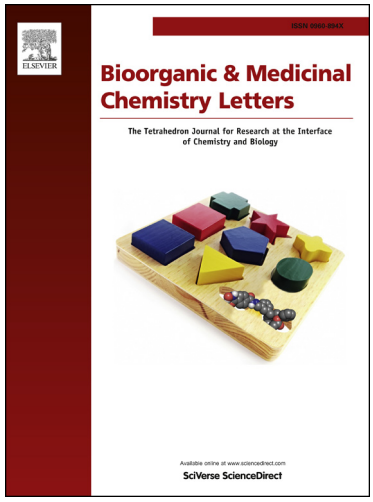


Synthesis And Binary QSAR Study of Antitubercular Quinolylhydrazides

Atul Manvar, Vijay Khedkar, Jignesh Patel, Vipul Vora, Narsinh Dodia, Gautam Patel, Evans Coutinho, Anamik Shah

PII: S0960-894X(13)00809-3
DOI: http://dx.doi.org/10.1016/j.bmcl.2013.06.076
Reference: BMCL 20640

To appear in: *Bioorganic & Medicinal Chemistry Letters*

Received Date: 14 March 2013
Revised Date: 8 June 2013
Accepted Date: 27 June 2013

Please cite this article as: Manvar, A., Khedkar, V., Patel, J., Vora, V., Dodia, N., Patel, G., Coutinho, E., Shah, A., Synthesis And Binary QSAR Study of Antitubercular Quinolylhydrazides, *Bioorganic & Medicinal Chemistry Letters* (2013), doi: http://dx.doi.org/10.1016/j.bmcl.2013.06.076

Synthesis And Binary QSAR Study of Antitubercular Quinolylhydrazides

Atul Manvar[a,b], Vijay Khedkar[d], Jignesh Patel[a], Vipul Vora[a], Narsinh Dodia[a], Gautam Patel[c], Evans Coutinho[d*] and Anamik Shah[a*]

[a]*National Facility for Drug Discovery, Department of Chemistry, Saurashtra University, Rajkot-360005, Gujarat, India.*
[b]*Present address: Centre for Synthesis & Chemical Biology, UCD School of Chemistry & Chemical Biology, University College Dublin, Belfield, Dublin 4, Ireland.*
[c]*HDH Pharma Inc., Morrisville, NC-27560.*
[d]*Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400098, India.*
[*]*Corresponding authors:*
Anamik Shah
Tel/fax: +91 281 2581013/2589609; Email: anamik_shah@hotmail.com
https://sites.google.com/site/anamikkantilalshah/home; http://chemdive.com/

Evans Coutinho
Tel: +91 2226670905, Email: evans@bcpindia.org

Abstract

In continuation with our previous work in anti-TB research area, in the present study we have demonstrated the structural diversity of quinolylhydrazides as potent antituberculars. The compound library was synthesized by molecular hybridization approach and tested *in vitro* against *M. tuberculosis* H_{37}Rv strains. Among the designed conjugates, the most promising molecules were found to exhibit 100% Growth Inhibition (GI) at MIC <6.25 µg/mL. Moreover, several analogs in the designed series were also turned out as excellent antituberculars. To probe the structural characteristics influencing on the SAR, the classification model was generated using a binary QSAR approach termed recursive partitioning (RP) analysis. The significant features outlined by the RP model act as a guide in order to design the 'lead' compound.

Keywords: Quinolones, Hydrazides, *Mtb*, H_{37}Rv, MABA, Binary QSAR

Mycobacterium tuberculosis (Mtb), the pathogen remains one of the most fatal infectious disease followed by the AIDS.[1-3] According to WHO statistics one third of the world's populations have been exposed to TB bacterium and new infections occurs at a rate of one per second.[4-7] Control and prevention of TB is a major task nowadays. Despite the need for better TB therapies no new clinical candidates have been developed in the last few decades.[8,9] Currently, TB is treated with a schedule of four drugs combination (e.g. isoniazide, rifampin, pyrazinamide, and ethambutol) and the treatment lasting up to 6-9 months.[10,11] There are several demerits of the existing treatment, such as exceedingly lengthy therapy, host toxicity, ineffectiveness against resistance strains etc.[12] Moreover, the drugs utilized for curing TB shows potential side effects such as thrombocytopenia, neuropathy, rashes fever, and drug induced hepatitis.[13] Furthermore, the evolution of its new virulent forms, multidrug resistant (MDR-TB) and extensively drug resistant (XDR-TB) have become a major threat of mankind.[14] The XDR-TB is virtually untreatable using current therapeutics and without strengthening of the current TB controls measures.[15] The significant challenges for TB control are increasing number of immunocompromized individuals with HIV infections, who are highly susceptible to the disease.[16] Consequently, there is a pressing need to develop a novel, potent, and fast-acting antituberculosis drugs having minimal toxicity profile that would administered in conjunction with antireteroviral drugs.[17,18] Currently, the global TB development pipeline has nine drugs.[19] Furthermore, various quinoline based motifs are also at preclinical and clinical stages for TB drug development. The fluoroquinolones, such as gatifloxacin and moxifloxacin targets DNA topoisomerase IV and DNA gyrase can be utilized as anti-TB agents, however, they often suffer from the resistance. The TMC207 is a highly potent anti-TB agent which is currently in phase II clinical trials **(Figure 1).**[20-22]

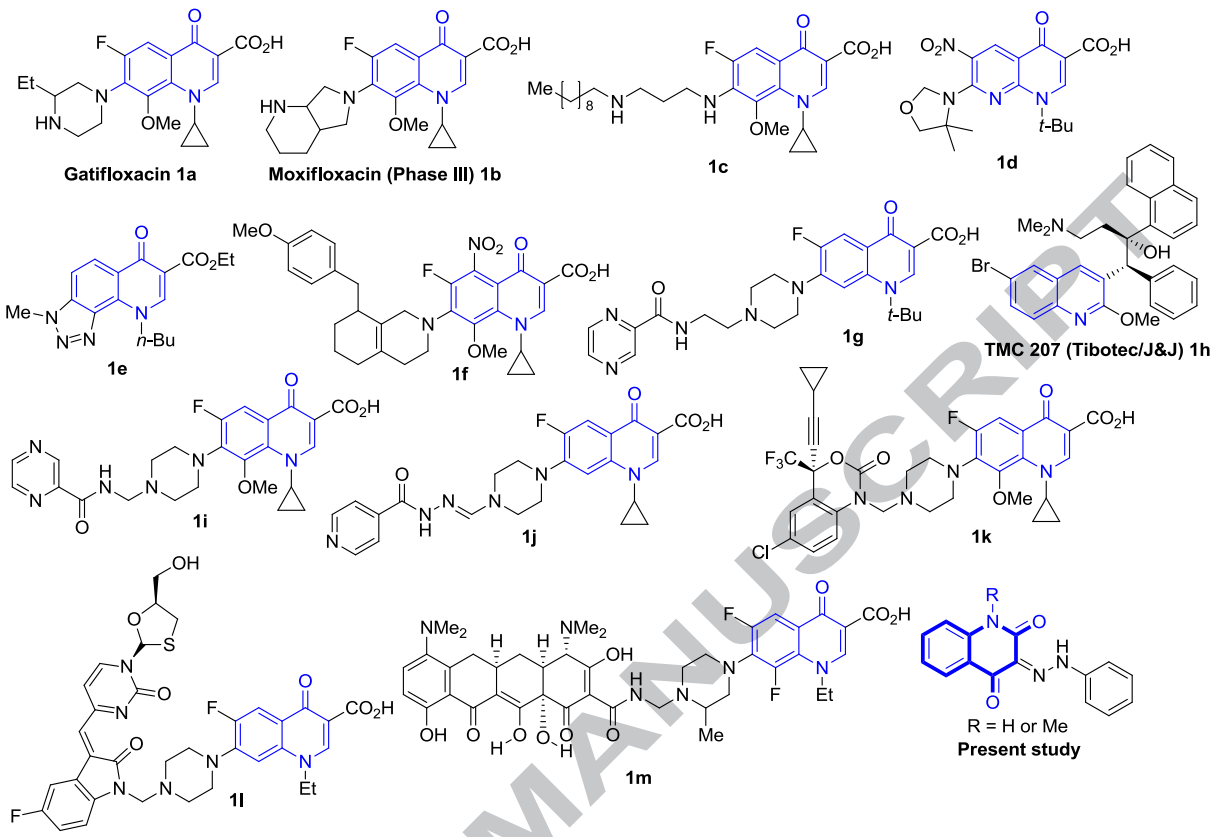
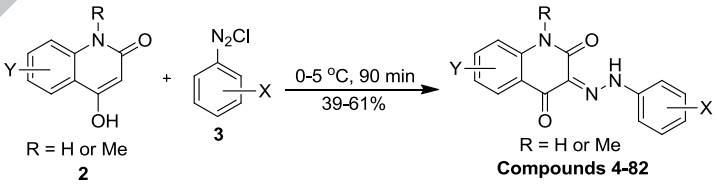


Figure 1: Quinoline based scaffolds at various stages for TB drug development (1).

In this context our group has investigated a wide range of heterocycles[23-27] as promising antituberculars. Among them the 3D-QSAR study and other *in silico* analysis based rational design of tetrahydropyrimidine[24] was identified as a 'lead' molecule. Inspired by these findings, in the current paper, we have envisioned for the synthesis of quinolylhydrazides as potent anti-TB agents by fragment based approach.

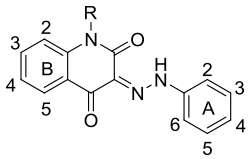


Scheme 1: Synthesis of 3-(2-phenylhydrazono)quinoline-2,4-(1*H*,3*H*)-dione derivatives.

At the outset, the synthetic strategy for the preparation of molecularly diverse hydrazides of quinolones is depicted in **Scheme 1**. The target molecules were prepared by straightforward Meerwein arylations.[28] The 4-hydroxyquinolones bearing electron-rich, electron-deficient as well as sterically hindered substituents were efficiently coupled with the corresponding diazonium

salts at lower temperature, afforded compounds **4-82** in moderate to good yields with a very high functional group tolerance (see supporting informations). The antimycobacterial activities of synthesized derivatives were assessed by employing the microplate alamar blue assay (MABA) against *Mtb* $H_{37}Rv$ strains[29] utilizing rifampicin as a reference drug.

Table 1: *In vitro* activity of compounds **4-82** against *M. tuberculosis* $H_{37}Rv$ strains.



Compounds 4-82

Compds	Substitutions			[a]GI (%)	[b]MIC (μg/mL)	[c]CLogP	Compds	Substitutions			[a]GI (%)	[b]MIC (μg/mL)	[c]CLogP
	R	Ring A	Ring B					R	Ring A	Ring B			
4	H	4-Br	3-NO_2	100	<6.25	2.643	**44**	H	4-F	3-Cl	50	<6.25	2.893
5	H	2-Ph	3-NO_2	100	<6.25	3.668	**45**	H	2,3-di Me	2-NO_2	48	>6.25	2.728
6	H	H	2-Cl	100	<6.25	2.750	**46**	H	3-Cl	3-Cl	47	>6.25	3.463
7	H	4-F	3-NO_2	100	<6.25	1.923	**47**	H	4-F	2-NO_2	46	>6.25	1.923
8	H	H	2-NO_2	99	<6.25	1.780	**48**	H	2-Me, 4-OMe	2-NO_2	45	>6.25	2.198
9	H	4-Cl	2-NO_2	99	<6.25	2.493	**49**	H	3-Me	H	42	>6.25	2.536
10	H	2-Me, 3-Cl	3-NO_2	98	<6.25	2.728	**50**	H	3-Cl	H	41	>6.25	2.750
11	H	4-Cl	4-Cl	98	<6.25	3.463	**51**	H	2-Me, 4-OMe	2,5-di Cl	40	>6.25	3.881
12	H	4-COMe	2, 3-di Cl	98	<6.25	2.782	**52**	Me	2-Me, 4-OMe	2-NO_2	40	>6.25	3.074
13	H	H	H	97	<6.25	2.037	**53**	H	3-Cl	2-Me	39	>6.25	3.249
14	Me	2-Me	H	97	<6.25	3.412	**54**	Me	4-Cl	H	38	>6.25	3.626
15	H	2-OMe	3-NO_2	96	<6.25	1.699	**55**	H	4-OMe	2-NO_2	37	>6.25	1.699
16	H	3-NO_2	4-Cl	96	<6.25	2.493	**56**	H	4-OMe	2,5-di Cl	36	>6.25	3.382
17	H	3-Me	2-NO_2	95	<6.25	3.035	**57**	H	2-Ph	2-NO_2	35	>6.25	3.668
18	H	3, 4-di Cl	4-Cl	95	<6.25	4.056	**58**	H	4-Me	2-Cl	34	>6.25	3.249
19	Me	2-NO_2	H	94	<6.25	3.412	**59**	H	3-Cl	2,5-di Cl	32	>6.25	4.176
20	H	4-Me	3-NO_2	92	<6.25	2.279	**60**	H	2-OMe	2-NO_2	32	>6.25	1.699
21	H	4-NO_2	4-Cl	91	<6.25	2.493	**61**	H	3-Me	2-Me	31	>6.25	3.035
22	H	4-NO_2	2, 3-di Cl	89	<6.25	3.086	**62**	H	2-Cl	3-Me	31	>6.25	3.249
23	H	4-COMe	2-NO_2	89	<6.25	1.219	**63**	H	4-COMe	4-Cl	29	>6.25	2.189
24	H	H	2, 3-di Cl	85	<6.25	3.343	**64**	H	3-NO_2	2,3-di Cl	27	>6.25	3.086
25	H	3-Cl, 4-F	2-NO_2	85	<6.25	2.636	**65**	H	4-COMe	2-Cl	24	>6.25	2.189
26	H	3-Cl	2-NO_2	82	<6.25	3.249	**66**	H	3-NO_2	2-Cl	24	>6.25	2.493
27	H	3-Cl	2, 3-di Cl	82	<6.25	4.056	**67**	H	2,6-di Me	2-NO_2	22	>6.25	2.778
28	H	2-NO_2	2-NO_2	80	<6.25	2.279	**68**	H	3,4-di Cl	2-Cl	21	>6.25	4.056
29	H	3-F	2, 3-di Cl	80	<6.25	3.468	**69**	H	2,3-di Cl	3-NO_2	21	>6.25	3.206
30	H	4-F	2, 5-di Cl	80	<6.25	3.606	**70**	Me	4-NO_2	H	20	>6.25	2.656
31	H	3, 4-di Cl	2-NO_2	79	<6.25	3.086	**71**	Me	4-COMe	H	20	>6.25	2.352
32	H	H	3-Cl	76	<6.25	2.750	**72**	H	3-Me	3-NO_2	19	>6.25	2.279
33	H	3-Me	3-Me	76	<6.25	3.035	**73**	H	4-OMe	2,3-di Cl	14	>6.25	3.262
34	H	4-F	4-Cl	71	<6.25	2.893	**74**	H	2,6-di Me	4-Cl	13	>6.25	3.748
35	H	4-Cl, 5-F	3-NO_2	70	<6.25	2.636	**75**	H	2-Me, 4-OMe	4-Cl	12	>6.25	3.168
36	H	4-NO_2	2-NO_2	65	<6.25	1.523	**76**	H	4-Me	2,5-di Cl	11	>6.25	3.962
37	H	2-NO_2	4-Cl	64	<6.25	2.493	**77**	Me	4,5-di Cl	H	10	>6.25	4.125
38	H	3-Me	2, 3-di Cl	62	<6.25	3.842	**78**	H	2-Cl	2-Cl	7	>6.25	3.463
39	H	4-Cl	2-NO_2	61	<6.25	2.493	**79**	H	4-COMe	3-Cl	6	>6.25	2.189
40	H	3-NO_2	2-NO_2	60	<6.25	1.523	**80**	H	4-OMe	3-Cl	4	>6.25	2.669
41	H	2-Me, 3-Cl	2-NO_2	56	<6.25	2.992	**81**	H	4-NO_2	3-Cl	1	>6.25	2.493
42	H	4-OMe	4-Cl	55	<6.25	2.669	**82**	H	H	4-OMe	0	>6.25	1.699
43	H	4-Me	2, 3-di Cl	54	<6.25	3.842							

[a]Growth Inhibition (GI) of virulent strains of *M. tuberculosis;* [b]MIC of Rifampicin: 0.015–0.125 μg/mL against *M. tuberculosis* $H_{37}Rv$ (97% inhibition); [c]CLogP is calculated on ChemDraw Ultra 12.0.

The structure activity relationship (SAR) of the molecular diversity is summarized in **Table 1.** The GI values were obtained in the range of 0-100% at MIC <6.25 and >6.25 µg/mL. The presence of versatile substituents on quinolylhydrazide skeleton was significantly affected on antitubercular activity. The unsubstituted analog **13** was found to be considerably potent (97% GI at MIC <6.25µM). Next, the influence of hydrophilic or lipophilic variants on the parent structure was probed. The analogs in which ring B constituted with NO_2 group, compounds **4, 6-10, 15, 17**, and **20** displayed superior activity. The compound **28** has shown 80% GI while, **36** and **40** exhibited almost 2-fold reductions in anti-TB activity. In the same context, the NO_2 containing quinolylhydrazides, for instance **41, 45, 47, 48, 55, 57, 60, 67, 69,** and **72** were found to be weakly actives. On the contrary, the presence of NO_2 group in ring A, in lieu of ring B, among the compounds **16, 21, 22, 37, 64, 66,** and **81**, only few analogs, such as **16, 21, 22** exhibited good activity. This trend has clearly indicated that not only the presence of NO_2 but its site-selectivity in the particular ring is also essential to generate a 'lead' molecule. Among N^1 methylated analogs, compounds **14** and **19** were proven to be actives. While, in comparison between **48** and **52** suggested that the *N*-methylation have not essential to design the potent antitubercular. We next scrutinized di-halogenated surrogates **11, 18, 27, 46, 59, 68,** and **78** have shown GI ranging from 7-98%, nevertheless CLogP were observed <5. The analogs **11** and **18** were found to be notably promising, while up to 4-fold reduction in anti-TB activity was observed within rest of the halogenated analogs. The incorporation of fluorine group may alter many of the physical properties of organic molecules, such as lipophilicity, metabolic stability, and conformational characteristics. For these reasons, we have rationally incorporated F group in conjunction with Cl either in ring A or B in compounds **29, 30, 34,** and **44**, nevertheless, no profound biological activities were noticed. Consequently, we may conclude that the blending of F and Cl was proven to be crucial in a search of prominent antitubercular of this class. In comparison between two sterically hindered analogs, compound **5** demonstrated 100% GI while, dramatic loss in activity was noticed in other regio-isomer **57**. When both rings comprising with Me group, compound **33** has shown satisfactory reduction in activity while **61** was found to be weakly potent. The molecules **38, 42, 43, 53, 62** exhibited moderate to poor activity. Notably promising compound **12** bearing COMe and Cl groups in rings A and B respectively. The least potent analog in the series, compound **82** constituted with OMe group in ring B, nevertheless CLogP is 1.699. In comparisons, the molecules **23, 63, 65,** and **79** in which 4-COMe in ring A

and NO_2 or Cl in ring B, a mere analog **23** was found active, whilst the biological profile of rest of the analogs were reduced up to 4-fold. The combination of OMe or Me and Cl, compounds **56, 58, 73-76** were found weakly potent. Finally, the effect of compound polarity was estimated by calculating the CLogP. The thumb rule for CLogP to a drug-like molecule must be less than '5' to by-pass a cell barrier. The CLogP seems to correlate some extent with lipophilicity and was found in the range of 1.219-4.176. Despite the lower potency, none of the analog has shown CLogP less than 5. Thus, from the SAR results we may confirm that the influences of electron density of the substituents, lipophilicity, as well as stereo-electronic properties are indispensable to design a 'lead' structure.

Next, to classify this dataset, a binary QSAR utilising the recursive partitioning (RP) analysis was performed. The RP is a simple statistical data analysis technique that seeks to decipher the elusive relationships in dataset involving thresholds, interactions, and non-linearities. The RP has an advantage of classifying biological activity data by considering an appropriate descriptors recursively. The non-linearity in the data impede the analysis that is based on linearity such as multiple linear regression (MLR), principal component analysis (PCA) or partial least squares (PLS) regression. The RP analysis is inherently faster than other grouping techniques for instance, clustering and the 2D nature makes facile for the data which is difficult to sieve in to usable divisions of classification. Furthermore, it has to be pointed out that the RP analysis is meant to identify the common characteristics of the binding modes in the data set and to recognize other compounds those fit the acquired rules.[30] The tree display of the SAR results help to identify the decisive factors that may categorize the dataset in to molecular classes with higher and lower potency. The algorithm accounts for all possible binary splits for each descriptor and to identify the optimal by splitting criterion–a Boolean function BN(M) of the molecular descriptors, returning 'true' or 'false'. The resulting subsets *recursively* split further with every new subset representing a *node* at which the splitting criterion is associated. The *branches* of the tree signify the truth values taken by the splitting criteria. The true response to any given splitting criteria follows a branch to the downside while the false response follows a branch to the upside in the partition tree. There can be several nodes consisting only 'active' molecules each representing different structural classes of known active. The goal is to produce the RP model with decision tree in which the final terminal nodes contain either only 'active' or

only 'inactive' molecules or in other words to produce terminal subsets of *maximal homogeneity* with respect to the *measured activities* of their members.

We probed a set of 79 molecules (compounds 4-82) for the RP analysis. The classification model was generated using the CSAR recursive partitioning technique incorporated in Cerius2 (version 4.8, Accelrys Inc)[31] software. The numerical descriptors that encode topological, geometric/structural, electronic, and thermodynamic properties were considered for deriving the RP model[32] and decision tree was generated from the results of the RP. The antitubercular activities of the molecules under study were described in terms of their growth inhibition (GI) values ranging from 0-100%. The entire dataset was split in to two classes as inactive (0) and active (1) with the class 1 containing 38 less active compounds having GI values <50% while the class 2 consisting 41 active compounds having GI values ≥50%. The correlation matrices were built for the descriptors and the descriptors with zero variance as well as columns containing 95% of zero values were eliminated. Moreover, the descriptors having cross-correlation coefficient >0.5 were eliminated as they have represented nearly same information. The resulting uncorrelated descriptors were used as independent variables (X) while the biological activity served as the dependent variable (Y) for the RP analysis. The activity classes were weighted equally (*weighing by classes)* wherein an active node is a node in which the fraction of actives exceeds the fraction of active molecules over the entire dataset (in case of *weighing by observables* a node with more than half of the members being active will be considered 'active'). The inactives categorized in an 'active' node while the actives classified in 'inactives' nodes are accounted as 'misclassified'. The splits were scored using Twoing rule scoring function to minimize the misclassification costs; the pruning factor was varied from 0 to 3; the values for maximum tree depth were varied between 5 and 10 while default values were used for maximum number of generic splits (30) and the number of knots per variable (20). The classification model was validated using cross validation, with the number of cross-validation groups set to 5.

The classification model was developed using 2D as well as 3D-physicochemical descriptors by varying the parameters discussed above seeking to improve the following parameters: 'Class % Observed Correct', a measure of intra-class prediction accounting for the compounds predicted correctly to be in respective classes as a percentage of a total number of compounds observed to be in each class; *'Overall % Predicted Correct',* so-called overall

prediction is the percent of ratio of a total number of compounds correctly classified to the number of compounds predicted in that class providing information on the accuracy of an overall prediction; the *'enrichment factor'* for a specific class is the ratio of percentage of compounds correctly predicted belong to that class over the original percentage of compounds belonging to that class. The statistical results of the RP analysis are summarized in **Table 2.**[33]

Table 2. Statistical results of recursive partitioning model.

Class	Value	Number of Compounds	%	Class % Obs Correct	Overall % Pred Correct	Enrichment
1	0	38	48.10	100	100	2.08
2	1	41	51.90	100	100	1.93

The results of the RP analysis were displayed as decision tree derived from the RP process splitting the whole dataset in to smaller subsets (nodes). The RP progresses sequentially examining all the descriptor variables to identify the best criterion for splitting the dataset in to 'active' or 'inactive' class. The 18-leaf RP decision tree was obtained for the quinolylhydrazides as depicted in **Figure 2.** The decision tree looks like the horizontal dendrogram formed with 18 terminal and 17 non-terminal nodes in which Class 1 (0) is plotted using Red while Class 2 (1) using Green. The terminal nodes 1, 3, 5, 7, 10, 12, 15, 17, and 18 represents class 2 (i.e. active) while the terminal nodes 2, 4, 6 8, 9, 11, 13, 14, and 16 represents class 1 (i.e. inactive). The analysis of decision tree shows that in class 1, i.e. the class of 'inactives', 38 out of the 38 compounds were correctly classified as belonging to class 1. For class 2 which is the class of 'actives', all 41 compounds were correctly classified as actives represented 100% intra-class predictivity. The overall prediction was represented by 'Overall % Predict Correct' was found to be 100% signified the accuracy of prediction when entire data set was being predicted with the RP model. The enrichment factor (2.08 for class 1 and 1.93 for class 2) also indicates that the RP model is statistically significant and can be used to classify new compound library. To evaluate the predictability and to avoid over fitting of data, the model was subjected to 5-fold cross-validation. This procedure leaves out 5% of the dataset and builds the model using remaining molecules which are then utilized to predict the activity of eliminated subset. The statistics of cross-validation are depicted in **Table 3.**

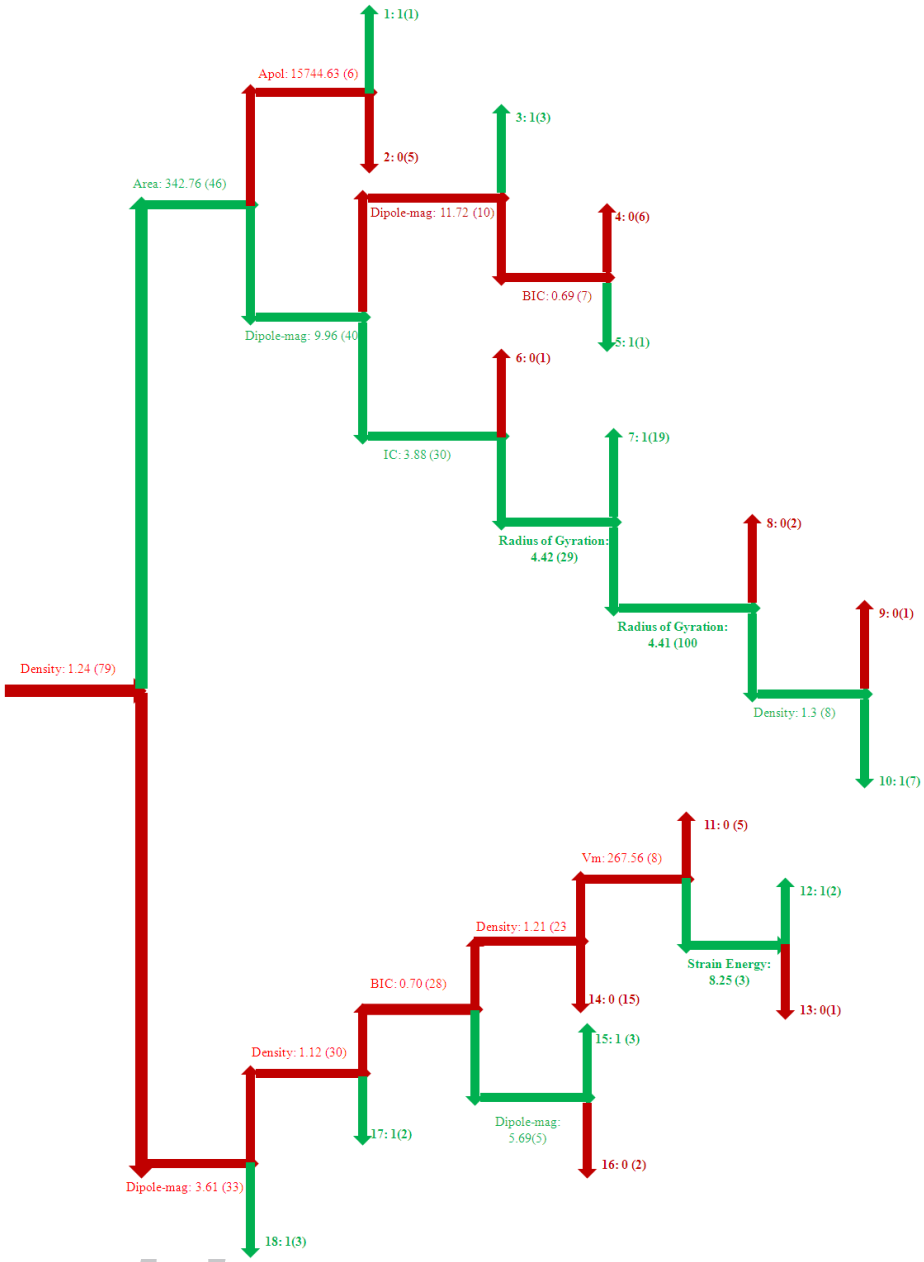


Figure 2. An 18-leaf decision tree derived from the Recursive Partitioning (RP) for prediction of antitubercular activity classes. Those marked 1:1, 2:0, 3:1, 4:0, 5:1, 6:0, 7:1, 8:0, 9:0: 10:1, 11:0, 12:1, 13:0, 14:0, 15:1, 16:0, 17:1 and 18:1 correspond to terminal nodes 1-18 and each terminal node corresponds to the value of 1 (active) or 0 (inactive).

Table 3. Statistical results of recursive partitioning model for 5-fold cross-validation.

Class	Value	Number of Compound	%	Class % Obs Correct	Overall % Pred Correct	Enrichment
1	0	38	48.10	52.63	55.56	1.16
2	1	41	51.90	60.98	58.14	1.12

The acceptable percentage of classification, 56% and 58% was accomplished for class 1 and class 2 with enrichment factor of 1.16 and 1.12, respectively indicated that the classification model would be able to classify and predict the activity (class) of new derivatives. The statistically significant cross-validation experiment signals for the given data set, a series of n partition trees can be derived based on subsets of size (1 - l/n), which correctly represents active apart from inactive in remaining l/n dataset. The key elements identified in the validation set were consistent with those from the entire set.

The first primary split was observed on the density (a spatial descriptor reflecting the type of atoms and how tightly they are packed in a molecule and also signify transport and melt behavior) and its cut-off value was 1.24. The dataset was split in to two branches-those molecules with density value less than 1.24 (compound **33**) followed a branch to downside while those (compound **46**) with density value above cut-off followed a branch to upside. The molecules with density greater than 1.24 were divided on the basis of area at the cut-off value of 342.76 with molecule (compound **6**) having an area greater than the cut-off values formed the branch to upside while those (compound **40**) with area less than 342.76 followed a branch downside. Six molecules with area greater than 342.76 were further split in to two branches on the basis of atomic polarizability (Apol). It is an electronic descriptor and is a measure of the relative tendency of charge distribution caused by the presence of nearby ion or dipole. It is also related to hydrophobicity and thus to the biological activity. The molecule **12** with Apol greater than 15744.63 formed the terminal node 1 and was correctly predicted as active, while the molecules **51, 56, 64, 73** and **76** formed the terminal node 2 was correctly classified as inactive by the RP model. Forty molecules with area less than 342.76 were split based on Dipole moment (Dipole-mag) with cut-off value of 9.96. The dipole moment is an electronic descriptor signifying the strength and orientation behavior of the molecule in an electrostatic field which can be related to the long-range ligand-receptor recognition and subsequent binding. Ten molecules **16, 18, 21, 30, 50, 59, 66, 68, 77,** and **81** with Dipole-mag greater than 9.96 followed a branch upside while those below cut-off formed a branch downside. These molecules were further split using the Dipole-mag with molecules having values greater than 11.72 formed the terminal node 3 while, the remaining seven molecules followed a branch downside. The terminal node 3 contains the three molecules **16, 18,** and **21** which were correctly classified as actives. The remaining seven molecules with Dipole-mag less than 11.72 but greater than 9.96 were

further split to form the terminal nodes 4 and 5 on the basis of BIC (Bonding Information Content). The molecules **50, 59, 66, 68, 77,** and **81** with BIC values greater than 0.69 formed the terminal node 4 and were predicted as inactive while molecule **30** having BIC less than 0.69 was correctly classified as active and formed the terminal node 5. Thirty molecules with Dipole-mag less than 9.96 were formed the branch downside were further split on the basis of IC (Information Content) values. The molecule **69** with IC value greater than 3.88 formed the terminal node 6 was correctly predicted as inactive by the RP model while, the remaining 29 molecules with IC values less than 3.88 were further split on the basis of Radius of Gyration with cut-off of 4.42. Nineteen molecules with Radius of gyration (a spatial descriptor related to the size of a molecule) greater than 4.42 formed the terminal node 7 while, the remaining 10 molecules with the descriptor value less than cut-off followed a branch downside. These were again split on the basis of radius of gyration with the molecules having radius of gyration greater than 4.41 but less than 4.42 formed the terminal node 8 and was found to contain two molecules **46** and **47** correctly predicted to be inactive. Eight molecules with Radius of Gyration less than 4.41 were now split using the descriptor variable density as the splitting criterion with molecule **78** having the density greater than 1.30 formed the terminal node 9 was correctly predicted as inactive. The remaining seven molecules **6, 24, 28, 32, 34, 37,** and **44** with density less than 1.30 formed the terminal node 10 and were correctly classified as active.

Switching to the first primary split thirty three molecules with density value less than 1.24 followed a branch downside. These molecules were further split in to two branches based on dipole-mag property. The molecules **5, 15,** and **20** with Dipole-mag values less than 3.61 formed the terminal node 18 and were correctly predicted as belonging to active class. The remaining 30 molecules followed a branch upside and were split on the basis of density with cut-off value as 1.12. The compounds **14** and **33** with density less than 1.12 formed the terminal node 17 while other twenty eight molecules followed a branch upside in the partition tree. These were further split on the basis of BIC with 5 molecules having the property value less than 0.7 followed a branch downside while those having BIC value greater than 0.7 followed a branch upside. A branch having molecules with BIC less than 0.7 was split in to two sub-nodes on the basis of Dipole-mag. The molecules **13, 19,** and **23** having the Dipole-mag values greater than 5.69 formed the terminal node 15 while, the molecules **57** and **67** with Dipole-mag property less than cut-off value formed the terminal node 16 and were correctly predicted by the RP model. In

viewing 23 molecules, out of 28 which showed the BIC value greater than 0.70 were split based on their density variable. Fifteen molecules with density values less than 1.21 formed the terminal node 14 and were correctly classified as inactive. The remaining eight molecules with density greater than 1.21 followed a branch upside and were split on the basis of molecular volume (Vm). It is a spatial descriptor representing the molecular volume inside the contact surface. It is calculated as a function of conformation and is associated with binding and transport of a molecule. Three molecules **17, 42,** and **63** with molecular volume less than 267.56 followed a branch downside and those with descriptor value greater than 267.56 formed the terminal node 11 containing the correctly predicted inactive compounds **55, 60, 65, 79,** and **80**. The molecules **17, 42,** and **63** were further split based on strain energy with correctly predicted active molecules **17** and **42** forming the terminal node 12 while molecule **63** having strain energy less than 8.15 formed the terminal node 13. Thus, the RP approach is sequentially dividing the dataset on the basis of their physico-chemical descriptor, which help to identify the variables need to be modified in order to arrive at the active molecule.

In conclusion, we have demonstrated the synthesis as well as antimycobacterial screening of a new family of hydrazide. A significant numbers of compounds were found highly potent in the preliminary screening assay. Furthermore, the recursive partitioning model derived for the quinolylhydrazides encodes the useful information pertaining to chemical environment around the molecules. The computational study predictions were also found to be in harmony with antimycobacterial activity data. Despite a high degree of structural similarity in the dataset, the classification model has been able to break down the large dataset in to a series of subsets those are enriched in either active or inactive molecules. This model will be evolved continuously by enlarging the dataset to generate a better hypothesis for designing novel candidates targeting tuberculosis. Further lead optimization and detailed biological study will be communicated in due course.

Acknowledgement

Financial support of the Department of Science and Technology (DST), New Delhi (India) under the project entitled "National Facility for Drug Discovery through New Chemical Entities (NCEs) Development and Instrumentation Support to Small Manufacturing Pharma Enterprise" is gratefully acknowledged. The authors are also thankful to Dr. Cecil Kwong for antimycobacterial screening at Tuberculosis Antimicrobial Acquisition and Coordinating Facility (TAACF), USA.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:

References

1. (a) Phillips, L. *Nature* **2013,** *493*, 14; (b) Mitton-Fry, M.J.; Hanna, D. In Neglected Diseases and Drug Discovery; Palmer, M.J.; Wells, T.N.C. (Eds.) **2012,** Chapter 9, pp 228-261.
2. Ginsberg, A.M.; Spigelman, M. *Nat. Med.* **2007,** *13*, 290.
3. Janin, Y.L. *Bioorg. Med. Chem.* **2007,** *15*, 2479 and references cited therein.
4. World Health Organization (WHO), global tuberculosis control **2010** http://www.who.int/tb/publications/global_report/2010/en/index.html.
5. Gutierrez-Lugo, M.T.; Bewley, C.A. *J. Med. Chem.* **2008,** *51*, 2606.
6. Nayyar, A.; Jain, R. *Curr. Med. Chem.* **2005,** *12*, 1873 and references cited therein.
7. Rawat, B.; Rawat, D. *Med. Res. Rev.* **2013,** *33*, 693 and references cited therein.
8. Saquib, M.; Gupta, M.K.; Sagar, R.; Prabhakar, Y.S.; Shaw, A.K.; Kumar, R.; Maulik, P.R.; Gaikwad, A.N.; Sinha, S.; Srivastava, A.K.; Chaturvedi, V.; Srivastava, R.; Srivastava, B.S. *J. Med. Chem.* **2007,** *50*, 2942.
9. Ahirrao, P. *Mini Rev. Med. Chem.* **2008,** *8*, 1441 and references cited therein.
10. Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. *Nature* **2011,** *469*, 483.
11. Blumberg, H.M.; Burman, W.J.; Chaisson, R.E.; Daley, C.L.; Etkind, S.C.; Friedman, L.N.; Fujiwara, P.; Grzemska, M.; Hopewell, P.C.; Iseman, M.D.; Jasmer, R.M.; Koppaka, V.; Menzies, R.I.; O'Brien, R.J.; Reves, R.R.; Reichman, L.B.; Simone, P.M.; Starke, J.R.; Vernon, A.A. *Am. J. Respir. Crit. Care Med.* **2003,** *167*, 603.
12. Yew, W.W.; Cynamon, M.; Zhang, Y. *Expert Opin. Emerging Drugs* **2011,** *16*, 1.
13. Ali, M.A.; Samy, J.G.; Manogaran, E.; Sellappan, V.; Hasan, M.Z.; Ahsan, M.J.; Pandian, S.; Yar, M.S. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 7000.
14. Sunduru, N.; Gupta, L.; Chaturvedi, V.; Dwivedi, R.; Sinha, S.; Chauhan, P.M.S. *Eur. J. Med. Chem.* **2010,** *45*, 3335.
15. Pieroni, M.; Lilienkampf, A.; Wan, B.; Wang, Y.; Franzblau, S.G.; Kozikowski, A.P. *J. Med. Chem.* **2009,** *52*, 6287.
16. Meng, Q.; Luo, H.; Chen, Y.; Wang, T.; Yao, Q. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 5372.
17. Meng, Q.; Luo, H.; Liu, Y.; Li, W.; Zhang, W.; Yao, Q. *Bioorg. Med. Chem. Lett.* **2009,** *19*, 2808.
18. Frieden, T.R.; Sterling, T.R.; Munsiff, S.S.; Watt, C.J.; Dye, C. *Lancet* **2003,** *362*, 887.
19. Wiles, J.A.; Bradbury, B.J.; Pucci, M.J. *Expert Opin. Ther. Pat.* **2010,** *20*, 1295.
20. Minovski, N.; Vracko, M.; Solmajer, T. *Mol. Divers.* **2011,** *15*, 417.
21. Lilienkampf, A.; Mao, J.; Wan, B.; Wang, Y.; Franzblau, S.G.; Kozikowski, A.P. *J. Med. Chem.* **2009,** *52*, 2109.
22. Villemagne, B.; Crauste, C.; Flipo, M.; Baulard, A.R.; Deprez, B.; Willand, N. *Eur. J. Med. Chem.* **2012,** *51*, 1.
23. (a) Kharkar, P.S.; Desai, B.; Gaveria, H.; Varu, B.; Loriya, R.; Naliapara, Y.; Shah, A.; Kulkarni, V.M. *J. Med. Chem.* **2002,** *45*, 4858; (b) Desai, B.; Sureja, D.; Naliapara, Y.; Shah, A.; Saxena, A.K. *Bioorg. Med. Chem.* **2001,** *9*, 1993; (c) Manvar, A.; Pissurlenkar, R.R.S.; Virsodia, V.; Upadhyay, K.; Manvar, D.; Mishra, A.; Acharya, H.; Parecha, A.; Dholakia, C.; Shah, A.; Coutinho, E.C. *Mol. Divers.* **2010,** *14*, 285.
24. (a) Virsodia, V.; Pissurlenkar, R.R.S.; Manvar, D.; Dholakia, C.; Adlakha, P.; Shah, A.; Coutinho, E.C. *Eur. J. Med. Chem.* **2008,** *43*, 2103; (b) Sarkar, D.; Sarkar, S.; Joshi, R.R.; Khedkar, V.M.; Pissurlenkar, R.R.; Coutinho, E.C.; Shah, A.K. PCT Int Appl, WO 2011/151701 A1 20111208, **2011**; (c) Sarkar, D.; Sarkar, S.; Joshi, R.R.; Khedkar, V.M.; Pissurlenkar, R.R.; Coutinho, E.C.; Shah, A.K. Indian Pat Appl IN 2010DE01244 A 20120302, **2012**.
25. (a) Manvar, A.; Malde, A.; Verma, J.; Virsodia, V.; Mishra, A.; Upadhyay, K.; Acharya, H.; Coutinho, E.C.; Shah, A. *Eur. J. Med. Chem.* **2008,** *43*, 2395; (b) Upadhyay, K.; Bavishi, A.; Thakrar, S.; Radadiya, A.; Vala, H.; Parekh, S.; Bhavsar, D.; Savant, M.; Parmar, M.; Adlakha, P.; Shah, A. *Bioorg. Med. Chem. Lett.* **2011,** *21*, 2547; (c) Upadhyay, K.; Manvar, A.; Rawal, K.; Joshi, S.; Trivedi, J.; Chaniyara, R.; Shah, A. *Chem. Biol. Drug Des.* **2012,** *80*, 1003; (d) Virsodia, V.; Shaikh, M.S.; Manvar, A.; Desai, B.; Parecha, A.; Loriya, R.; Dholariya, K.; Patel, G.; Vora, V.; Upadhyay, K.; Denish, K.; Shah, A.; Coutinho, E.C. *Chem. Biol. Drug Des.* **2010,** *76*, 412.
26. Manvar A.; Bavishi, A.; Radadiya, A.; Patel, J.; Vora, V.; Dodia, N.; Rawal, K.; Shah, A. *Bioorg. Med. Chem. Lett.* **2011,** *21*, 4728.
27. Khunt, R.C.; Khedkar, V.M.; Chawda, R.S.; Chauhan, N.A.; Parikh, A.R.; Coutinho, E.C. *Bioorg. Med. Chem. Lett.* **2012,** *22*, 666.

28. For review see: (a) Hari, D.P.; König, B. *Angew. Chem. Int. Ed.* **2013,** *52,* 4734 and references cited therein (b) Murineddu, G.; Loriga, G.; Gavini, E.; Peana, A.T.; Mulé, A.C.; Pinna, G.A. *Arch. Pharm. Pharm. Med. Chem.* **2001,** *334,* 393.
29. Collins, L.A.; Franzblau, S.G. *Antimicrob. Agents Chemother.* **1997,** *41,* 1004.
30. (a) Strobl, C.; Malley, J.; Tutz, G. *Psychol. Methods* **2009,** *14,* 323; (b) Choi, S.-Y.; Shin, J.-H.; Ryu, C.-K.; Nam, K.-Y.; No, K.-T.; Park Choo, H.-Y. *Bioorg. Med. Chem.* **2006,** *14,* 1608; (c) Young, S.S.; Hawkins, D.M. *Methods Mol. Biol.* **2004,** *275,* 317; (d) Rusinko III, A.; Young, S.S.; Drewry, D.H.; Gerritz, S.W. *Comb. Chem. High Throughput Screening* **2002,** *5,* 125; (e) Blower, P.; Fligner, M.; Verducci, J.; Bjoraker, J. *J. Chem. Inf. Comput. Sci.* **2002,** *42,* 393; (f) Zmuidinavicius, D.; Didziapetris, R.; Japertas, P.; Avdeef, A.; Petrauskas, A. *J. Pharm. Sci.* **2003,** *92,* 621; (g) Morshed, M.N.; Cho, Y.S.; Seo, S.H.; Han, K.-C.; Yang, E.G.; Pae, A.N. *Bioorg. Med. Chem.* **2011,** *19,* 907; (h) Chen, X.; Rusinko, A.; Young, S.S. *J. Chem. Inf. Comput. Sci.* **1998,** *38,* 1054; (i) Rusinko, A.; Farmen, M.W.; Lambert, C.G.; Brown, P.L.; Young, S.S. *J. Chem. Inf. Comput. Sci.* **1999,** *39,* 1017; (j) Lu, J.; Zheng, M.; Wang, Y.; Shen, Q.; Luo, X.; Jiang, H.; Chen, K. *J. Comput.-Aided Mol. Des.* **2011,** *25,* 885; (k) Horvath, D. *SAR QSAR Environ. Res.* **2001,** *12,* 181; (l) Burton, J.; Ijjaali, I.; Barberan, O.; Petitet, F.; Vercauteren, D.P.; Michel, A. *J. Med. Chem.* **2006,** *49,* 6231; (m) Deconinck, E.; Zhang, M.H.; Coomans, D.; Vander Heyden, Y. *J. Chem. Inf. Model* **2006,** *46,* 1410.
31. Accelrys, I. Cerius2. San Diego, CA, USA; **1998.**
32. See, Table 1S in supporting informations.
33. For activity prediction, see Table 2S in supporting informations.

Graphical abstract

Synthesis And Binary QSAR Study of Antitubercular Quinolylhydrazides

Atul Manvar, Vijay Khedkar, Jignesh Patel, Vipul Vora, Narsinh Dodia, Gautam Patel, Evans Coutinho, and Anamik Shah