ r^2_{m(test)} = 0.847,\ r^2_{m(overall)} = 0.904 \quad (2)$$

The best combination of descriptors according to specified stepping criteria (F to enter 4, F to remove 3.9) were found to be maxHBint8 (maximum E-state descriptor of strength for potential hydrogen bonds of path length 8), SHBint9 (sum of E-state descriptor of strength for potential hydrogen bonds of path length 9) and minHbd (minimum E-state for strong hydrogen bonds donars). The electro topological state (E-state) hydrogen bonding descriptors of the eq. 2 contribute negatively towards vero cell toxicity. It is clear that the stepwise equation is mainly governed by the E-state descriptors of hydrogen bonding. According to the relative importance of descriptors maxHBint8 contribute mostly towards vero cell toxicity and it was observed that all bis-cyclic piperazines (compound no. 20, 22, 23, 24, 26, 27, 28, 30, 31, 32, 34, 35, 37, 38, 39, 42, 44, 45, 47, 49, 50, 51, 52, 55, 56) and all chiral pentaamines (compound no. 2, 3) in training set having zero value for maxHBint8 descriptor thus these compounds show greater values of vero cell toxicity in comparison to bis-cyclic guanidines and bis-cyclic thioureas. All the bis-cyclic guanidines and bis-cyclic thioureas were found to be with a definite positive value and compound no. 19 having least value for maxHBint8 descriptor thus it contributes greater towards vero cell toxicity in comparison to rest of the bis-cyclic guanidines (compound no. 7, 9, 10, 12, 13, 14, 15, 16, 18) and bis-cyclic thioureas (compound no. 20). Bis-cyclic piperazines (compound no. 2, 3) with a definite positive value for SHBint9 and minHbd descriptors showed higher toxicity value due to zero value of maxHBint8 which contributes mostly towards vero cell toxicity. The statistical parameters of the equation 2 along with other models are listed in Table 3 which indicates robustness of the model.

Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

FA-PLS

For FA-PLS initial reduction of number of descriptors was done using factor analysis. From pool of 213 descriptors we reduced the number of descriptors to 107 by selecting the variables in highly loaded factors. The final equation (Eq. 3) was obtained with one component, optimized with Q^2 and showed good internal and external parameters which are listed in Table 3.

$$pIC_{50} = 8.909 - 0.462SHBd - 0.050SHBint2 - 0.154SHBint8 - 0.567SHdNH - 0.343maxHBint8$$

$$n_{training} = 38,\ R^2 = 0.888,\ R^2_a = 0.857,\ Q^2 = 0.877,\ r^2_{m(LOO)} = 0.877,\ PRESS = 1.540,\ F = 285.610$$

$$(df\ 1, 36),$$

$$n_{Test} = 19,\ Q^2_{F1} = 0.936,\ Q^2_{F2} = 0.936,\ Q^2_{F3} = 0.872,\ r^2_{test} = 0.842,\ r^2_{overall} = 0.868 \quad (3)$$

The E-state atom type descriptors namely SHBd (sum of E-states for strong hydrogen bond donars), SHBint2 (sum of E-state descriptors of strength for potential hydrogen bonds of path length 2), SHBint8 (sum of E-state descriptors of strength for potential hydrogen bonds of path length 8), SHdNH (sum of atom-type H E-state: =NH) and maxHBint8 were obtained in FA-PLS equation and all these descriptors have negative contribution to the vero cell toxicity. Bis-cyclic guanidines (compound no. 7, 9, 10, 12, 13, 14, 15, 16) with higher value for SHBd, SHBint2, SHBint8, SHdNH and maxHBint8 descriptors were found to be less toxic to vero cell. And on the contrary chiral pentaamines (compound no. 2, 3), bis-cyclic thioureas (compound no. 18, 19) and bis-cyclic piperazines (compound no. 20, 22, 23, 24, 26, 27, 28, 30, 31, 32, 34, 35, 37, 38, 39, 42, 44, 45, 47, 49, 50, 51, 52, 55, 56) with zero or lesser value for E-state atom type descriptors (SHBd, SHBint2, SHBint8, SHdNH and maxHBint8)showing high value of vero cell toxicity.

All chiral pentaamines and bis-cyclic piperazines in training set were found to be with zero value for E-state descriptors of strength for potential H-bond (SHBint2, SHBint8, maxHBint8) and atom type H E-state: =NH (SHdNH) descriptor hence showing greater vero cell toxicity in comparison to bis-cyclic guanidines and bis-cyclic thioureas. In addition, bis-cyclic thioureas were also having zero value for SHdNH descriptor due to absence of imine (=NH) group and hence showing greater vero cell toxicity in comparison to bis-cyclic guanidines.

The E-state descriptor for strong H-bond donor (SHBd) was found to be with greater values for all bis-cyclic guanidines in comparison to chiral pentaamines, bis-cyclic thioureas and bis-cyclic piperazines, thus bis- cyclic guanidines showing lesser vero cell toxicity among all compounds in training set.

Further test on external validation

The external validation statistics of the developed models were further verified by using different external validation parameters proposed by Golbraikh and Tropsha {i.e., (i) $Q^2 > 0.5$, (ii) $r^2 > 0.6$, (iii) r_0^2 or $r_0'^2$ is close to r^2, such that $[(r^2 - r_0^2)/r^2]$ or $[(r^2 - r_0'^2)/r^2] < 0.1$ and $0.85 \leq k \leq 1.15$ or $0.85 \leq k' \leq 1.15$} [38, 41] and Roy et al [average r_m^2 ($\overline{r_m^2}$) should be > 0.5 & delta r_m^2 (Δr_m^2) should be < 0.2] [39, 42] and by Insubria group [concordance correlation coefficient (CCC) with cut off value of 0.85] [33, 40]. All the developed models satisfy all the statistical criteria (listed in Table 3, 4 and 5) as proposed by the above authors.

Table 3: Statistical quality and different validation metrics of developed model

Type of statistical method	R^2	Q^2	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	CCC
FA-MLR	0.890	0.877	0.942	0.941	0.883	0.966
Stepwise	0.947	0.918	0.933	0.933	0.866	0.958
FA-PLS	0.888	0.877	0.936	0.936	0.872	0.963

Table 4: External validation characteristics of developed model according to Golbraikh and Tropsha [41]

Parameters	Types of statistical method		
	FA-MLR	Stepwise	FA-PLS
r^2	0.950	0.943	0.952
r_0^2	0.943	0.933	0.939
$r_0'^2$	0.928	0.911	0.917
$(r2 - r_0^2)/r^2$	0.007	0.011	0.014
$(r2 - r_0'^2)/r^2$	0.023	0.034	0.037
k	1.100	1.116	1.138
k/	0.864	0.845	0.836

Table 5: Validation characteristics of developed model according to r2m metrics [42, 43]

Parameters	Type of statistical method		
	FA-MLR	Stepwise	FA-PLS
$r^2_{m(test)}$	0.868	0.847	0.842
$r'^2_{m(test)}$	0.807	0.774	0.774
$\overline{r^2_{m(test)}}$	0.838	0.811	0.808
$\Delta r^2_{m(test)}$	0.061	0.073	0.068
$r^2_{m(LOO)}$	0.868	0.901	0.877
$r'^2_{m(LOO)}$	0.770	0.862	0.770
$\overline{r^2_{m(LOO)}}$	0.819	0.881	0.824
$\Delta r^2_{m(LOO)}$	0.099	0.039	0.107
$r^2_{m(overall)}$	0.879	0.904	0.868
$r'^2_{m(overall)}$	0.782	0.829	0.770
$\overline{r^2_{m(overall)}}$	0.830	0.867	0.819
$\Delta r^2_{m(overall)}$	0.097	0.074	0.098

Overview and conclusion

For the robustness and statistical significance of the developed models initial division of dataset was done for training and test set compounds selection by using K means clustering techniques. Three chemometric tools (FA-MLR,

Volume : 4 | Issue : 9 | September 2014 | ISSN - 2249-555X

Stepwise MLR, and FA-PLS) were used for model building. All the models are statistically robust both internally (Q^2: 0.877-0.918) and externally (Q^2_{F1}: 0.933-0.942; Q^2_{F2}: 0.933-0.941; Q^2_{F3}: 0.866-0.883) and satisfy the criteria of acceptable QSAR model proposed by different groups [31, 32, 33].

The models of cytotoxicity (FA-MLR, Stepwise MLR, FA-PLS) indicates the importance of hydrogen bonding parameters (maxHBint2 maxHBint8 SHBint9 minHbd SHBd, SHBint2, SHBint8, SHdNH). All the models indicate less hydrogen bonding potentials of compounds responsible for greater cytotoxicity towards vero cell. Bicyclic guanidines and thioureas show less cytotoxicity contrary to bicyclic piperazines and chiral pentaamines based on their hydrogen bonding potential. Cytotoxicity assays were performed in final stage of lead screening and the theoretical model could be helpful in designing and drug development process.

REFERENCE

1. Corbett, E.L., Dye, C., Maher, D., Raviglione, M.C., Walker, N., Watt, C.J., Williams, B.G. (2003) The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch Intern Med 163: 1009-1021. | 2. Harries, A.D., Mahler, D. (1996) TB/HIV, a clinical manual. World Health Organization: Geneva. | 3. World Health Organization report: Global tuberculosis control (2011) ISBN 978 92 4 156438 0. | 4. Shah, N.S., et al. (2007) Worldwide emergence of extensively drug-resistant tuberculosis. J Emerg Infect Dis 13: 380-387. | 5. Aparna, V., Desiraju, G.R., Gopalakrishnan, B., Jeevan, J. (2005) A virtual screening approach for thymidine monophosphate kinase inhibitors as antitubercular agents based on docking and pharmacophore models. J Chem inf model 45: 1101–1108. | 6. Coutinho, E., Jain, R., Malde, A., Nayyar, A. (2006) 3D-QSAR study of ring-substituted quinoline class of anti-tuberculosis agents. Bioorg and Med Chem 14: 847–856. | 7. Baldi, P., Gago, G., Gramajo, H., Kurth, D., Lin, T.W., Melgar, M.M., Purdon, J., Swamidass, S.J., Tsai, S.C., Tseng, T. (2006) Structure-based inhibitor design of AccD5, an essential acyl-CoA carboxylase carboxyltransferase domain of Mycobacterium tuberculosis. Pro Nat Acad Sci U.S.A. 103: 3072–3077. | 8. Geetha Babu, S.K., Mukesh, D., Sivakumar, P.M. (2007) QSAR studies on chalcones and flavonoids as anti-tuberculosis agents using genetic function approximation (GFA) method. Chem Pharm Bullet 55: 44–49. | 9. Kumar, A., Panda, G., Shagufta, Siddiqui, M.I. (2007) CoMFA and CoMSIA 3D-QSAR analysis of diaryloxy methano-phenanthrene derivatives as anti-tubercular agents. J Mol Model 13: 99–109. | 10. Kumar, A., Siddiqi, M.I. (2008) CoMFA based de novo design of pyrrolidine carboxamides as inhibitors of enoyl acyl carrier protein reductase from Mycobacterium tuberculosis. J Mol Model 14: 923–935. | 11. Cho, Y., Ioerger, T.R., Sacchettini, J.C., (2008) Discovery of novel nitrobenzothiazole inhibitors for Mycobacterium tuberculosis ATP phosphoribosyl transferase (HisG) through virtual screening. J Med Chem 51: 5984–5992. | 12. Dolezal, R., Kubicova, L., Kunes, J., Petrlikova, E., Waisser, K. (2009) N-benzylsalicylthioamides: highly active potential antituberculotics. Arch Pharm 342: 113–119. | 13. Brandt, C.A., Felli, V.M., Fernandes, J.P., Ferreira, E.I., Pasqualoto, K.F. (2010) QSAR modeling of a set of pyrazinoate esters as antituberculosis prodrugs. Arch Pharm 343: 91–97. | 14. Acharya, H.D., Coutinho, E.C., Dholakia, C.D., Manvar, A.T., Manvar, D.R., Mishra, A.K., Parecha, A.R., Pissurlenkar, R.R.S., Shah, A.K., Virsodia, V.R. (2010) Synthesis, in vitro antitubercular activity and 3D QSAR study of 1,4-dihydropyridines. Mol Divers 14: 285–305. | 15. De-sousa, J.A., Kovalishyn, V., Leitao, R.E., Ventura, C. (2011) QSAR modeling of antitubercular activity of diverse organic compounds. Chemom Intell Lab Syst 107: 69-74. | 16. Chauhan, N.A., Chawda, R.S., Coutinho, E.C., Khedkar, V.M., Khunt, R.C., Parikh, A.R. (2012) Synthesis, antitubercular evaluation and 3D-QSAR study of N-phenyl-3-(4-fluorophenyl)-4-substituted pyrazole derivatives. Bioorg Med Chem 22: 666-678. | 17. Bjorn, S., Jon, N.M., Lise, L.G. (2002) Synthesis and antimycobacterial activity of 6-arylpurines: the requirements for the N-9 substituent in active antimycobacterial purines. J Med Chem 45: 1383-1386. | 18. Antonio, M., Asunción, B., Beatriz, S., Esther, V., Ignacio, A., Lidia, M.L., Raquel, V., Saioa, A., Silvia, P.S. (2009) Selective activity against Mycobacterium tuberculosis of new quinoxaline 1,4-di-N-oxides. Bioorg Med Chem 17: 385–389. | 19. Adams, L., Anathan, S., Biswas, A., Brooks, J., Franzblau, S., Krahenbuh, J., Kwong, C., Maddry, J., Michael, M., Miller, L., Orme, I., Poffenberger, A., Reynolds, R., Rouse, D., Secrist, J.; Winfield, D. (2001) Search for new drugs for treatment of tuberculosis. Antimicrob Agents Chemother 45: 1943-1946. | 20. Arutyunyan, S., Houghten, R.A., Jon, A., Nefzi, A. (2009) Parallel synthesis of chiral pentaamines and pyrrolidine containing bis-heterocyclic libraries. Multiple scaffolds with multiple building blocks: a double discovery for the identification of new antitubercular compounds. Bioorg Med Chem 19: 5169-5175. | 21. CS ChemOffice is software of CambridgeSoft Corporation, USA. | 22. Yap, C.W. (2011) PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints. J Comp Chem 32: 1466-1474. | 23. Roy, K., Roy, P.P. (2009) Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise, MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. Europ J Med Chem 44: 2913-2922. | 24. Cochran, W.G., Snedecor, G.W. (1967) Statistical Methods. Oxford & IBH, New Delhi, India. | 25. Eriksson, L., Wold, S. (1995) Validation tools. In: Waterbeemd, H. van de, editor. Chemometric Methods in Molecular Design. Weinheim, VCH, p. 312 – 317. | 26. Debnath, A.K. (2001) Quantitative structure-activity relationships: a versatile toll in drug design. In: Ghose, A.K., Viswanadhan, V.N., editors. Combinatorial Library Design and Evaluation. Marcel Dekker, New York, USA, p. 73 – 129. | 27. Roy K. (2007) On some aspects of validation of predictive quantitative structure-activity relationship models. Expert Opin Drug Discov 2: 1567-1577. | 28. Blair, R.M., Branham, W. S., Dial, S.L., Fang, H., Moland, C.L., Perkins, R., Sheehan, D.M., Shi, L.M., Tong, W., Wu, J. (2001) QSAR models using a large diverse set of estrogens. J Chem Inf Comput Sci 41: 186-195. | 29. Ebert, R.U., Chen, J., Kuhne, R., Schuurmann, G., Wang, B. (2008) External validation and prediction employing the predictive squared correlation coefficients, test set activity mean Vs training set activity mean. J Chem Inf Model 48: 2140-2145. | 30. Ballabio, D., Consonni, V., Todeschini, R. (2009) Comments on the definition of the Q2 parameter for QSAR Validation. J Chem Inf Model 49: 1669-1678. | 31. Roy, K., Roy, P.P. (2008) Comparative QSAR studies of CYP1A2 inhibitor flavonoids using 2D and 3D descriptors. Chem Biol Drug Des 72: 370-382. | 32. Mitra, I., Paul, S., Roy, K., Roy, P.P. (2009) On two novel parameters for validation of | predictive QSAR models. Molecules 14: 1660-1701. | 33. Chirico, N., Gramatica, P. (2011) Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. J Chem Inf Model 51: 2320-2335. | 34. http://203.200.173.43:8080/rmsquare/ (access on: 15/12/2012) | 35. Darlington, R. B. (1990) Regression and Linear Models. McGraw-Hill, New York, USA. | 36. Franke, R. (1984) Theoretical Drug Design Methods, Elsevier, Amsterdam, Netherland. | 37. Franke, R., Gruska, A. (1995) Chemometric methods in molecular design. In: Van de Waterbeemd H., editor. VCH, Weinheim, p. 113-163. | 38. Wold, S. (1995) Chemometric methods in molecular design. In: Van de Waterbeemd H., editor VCH, Weinheim, p. 195-218. | 39. SPSS is statistical software of SPSS Inc., IL, USA. | 40. MINITAB is statistical software of Minitab Inc., USA. | 41. Golbraikh, A., Tropsha, A. (2002) Beware of q2!. J Mol Graph Model 20: 269-276. | 42. Das, R., Mitra, I., Ojha, P.K., Roy, K. (2011) Further exploring metrics for validation of QSPR models. Chemom Intell Lab Syst 107: 194-205. | 43. Das, R.N., Kabir, H., Kar, S., Mitra, I., Ojha, P.K., Roy, K. (2012) Comparative studies on some metrics for external validation of QSPR models. J Chem Inf Model 52: 396-408. |