



Received on 31 October, 2013; received in revised form, 21 February, 2014; accepted, 13 March, 2014; published 01 April, 2014

GQSAR APPROACH TO STUDY THE EFFECT OF VARIOUS GROUPS ON SUBSTITUTED QUINOLINES SHOWING ANTI-TUBERCULOSIS ACTIVITY

B.J. Mylliemngap¹, A. Borthakur², K. Ingale³, S. Karanam³, D. Velmurugan⁴, A. Bhattacharjee*¹

Department of Biotechnology and Bioinformatics, North Eastern Hill University¹, Shillong-793022, Meghalaya, India

Division of Microbiology, Indian Agricultural Research Institute², Pusa, New Delhi-110012, India

Vlife Sciences Technologies Pvt. Ltd.³, Baner, Pune-411045, Maharashtra, India

Centre for Advanced studies in Crystallography and Biophysics, University of Madras⁴, Guindy (Maraimalai) Campus, Chennai- 600 025, Tamil Nadu, India

Keywords:

Anti-tuberculosis, GQSAR, quinolines

Correspondence to Author:

A. Bhattacharjee

Department of Biotechnology and Bioinformatics, North Eastern Hill University, Shillong-793022, Meghalaya, India

E-mail: atanubioinfo@gmail.com

ABSTRACT: In this paper, an attempt was made to develop a Group-Quantitative Structure Activity Relationship (GQSAR) model on a series of substituted Quinoline derivatives as anti-Tubercular agents. The dataset was fragmented based on substitution points on the quinolone ring. Various 2D descriptors were calculated and used in the present analysis. For model validation, the dataset was divided into various training and test sets using sphere exclusion method. The developed G-QSAR models were found to be statistically significant with respect to training ($r^2 > 0.7$), cross-validation ($q^2 > 0.6$), and external validation ($\text{pred}_r^2 > 0.5$). The developed GQSAR model suggests that substitutions on fragment R4 have an important role in determining biological activity.

INTRODUCTION: Each year more lives are lost due to *Mycobacterium tuberculosis* causing the deadly disease tuberculosis affecting the lungs (pulmonary TB) and other sites as well (extrapulmonary TB). There are millions of new cases of tuberculosis being diagnosed with 1.45 million deaths occurring mainly in developing countries¹. With the available treatment such as direct observed treatment short-course (DOTS) making a major progress in TB care and control, the tubercle bacillus is able to survive and continues to claim more lives¹⁻².

The extensive appearance of drug resistant strains with the deadly synergy of human immunodeficiency virus (HIV) has increased the incidence of tuberculosis in both developing and industrialized countries³. TB and HIV infections complement each other and approximately 0.35 million adults are infected with both pathogens¹.

The enigma of its dormancy and capability of infection in the latent phase are the prime reasons for which most of the treatments have failed against it as a result of which one third of the world population is infected⁴.

The emergence of drug resistance due to the extensive period of treatment in which patients fail to complete the therapy and the unique property of mycobacterium to become persistent or dormant over very long periods indicate an urgent need to develop new anti-tuberculosis drugs⁵⁻⁶.

<p>QUICK RESPONSE CODE</p> 	<p>DOI: 10.13040/IJPSR.0975-8232.5(4).1345-51</p> <hr/> <p>Article can be accessed online on: www.ijpsr.com</p> <hr/> <p>DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.5(4).1345-51</p>
-----------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Quinolines exhibit potent *in vitro* and *in vivo* anti-tubercular activity by both bacterial type II topoisomerase, DNA gyrase and topoisomerase IV, which are essential enzymes in catalyzing DNA supercoiling and decatenation reactions⁷.

The present study is aimed to understand the structure activity of a few quinolines and obtaining a predictive quantitative structure activity relationship (QSAR) models. Among the present quinolines used as drugs for TB, Diarylquinoline R207910 (called as DARQ), targeted to the protonpump of bacterial ATP synthase has inhibited mycobacterial growth very effectively⁸. With the aim of developing new anti-TB drugs, quinoline class of compounds has shown promising drug potentiality by inhibiting both replicating and non-replicating mycobacteria⁹. This study would help in optimizing the structural-chemistry of various quinoline derivatives so as to enhance their anti-TB activity for the development as drugs later on. A stochastic search method such as an artificial neural network, k-Nearest Neighbor-Molecular Field Analysis (kNN-MFA), multiple regressions (MR) and partial least square regression (PLSR) was developed to provide an insight to the various interactive fields of different compounds in concordance with the *in vitro* experimental data.

QSAR approaches have been developed through the years and a more recent QSAR method Group-Based QSAR (G-QSAR) proposed to use

descriptors by evaluating for the fragments of the molecules generated using specific fragmentation rules defined for a given dataset¹⁰ is used to generate models that could provide a valuable reference in the design of pharmaceuticals with improved anti-tubercular activity.

MATERIALS AND METHODS: The current study of GQSAR study was performed using V-Life MDS version 4.1. GQSAR method utilizes fragment based descriptors of molecules, for the current dataset a common pattern based on substitutions was identified and the molecules were fragmented as shown in (Fig. 1a).

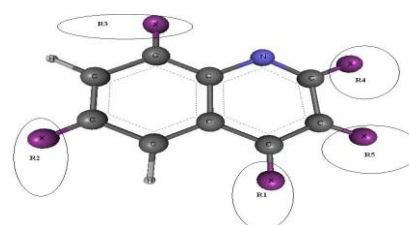
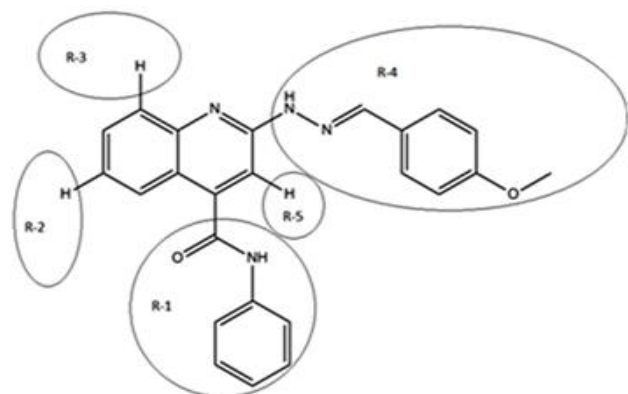


FIG. 1A: COMMON TEMPLATE FOR FRAGMENTATION OF THE COMPOUNDS

Dataset for G-QSAR study: Substituted Quinolines with known anti-tuberculosis activity were considered for the GQSAR study, activity of the compounds was taken from reported work given as (Minimum Inhibitory Concentration) MIC values¹¹⁻¹⁴. The biological activity (MIC) of the molecules were converted to their corresponding pMIC (-logMIC) values and used as dependent variables in the QSAR calculations (Table 1).

TABLE 1: STRUCTURE, IUPAC NAME AND BIOLOGICAL ACTIVITIES OF THE MOLECULES FOR GQSAR

Sr. No	IUPAC name	MIC (μM or $\mu\text{g/ml}$)	pMIC (μM)
1	(2,8-Bis-trifluoromethylquinolin-4-yl)-(2-piperidin-1-yl-ethyl) amine	3.13	-0.903
2	6-bromo-3-(imidazol-1-yl-phenyl-methyl)-2-methoxy-quinoline	6.25	-1.20
3	6-Bromo-2-methoxy-3-{phenyl-[4-(3-trifluoromethylphenyl)-piperazin-1-yl]-methyl}-quinoline	6.25	-1.051
4†	6-Bromo-2-methoxy-3-(phenyl-pyrazol-1-ylmethyl)-quinoline	6.25	-1.20
5	6-[[[(6-Bromo-2-methoxy-quinolin-3-yl)-phenylmethyl]-amino]-chromen-2-one	6.25	-1.114
6	N4-Phenyl-2-(N ⁴ -4-methoxybenzylidenehydrazino)-4-quinolinecarboxamide	6.25	-1.198
7	2-[(E)-{2-[2,8-Bis(trifluoromethyl)quinolin-4-yl] hydrazinylidene} methyl]-5-fluorophenol	50	-1.699
8	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2-(3-(1,1,2,2- tetrafluoroethoxy) benzylidene) hydrazine	12.5	-1.097
9	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2-(3-hydroxy- 4-methoxybenzylidene) hydrazine	12.5	-1.097
10	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2-(2-fluoro-4- methoxybenzylidene) hydrazine	6.25	-0.796
11	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2-((thiophen- 2-yl)methylene) hydrazine	50	-1.699
12	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2- ((pyridin-3-yl)methylene) hydrazine	12.5	-1.097
13	(E)-1-(2,8-Bis(trifluoromethyl)quinolin-4-yl)-2- (cyclohexylmethylene) hydrazine	6.25	-0.796
14	(E)-1-(2,8-Bis(trifluoromethyl) quinolin-4-yl)-2- (3-(trifluoromethoxy)benzylidene)	12.5	-1.097



iii

FIG. 1B: FRAGMENTATION PATTERN BASED ON THE TEMPLATE FOR (i) COMPOUNDS 2 TO 5 (ii) 1 AND 7 TO 35 (iii) COMPOUND 6

Calculation of fragment based 2D descriptors:

Descriptors for QSAR model generation were then calculated for each fragment of a given molecule, which include 2D and 3D descriptors falling under the class of Physico-Chemical and Alignment Independent descriptors. Invariable columns (those columns that had the same numerical value throughout and thus invariable) not contributing to the QSAR were removed.

Data selection for GQSAR: In order to generate a QSAR model biological activity was selected as a dependent variable and the fragment based 2D descriptors as independent variables. The training and test sets were generated using the sphere exclusion method which allows construction of the training set covering all the descriptor space areas occupied by representative points¹⁶ and the dissimilarity value was set to 2.7. The unicolon statistics of the training and test sets was calculated to ensure correct classification of the training and test sets.

G-QSAR model building: A pool of descriptors were selected for building a quantitative model, and after removing the invariability a total of 565 descriptors encompassing all the fragments of the molecules was finally retained to build the QSAR model. Variable selection methods coupled with a various model building method such as k-nearest-neighbor (kNN), partial least square regression (PLSR) and multiple regression (MR) were

selected from GQSAR model building. Cross-correlation limit as 0.5 with the term selection criteria as r^2 for MR, PLSR and q^2 for kNN-MFA and F-test 'in' value 4.0 and F-test 'out' value 3.0 were used as additional parameters for Stepwise variable selection method. The number of nearest neighbors in kNN was within the range of 2-5 and the distance-based weighted average as a method for prediction of biological activity was selected for the study.

Simulated annealing was used as variable selection method for PLSR (SA-PLSR) method with variance cut-off of 0.5 and cross correlation of 0.7 with maximum and minimum temperature of 100 and 0.01, respectively.

RESULTS AND DISCUSSION: Substituted quinolines reported to show anti-TB activity (pMIC) in terms of micro molar concentrations were taken for GQSAR study. Fragmentation in GQSAR is done to obtain site specific clues for designing of new molecules. The congeneric series of molecules was fragmented based on a common template (**Fig. 1a**) resulting in five fragments.

Fragment based 2D descriptors like Physicochemical, Chi, Chiv, Kappa, estate counts, Estate indices and topological alignment independent descriptors were calculate for all five fragments. The data was classified in training and test set using sphere exclusion method, training set comprised of 30 compounds while the test set of 5 compounds. The uni-column statistics study showed that the maximum of the test was less than the maximum of training and minimum of test was greater than the minimum of training set (**Table 2**) which also indicated that the test set was interpolative/ subset of the training set.

The mean of the test set was higher than the mean of training set and showed the presence of relatively more active molecules compared to the inactive ones. Also, a relatively higher standard deviation in the training set indicated that the training set had a widely distributed activity as compared to the test set.

TABLE 2: UNI-COLUMN STATISTICS

Column Name	Average	Max	Min	StdDev	Sum
pMIC_training	-1.1164	-0.4940	-1.6990	0.3287	-3.4910
pMIC_test	-0.8766	-0.4940	-1.2000	0.2795	-4.3830

Multiple QSAR models as shown in table 3 were generated coupling variable selection methods like StepWise variable selection and Simulated Annealing method with model building methods like Multiple regression, Partial least square regression, and k-nearest neighbor. The Multiple Regression and Partial least square model show less statistical model, the estimated accuracy of 78.64% ($r^2 = 0.7864$) and 78.57% ($r^2 = 0.7857$), an internal validation accuracy of 60.61% ($q^2 = 0.6061$) and 60.15% ($q^2 = 0.6015$) but a very less external predictivity for the external test set of 17.61% ($\text{pred}_r^2 = 0.1761$) and 20.52% ($\text{pred}_r^2 = 0.2052$) respectively.

The same data when subjected to kNN method resulted in a statistically significant model with q^2 around 0.6996 and good external predictivity of the test set with $\text{pred}_r^2 = 0.5347$. A model using random method of variable selection, simulated annealing, coupled with PLSR to improve results was statistically significant in terms of activity-descriptors relationship accuracy of 83.63% (0.8363) and an internal predictivity of 48.56% ($q^2 = 0.4856$). The external predictive ability of 51.54% ($\text{pred}_r^2 = 0.5154$) (**Table 3**).

TABLE 3: MODEL STATISTICS OF THE QSAR ANALYSIS

	MR	PLSR	kNN	SA-PLSR
Optimum Components	NA	3	NA	4
n	30	30	30	30
Degree of freedom	23	26	24	25
r ²	0.7864	0.7857	---	0.8363
q ²	0.6061	0.6015	0.6996	0.4856
F test	14.1156	31.7736	---	31.9200
r ² se	0.1705	0.1607	---	0.1432
q ² se	0.2316	0.2191	0.1801	0.2539
pred_r ²	0.1761	0.2052	0.5347	0.5154
pred_r ² se	0.3515	0.3453	0.2642	0.2696
Selected descriptors	R4-T_N_O_4 R4-T_2_S_1 R4-SdssCE-index R4-SaaCHE-index R4-T_2_N_1 R4-SaaaCE-index	R4-T_N_O_4 R4-T_2_S_1 R4-SdssCE-index R4-SaaCHE-index R4-T_2_N_1 R4-SaaaCE-index	R4-T_C_N_4 R4-SaaNHE-index R2-HydrogensCount R4-T_N_O_4 R4-T_2_N_0	R4-T_N_O_5 R5-chiV0 R4-T_C_N_5 R4-chiV3Cluster R2-PolarSurfaceAreaExcludingPandS R4-T_N_O_4 R4-T_2_F_2 R2-FluorinesCount R4-SaaSE-index R2-T_2_C_7

The experimental and predicted activity of the compounds of training and test set are reported in **Table 4**. The nearness in experimental and predicted activity of the model was shown as the plots for SA-PLSR (**Fig. 2 and 3**). The robustness of the models was also evaluated based on the F-value and the standard error values. The F-value of SA-PLSR model was 31.92 indicating the model is significant. The SA-PLS model generated had 4 optimum components, and 10 descriptors. The contribution plot of all descriptors is shown in **fig. 4**. The contribution plot shows that the contribution of fragment R4 is maximum (~80 %) followed by fragment R2 (~14%) and R5 (~6%), indicating modifying substitution at R4 can improve biological activity significantly.

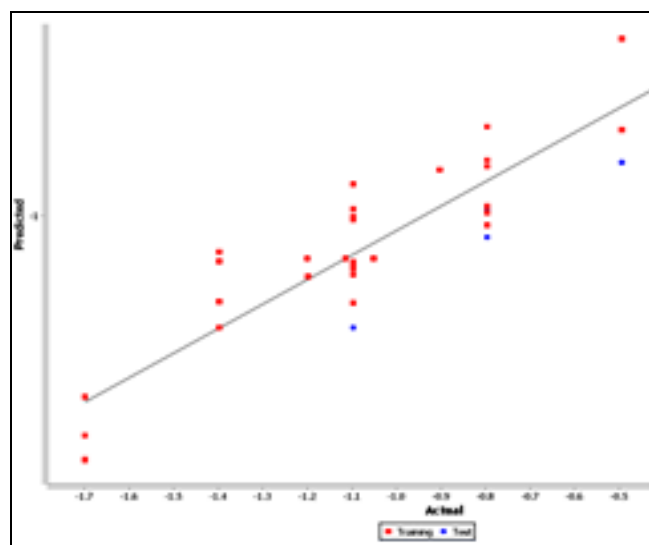


FIG. 2: FITNESS PLOT FOR SA-PLSR

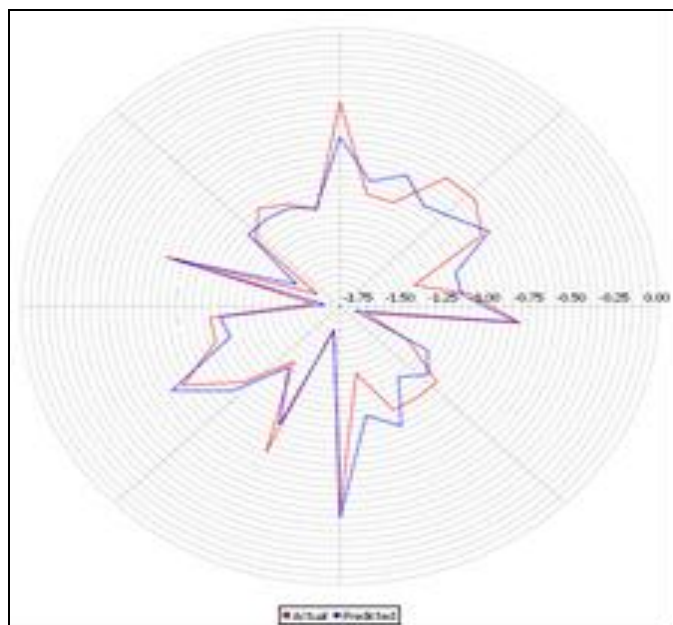


FIG. 3: ACTUAL AND PREDICTED ACTIVITY DISTRIBUTION OF THE TRAINING SET GENERATED BY SA-PLSR MODEL

Descriptors like ChiV3Cluster (~8 %), T_2_F_2 (~5%) and SaaSE-Index (~15%) at fragment R4 are positively contributing, while T_N_O_5 (~6%), T_C_N_5 (~25%) and T_N_O_4 (~22%) are negatively contribution towards activity. The contributions suggest decreasing the distance between Nitrogen atom and Oxygen atom below 4 bonds while increasing the aromatic fluorine and sulfur atoms can increase biological activity.

Contributions descriptors at Fragment R2 show fluorine count is positively contributing while Polarsurfacearea excluding Phosphorus and Sulfur and T_2_C_7 are negatively contributing, indicating increasing fluorine atoms without increasing polar surface area. Fragment R5 has very low contribution towards activity, descriptor chiV0 (~6%) at fragment R5 is negatively contributing towards activity.

TABLE 4: EXPERIMENTAL AND PREDICTED ACTIVITY OF THE COMPOUNDS

Compounds	pMIC	Pred_MR	Pred_PLS	Pred_kNN	Pred_SA-PLS
1	-0.903	-1.06689	-1.06366	-0.796	-0.85552
2	-1.2	-1.06689	-1.06366	-1.114	-1.13469
3	-1.051	-1.06689	-1.06366	-1.114	-1.13469
4 †	-1.2	-1.06689	-1.06366	-1.114	-1.13469
5	-1.114	-1.06689	-1.06366	-1.051	-1.13469
6	-1.198	-1.27975	-1.27162	-1.114	-1.19103
7	-1.699	-1.67396	-1.68676	-1.699	-1.76895
8	-1.097	-1.10398	-1.0999	-1.097	-1.16657
9	-1.097	-1.09587	-1.09197	-1.097	-1.14436
10	-0.796	-1.07683	-1.07337	-1.097	-0.99007
11	-1.699	-1.699	-1.71281	-1.699	-1.56834
12	-1.097	-1.01036	-1.00843	-0.796	-1.18503
13	-0.796	-0.88938	-0.89023	-1.097	-0.72043
14	-1.097	-1.11424	-1.10992	-1.097	-1.27379
15	-0.796	-0.88938	-0.89023	-0.796	-1.02884
16	-1.699	-1.66458	-1.6776	-1.699	-1.76498
17	-0.494	-0.66894	-0.67486	-0.796	-0.73014
18	-0.796	-0.70659	-0.71165	-0.796	-0.84508
19	-1.398	-1.13975	-1.13484	-1.097	-1.26901
20	-1.097	-0.95665	-0.95595	-1.097	-1.15453
21	-1.097	-1.19284	-1.18759	-1.097	-1.00197
22	-1.097	-1.01351	-1.01264	-1.097	-1.35181
23 †	-0.796	-1.46051	-1.44965	-1.097	-1.06706
24	-1.699	-1.75847	-1.76422	-1.699	-1.69
25 †	-0.494	-0.60178	-0.60925	-0.494	-0.44352
26	-1.398	-1.40613	-1.39639	-1.097	-1.3515
27	-1.097	-1.17386	-1.16011	-1.097	-0.90056
28	-1.398	-1.21717	-1.20275	-1.097	-1.11309
29 †	-1.097	-1.2294	-1.21523	-1.398	-1.01149
30	-0.796	-1.15694	-1.14549	-1.097	-0.97028
31	-1.398	-0.97417	-0.96846	-1.699	-1.14332
32	-0.494	-0.56383	-0.58505	-0.796	-0.83228
33	-0.796	-0.80547	-0.82149	-0.796	-0.8251
34	-1.097	-0.99186	-1.0021	-1.097	-0.97945
35 †	-0.796	-0.71925	-0.73485	-1.097	-0.97945

† Compounds of Test Set

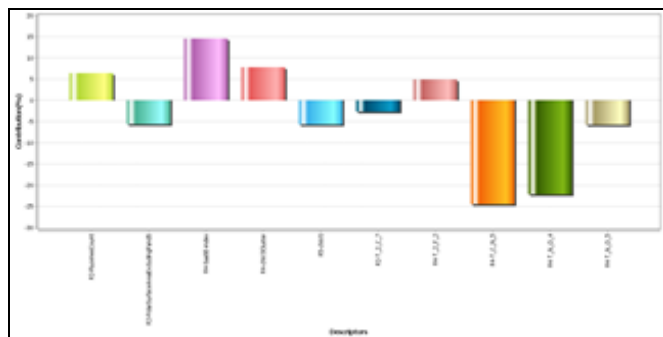


FIG. 4: CONTRIBUTION OF VARIOUS DESCRIPTORS IN THE SA-PLSR MODEL.

Detailed analysis of all the generated QQSAR models indicate that fragment R4 is a major contributor towards anti-tubercular activity. In fragment R4, nitrogen atom and its connectivity towards other atoms is decisive feature, also it has been seen that presence of aromatic ring at fragment R4 is important.

CONCLUSION: Current study includes GQSAR analysis of substituted quinoline derivatives; the GQSAR models generated show statistically significant and robust model with good external predictive ability. Model SA-PLSR provides site specific clues for design of compounds with better activity.

ACKNOWLEDGEMENT: The authors are thankful to the Department of Biotechnology (DBT), Government of India for providing the necessary fund and facilities needed for this work.

REFERENCES:

- World Health Organization: The sixteenth global report on tuberculosis, 2011.
- Bloom BR: Tuberculosis: pathogenesis, protection, and control. ASM Press, Washington D.C., Edition 1, 1994: 3-11.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S and Barrell BG: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393(6685): 37-544.
- Dye C, Watt C J, Bleed D: Low access to a highly effective therapy: a challenge for international tuberculosis control. *Bull World Health Organ* 2002; 80(6):437-444.
- Bhattacharjee A, Mylliemngap BJ and Velmurugan D: 3D QSAR studies on fluoroquinolones derivatives as inhibitors for tuberculosis. *Bioinformation* 2012; 8(8):381-387.
- Chopra P, Meena LS and Singh Y: New drug targets for *Mycobacterium tuberculosis*. *Indian J Med Res* 2003; 117:1-9.
- Prajapati K, Singh S, Pathak AK and Mehta P: QSAR analysis on some 8-methoxy quinoline derivatives as H37RV (MTB) inhibitors. *International Journal of Chem Tech Research* 2011; 193:408-422.
- Andries K, Verhasselt P, Guillemont J, Gohlmann HW, Neefs JM, Winkler H, Van Gestel J, Timmerman P, Zhu M, Lee E, Williams P, de Chaffoy D, Huitric E, Hoffner S, Cambau E, Truffot-Pernot C, Lounis N and Jarlier V: A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. *Science* 2005; 307(5707):223-227.
- Lilienkampf A, Mao J, Wan B, Wang Y, Franzblau SG and Kozikowski AP: Structure-activity relationships for a series of quinoline-based compounds active against replicating and nonreplicating *Mycobacterium tuberculosis*. *J Med Chem* 2009; 52(7):2109-2118.
- Ajmani S, Jadhav K and Kulkarni SA: Group-Based QSAR (G-QSAR): Mitigating Interpretation Challenges in QSAR. *QSAR Comb Sci* 2008; 28(1):36-51.
- Mital A, Negi VS and Ramachandran U: Synthesis and antimycobacterial activities of certain trifluoromethyl-aminoquinoline derivatives. *Arkivoc* 2006; 10:220-227.
- Upadhyaya R S, Vandavasi JK, Vasireddy NR, Sharma V, Dixit SS and Chattopadhyaya J: Design, synthesis, biological evaluation and molecular modelling studies of novel quinoline derivatives against *Mycobacterium tuberculosis*. *Bioorg Med Chem* 2009; 17:2830-2841.
- Eswaran S, Adhikari AV, Chowdhury IH, Pal NK and Thomas KD: New quinoline derivatives: Synthesis and investigation of antibacterial and antituberculosis properties. *Eur J Med Chem* 2010; 45(8):3374-3383.
- Nayyar A and Jain R: Synthesis and anti-tuberculosis activity of 2, 4-disubstituted quinolines. *Indian J Chem* 2008; 47:117-128.
- Halgren T A: Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization and Performance of MMFF94. *J Comp Chem* 1996; 17:490-519.
- Golbraikh A and Tropsha A: QSAR modeling using chirality descriptors derived from molecular topology. *J Chem Inf Comput Sci* 2003; 43:144-154.

How to cite this article:

Mylliemngap BJ, Borthakur A, Ingale K, Karanam S, Velmurugan D, Bhattacharjee A: GQSAR approach to study the effect of various groups on substituted quinolines showing anti-tuberculosis activity. *Int J Pharm Sci Res* 2014; 5(4): 1345-51. doi: 10.13040/IJPSR.0975-8232.5(4).1345-51

All © 2013 are reserved by International Journal of Pharmaceutical Sciences and Research. This Journal licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

This article can be downloaded to **ANDROID OS** based mobile. Scan QR Code using Code/Bar Scanner from your mobile. (Scanners are available on Google Playstore)