Original article

Design, synthesis and biological evaluation of novel isoniazid derivatives with potent antitubercular activity

CrossMark

Filomena Martins [a,*], Susana Santos [a], Cristina Ventura [a,b], Ruben Elvas-Leitão [a,c], Lídia Santos [a,1], Susana Vitorino [a,2], Marina Reis [a,b], Vanessa Miranda [a,3], Henrique F. Correia [a], João Aires-de-Sousa [d], Vasyl Kovalishyn [d,4], Diogo A.R.S. Latino [d,e], Jorge Ramos [f], Miguel Viveiros [f]

[a] Centro de Química e Bioquímica (CQB), Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa, Ed. C8, Campo Grande, 1749-016 Lisboa, Portugal
[b] Instituto Superior de Educação e Ciências, Alameda das Linhas de Torres 179, 1750-142 Lisboa, Portugal
[c] Área Departamental de Engenharia Química, Instituto Superior de Engenharia de Lisboa (ISEL), Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro, 1959-007 Lisboa, Portugal
[d] REQUIMTE and CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte de Caparica, 2829-516 Caparica, Portugal
[e] Centro de Ciências Moleculares e Materiais (CCMM), Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa, Ed. C8, Campo Grande, 1749-016 Lisboa, Portugal
[f] Grupo de Micobactérias, Unidade de Microbiologia Médica, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira 100, 1349-008 Lisboa, Portugal

A B S T R A C T

The disturbing emergence of multidrug-resistant strains of *Mycobacterium tuberculosis* (*Mtb*) has been driving the scientific community to urgently search for new and efficient antitubercular drugs. Despite the various drugs currently under evaluation, isoniazid is still the key and most effective component in all multi-therapeutic regimens recommended by the WHO. This paper describes the QSAR-oriented design, synthesis and *in vitro* antitubercular activity of several potent isoniazid derivatives (iso-nicotinoyl hydrazones and isonicotinoyl hydrazides) against H37Rv and two resistant *Mtb* strains. QSAR studies entailed RFs and ASNNs classification models, as well as MLR models. Strict validation procedures were used to guarantee the models' robustness and predictive ability. Lipophilicity was shown not to be relevant to explain the activity of these derivatives, whereas shorter $N-N$ distances and lengthy sub-stituents lead to more active compounds. Compounds **1**, **2**, **4**, **5** and **6**, showed measured activities against H37Rv higher than INH (*i.e.*, MIC $\leq$ 0.28 μM), while compound **9** exhibited a six fold decrease in MIC against the *katG* (S315T) mutated strain, by comparison with INH (*i.e.*, 6.9 *vs.* 43.8 μM). All compounds were ineffective against H37Rv$_{INH}$ ($\Delta katG$), a strain with a full deletion of the *katG* gene, thus

corroborating the importance of KatG in the activation of INH-based compounds. The most potent compounds were also shown not to be cytotoxic up to a concentration 500 times higher than MIC.

1. Introduction

With nearly one-third of the global population infected with the *Mycobacterium tuberculosis* bacilli (*Mtb*), tuberculosis (TB) is still a major cause of mortality and morbidity. The World Health Organization (WHO) estimated for 2012 8.6 million new cases and 1.3 million deaths due to TB, of which 320.000 among HIV-positive [1]. It was also reported that about 3.6% of the new patients and 20% of the previously treated ones had multidrug-resistant TB (MDR-TB), which is caused by *Mtb* strains resistant to the most effective first-line drugs, isoniazid (INH) and rifampicin (RIF). Furthermore, as by September 2013, 92 countries had reported at least one case of extensively drug-resistant TB (XDR-TB), a form of TB resistant to INH and RIF, as well as to any fluoroquinolone and to any of the second-line injectable drugs (amikacin, kanamycin or capreomycin). Recent WHO updates claim that approximately 9.6% of all MDR-TB cases are XDR-TB [2]. To make the scenario even more serious, a new designation, "totally drug-resistant" tuberculosis (TDR-TB), although not yet recognized by the WHO, has been used in the literature to describe an allegedly new (or deadlier) form of TB, resistant to all first- and second-line drugs [3–5]. These resistant strains are emerging as a major public health problem, threatening the success of the DOTS program (Directly-Observed Treatment Short-course), the WHO program launched in 1994 for monitoring TB detection, treatment and drug supply [6,7].

Besides the need to make the treatment of drug-sensitive TB shorter, simpler, safer and more accessible, also treatment of MDR- and XDR-TB must be greatly improved to become more efficient, less toxic, better tolerated and less expensive, as well as compatible with antiretroviral therapies due to the increasing number of patients co-infected with HIV/AIDS [8–11]. Current drugs against TB are mostly inadequate to address this fight against tuberculosis.

After decades of stagnation in research for new antitubercular drugs, a new interest from pharmaceutical companies and consortia, a more consistent support from governmental institutions and some effective funding from international non-governmental organizations, have led to the appearance of several classes of molecules which trigger different biological pathways such as cell wall synthesis, nucleic acid synthesis, protein synthesis or membrane energy production and which are now under various stages of clinical trials [11–13].

Despite the accelerated approval granted by FDA to bedaquiline, the first drug to seek a specific indication for (adult) MDR-TB in the USA and also the first drug with a new mechanism of action (MOA) for treating TB in over 40 years [14–16], the fact remains that isoniazid is still the most effective drug against wild type *Mtb* and is therefore the treatment of choice for TB.

The clinical usefulness of isoniazid (isonicotinic acid hydrazide) was discovered in the 1950's when its specific high activity against *Mtb* was first noticed, higher than any other compound used at the time [17,18]. INH is indeed particularly active against *Mtb* with a minimum inhibitory concentration (MIC) of 0.05 μg/mL [18].

It is thought that INH enters *Mtb* through passive diffusion through the cell-wall and that it is active only against dividing bacteria [19,20]. The first study regarding the mechanism of action of INH was published in 1970 by Winder and Collins [21], connecting INH with mycolic acid biosynthesis. In order for INH to be effective against *Mtb* it needs to be activated by the multifunctional catalase–peroxidase enzyme KatG into a range of activated species, such as an isonicotinoyl radical, that can acylate numerous compounds [22,23]. The generally accepted mechanism of action of INH postulates that the isonicotinoyl radical binds to the nicotinamide adenine dinucleotide (NAD^+) and the resulting adduct inhibits the enoyl-ACP reductase InhA, a NADH-dependent enoyl-acyl carrier protein (enoyl-ACP) reductase of the fatty acid synthase type II system (FASII). This inhibition causes accumulation of long-chain fatty acids, inhibition of mycolic acid biosynthesis and, ultimately, cell death [18,24].

The main mechanism of resistance to isoniazid resides in the presence of mutations in its activator, KatG, product of the *katG* gene [25], whereas mutations in the *inhA* gene represent the second most common cause of resistance. Altogether, mutations in these two genes are responsible for approximately 75% of all cases of *Mtb* resistance to isoniazid in the clinical setting [26]. Resistance to INH has also been associated with mutations in several other genes (*e.g.*, *ndh*, *kasA* and *oxyR*–*ahpC* intergenic region) [18], but their direct association with resistance is still not clear. It has also been demonstrated very recently that drug efflux contributes likewise to the overall resistance, particularly in the case of acquired isoniazid resistance [27,28], much in the same way as active-efflux plays a role in the failure of cancer chemotherapy [29].

As mentioned before, isoniazid remains a key component in all multiple drug treatment regimens recommended by the WHO albeit resistant isolates are rapidly generated during monotherapy or inappropriate treatment [30]. Similarly to some other candidates in the pipeline that were chemically tailored from drugs to which *Mtb* was already resistant, departing from isoniazid might also bring new insights into the development of new antitubercular agents. Hence, improvement of INH by introducing chemical modifications in its core structure in order to enhance the biological response against *Mtb* and/or circumvent resistance phenomena continues to be an interesting scientific challenge.

To face this challenge, *in vitro*, *in vivo* or *in silico* methods are being used in early stages of drug development to avoid possible failures, especially those related with drug metabolism, pharmacokinetic profiles and toxicity issues [31,32]. In particular, quantitative structure–activity relationships (QSAR), is an *in silico* methodology extensively used in Medicinal Chemistry to find relationships between molecular properties of chemical compounds and measured biological activities [32–34].

With the purpose of establishing robust and predictive MLR- and NN-based QSAR models to assist us in the design of new active compounds, we have developed a systematic analysis of a large data set of potentially antitubercular agents [35–37]. In this work we describe the details of the QSAR-oriented design, synthesis and *in vitro* antibacterial activity evaluation of 13 compounds derived from the INH core structure (Fig. 1) which were assessed against the *wt Mtb* H37Rv strain and two mutated strains, one carrying solely a *katG* S315T mutation (the most frequent mutation related with INH resistance), and the other resulting from a full deletion of the *katG* gene (Δ*katG*) [28].

2. Methods

2.1. Data sets

Different blocks of data, based on two independent approaches, were used to build anti-tubercular activity models in order to design novel potentially active isoniazid derivatives, namely i –
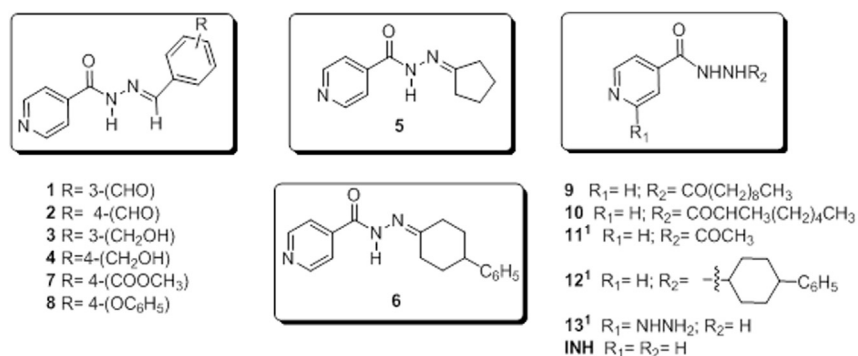
Fig. 1. Structures of compounds synthesized in this work. [1]Compounds 11, 12 and 13 have been synthesized for comparative purposes.

"RFs and ASNNs data sets", used to develop classification models that were integrated into a consensus model to predict the propensity of a compound to be active [36,38]. These classification models were applied to a "virtual data set" of compounds assembled using the INH molecule as a common scaffold; ii – "MLRs data sets", used to develop regression models for the design of another set of potentially active compounds. Additionally, these MLR models were also used for a quantitative evaluation of some selected compounds proposed as active by the classification models.

2.1.1. RFs and ASNNs data sets

Four different data sets were used to establish classification models with Associative Neural Networks (ASNNs) and Random Forests (RFs) methods. All compounds were divided into two classes, according to their MIC values, i.e., their minimum inhibitory concentrations: active (MIC ≤ 5 μM) and inactive (MIC > 5 μM).

Data set i consisted of 173 hydrazide derivatives with known MIC values retrieved from the literature, 136 compounds tested against the BCG Mtb strain [39] and 37 against the H37Rv [40–43]. Data set ii consisted of 372 INH derivatives obtained from two sources: 259 compounds from the NIAID and GVK databases [44], evaluated against H37Rv, and 113 retrieved from several other sources [39–43,45–47], of which 53 were tested against BCG and 60 against H37Rv. Data set iii corresponded to the 113 INH derivatives subset of data set ii whereas data set iv corresponded to the remaining 259 INH derivatives. An additional external evaluation set composed of 45 INH derivatives tested against the H37RV strain, data set v, was also selected from the literature [48–50]. All compounds were stored in SMILES and MDL SDF formats.

RFs and ASNNs data sets are given in Table S1 in Supporting Information.

2.1.2. Virtual data set

A set of 430 drug-like molecules was generated with an in-house developed program starting from 10 skeletons comprising derivatives of two INH families, isonicotinoyl hydrazides and isonicotinoyl hydrazones, and 43 common substitution patterns (see Table S2, in Supporting Information). This set was then screened by the developed classification models.

2.1.3. MLR data sets

For building the MLR models, biological data were retrieved from literature for 45 compounds tested against the BCG strain [39], and 37 compounds tested against the H37Rv strain [40–43], in a total of 82 isoniazid derivatives. Although data refer to two different strains, no systematic deviation among the various sets of MIC values that could be attributed to data inconsistency was ever noticed in the setting up of the models either in this work or in our previous investigations involving INH derivatives [35–37].

Additionally, a set of 13 compounds from the Novartis TB public access database, with MIC values measured against the H37Rv strain, were used to make a further external validation of the established models [51].

2.2. Molecular descriptors

2.2.1. Molecular descriptors for RFs and ASNNs models

A set of 3224 theoretical molecular descriptors divided in 22 types was calculated using the DRAGON program [52] for all mentioned RFs and ASNNs data sets. Constant and near-constant descriptors were deleted.

A set of 1270 descriptors, for the above mentioned data sets, was also calculated with the ADRIANA.Code program [53]. For this purpose, the 3D structures were generated with the embedded CORINA program.

The type of descriptors used in each case is presented in Table S1 of Supporting Information.

2.2.2. Molecular descriptors for MLR models

Twenty-eight molecular descriptors were calculated for each of the 82 isoniazid derived compounds: 5 Abraham's descriptors calculated with the Absolv program [54–56], and 8 energetic, 9 geometrical, 3 structural (Verloop's parameters), and 2 electronic parameters calculated with the Molecular Modeling Pro Plus (MMP$^+$) software [57]. Additionally, ClogP, taken as a measure of lipophilicity, was calculated using ChemDraw Ultra, 11.0.1 [58]. For descriptors dependent on 3D structures, the geometry of each compound was previously optimized using the MM2 method, a molecular mechanics method incorporated in the software. Partial charges were calculated using the Del Re method and dipole moments by a combination of PEOE and Huckel/4 methods also included in the referenced software. Molecular structures were drawn with ChemDraw Ultra, and saved as SMILES strings and are available as Supporting Information in Table S3. Descriptor values for all 82 compounds, as well as the corresponding MIC values from literature, are also given as Supporting Information in Table S4.

2.3. Classification and regression techniques

2.3.1. Random Forests

A Random Forest [59], RF, is an ensemble of unpruned classification trees which is created using bootstrap samples of the training set and random subsets of descriptors to define the best split at each node. The final prediction for an object from a RF is obtained by majority voting of the individual trees. The

bootstrapping technique is used to internally assess the model's performance from the prediction errors for the objects left out in the bootstrap procedure (out-of-bag estimation, OOB). RF can also be used for descriptor selection. The RF algorithm estimates the importance of a descriptor by determining how much the prediction error increases when data for that descriptor is permuted keeping all others unchanged. RFs were grown with R program v. 2.10.0 and 2.11.0 [60] using the Random Forest library. Forest size and number of descriptors selected at each split were optimized, being the number of trees set to 1000 and the number of descriptors tried at each split set to 30.

2.3.2. Associative neural networks

Associative neural networks (ASNNs) [61] integrate an ensemble of feed-forward neural networks (FFNNs) with a memory of experimental data. The ensemble consists of independently trained FFNNs, which contribute to a single prediction. The final prediction for an object from an ASNN is obtained from (a) the outputs produced by the ensemble of individual NNs and (b) the most similar cases in the memory (here, the training set). Further details on the application of this technique can be found in Supporting Information, Section 1.

2.3.3. Multiple linear regressions

In the context of this work, a multiple linear regression (MLR) methodology was used to establish a mathematical relationship between the biological activity of the target compounds (Y) and a set of molecular descriptors (X_i) that numerically encode the features of the compounds' chemical structure and which can be expressed in general terms as follows:

$$Y = a_0 + \sum_{i=1}^{N} a_i X_i + \zeta \tag{1}$$

where ζ stands for the regression residuals and a_0 and a_i are the regression coefficients obtained by minimizing the sum of the squared residuals. The order of magnitude of the regression coefficients reflects the degree of influence of the molecular descriptors, whereas their sign can be interpreted in terms of positive or negative contributions to the activity. A correct use of a MLR-QSAR approach presupposes that the descriptors are not intercorrelated to avoid redundancy in descriptor information [62] and that the number of compounds for model building is at least four times the number of descriptors used, to avoid overfitting and therefore chance correlations [62,63]. The success of a MLR-QSAR approach in the development of interpretative and predictive model equations requires representative, homogeneous and good quality data, a well-defined domain of applicability [64–66], and strict validation procedures [63,67,68].

2.4. Selection of descriptors for ASNNs and RF models

The selection of descriptors to develop RF and ASNN classification models was performed in two steps: (a) using the ranking of descriptors established by RFs and (b) using pruning methods implemented in the ASNN program. A pre-selection of the 300 to 400 most important descriptors obtained from RF models was performed and used in the next step where the ASNN pruning methods were applied. Details on the ASNNs pruning methods are given in Supporting Information, Section 2.

The influence of the number of selected descriptors on the ASNN model quality was checked on the basis of the leave-one-out (LOO) results for the training sets (please refer to Supporting Information, Section 1). Pre-selecting less than 300−400 descriptors by RF

decreased the ASNN models predictive ability. With the final sets of descriptors selected by ASNNs pruning methods for each one of the data sets, ASNNs and RFs were used to obtain the final models.

2.5. Validation procedures

2.5.1. Validation of RFs and ASNN models

Both ASNNs and RF models were first internally validated on the basis of the training set by their intrinsic validation procedures, namely out-of-bag (OOB) and leave-one-out method, respectively (see Section 2.3). Furthermore, models were submitted to an external validation using an independent test set (*data set v*).

The traditional way of summarizing "lack of fit" in QSAR models is to use the root mean squared error (*RMSE*) and the mean absolute error (*MAE*) between predicted and experimental activities. For ASNNs, the *RMSE* is calculated by the expression

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{k} \left(Y_{ij} - O_{ij}\right)^2}{N \times k}} \tag{2}$$

where Y_{ij} and O_{ij} are the predicted and measured activity j for molecule i, respectively, N is the number of compounds, and k is the number of output neurons.

To assess the classification ability and to separately monitor the classification performance of the two classes (active and non-active), sensitivity (Sn), specificity (Sp), and overall accuracy (Ac) were calculated as follows:

$$Sn = TP/(TP + FN) \tag{3}$$

$$Sp = TN/(TN + FP) \tag{4}$$

$$Ac = (TP + TN)/(TP + FN + TN + FP) \tag{5}$$

where TP, FP, TN and FN are true positives, false positives, true negatives and false negatives, respectively. In general, the overall accuracy, Ac, is always used to measure the predictive power of classification models, as the number of active and inactive compounds is approximately the same in all data sets.

2.5.2. Validation procedures used on MLR models

For MLR validation purposes, data was split into a training set and an independent test set with similar variability both in the dependent variable, *i.e.*, the biological activity, and in the chemical/structural multidimensional space covered by the descriptors. Training/test set splitting took also into consideration the fact that test set compounds must exhibit structural similarity with their training set counterparts to hamper predictions by extrapolation which are more likely to be unreliable [62,66]. The training set was used to derive the models, whereas the test set (and, for that matter, also the Novartis external data set) was used to evaluate the models' predictive ability.

The best MLR models were found by a forward stepwise procedure starting with equations containing a single descriptor and adding up terms, one at a time. All combinations of the 28 molecular descriptors were tested and descriptors were retained or disregarded in each step according to rigorous statistical criteria (see below).

Also, the intercorrelation among descriptors used in each regression was always checked (for descriptors to be considered non-redundant, r^2 between any two descriptors should be smaller than 0.5 and R^2 for a given descriptor against a linear combination of all others, should be lower than 0.8) [69].

Additionally, models were examined for anomalous observations, the so-called outliers. These allegedly suspicious values were fully assessed using two criteria: the conventional criterion of $|Y_{calc} - Y_{exp}| > 2\,SD$, where SD stands for the standard deviation of the fit, and a more refined measure, the Cook's distance [70,71].

The goodness-of-fit of the best models was also evaluated by checking their linearity through residual analysis and normal probability plots. The models' robustness (evaluated with training set compounds) was assessed by traditional statistical criteria, namely, the determination coefficient, R^2, the standard deviation of the fit, *SD*, the *F* statistics, and the significance level (*SL*) of each adjusted parameter (parameters were kept in the regressions only if $SL > 95\%$).

Internal validation was further gauged by the leave-many-out (LMO) cross validation correlation coefficient, Q^2, given by:

$$Q^2 = 1 - \frac{\sum_{i=1}^{\text{training}} \left(Y_i - \widehat{Y}_i\right)^2}{\sum_{i=1}^{\text{training}} \left(Y_i - \overline{Y}_i\right)^2} \tag{6}$$

where Y_i, $\widehat{Y}_i$ and $\overline{Y}_i$ are the measured, predicted and averaged (over the whole data set) values of the dependent variable, respectively. A Q^2_{LMO} (average) > 0.6 is usually taken as a good measure of the robustness of the model [35,63,72]. Root mean squared errors (*RMSE*) were also calculated.

The predictive power of the MLR models was evaluated through several demanding and complementary external validation criteria, using precautionary threshold values, *i.e.*, values significantly higher than those normally applied in current QSAR studies [73–76]. The "measures" of predictive ability calculated in this work were the external $Q^2_{ext}(F1)$ statistics [62,74], the r^2_m metrics (and also the corresponding Δr^2_m) [73,76,77] and the concordance correlation coefficient, *CCC* [74,75]. Details of these metrics are given in Supporting Information, Section 3. Additionally, scatter plots of predicted *vs.* experimental data were also obtained, as recent studies have recommended the visual inspection of these plots as important complementary indicators of model predictivity [75].

The predictive capacity of each model was further evaluated by computing various measures of fit between predicted and experimental values, namely, the average error (*AE*), the absolute average error (*AAE*) and the already referred *RMSE* which should all be as low as possible. Finally, models had still to comply with the following criteria:

$$R^2 > 0.6; \quad \frac{R^2 - R_0^2}{R^2} < 0.1; \quad 0.85 < m < 1.15$$

where R^2 and R_0^2 have the same meaning as before, and *m* is the slope of the regression between predicted and experimental values [63,67].

A Y-randomization was also performed to eliminate the possibility of chance correlation in the best MLR models. A randomization experiment was carried out by randomly shuffling the Y-column values (log (1/MIC)) twenty five times while keeping the descriptor matrix unchanged. Each randomized model was subsequently used to make predictions for the test set, being the corresponding R^2 and *RMSE* calculated each time. Both an increase in *RMSE* and a substantial decrease in R^2 *vis-à-vis* the non-randomized model support the robustness and statistical reliability of the original model. To help judging if the difference between the original R^2 and the average R^2 for the randomized models, R_r^2, was significant, a further statistical parameter, the so-called Todeschini $^cR_p^2$ parameter, was also computed [78]:

$$^cR_p^2 = R\sqrt{\left(R^2 - R_r^2\right)} \tag{7}$$

QSAR models that present $^cR_p^2$ values higher that 0.5 are considered robust and reliable and not merely a result of chance correlation.

A further important aspect to take into consideration in this type of analysis is the applicability domain (AD) of the obtained QSAR models, which is essential to guarantee that they make reliable predictions [62,64,65,79]. To assess the AD in the developed MLR models, two independent filtering methods were considered: i − the range of individual descriptors, and ii − the leverage approach. In the first case, the descriptors' range for training set compounds was calculated and the minimum and maximum values of the descriptors were used to define their interval of validity. Test set (and designed) compounds descriptors were then compared with the corresponding average values in the training set to ensure that predictions were performed for compounds whose molecular descriptors fell within the established limits. In the second case, the leverage value, *h*, of each compound was determined as described elsewhere [64,79,80], thus providing a measure of the distance of the compound from the centroid of the training set. A "warning leverage", h^*, is generally fixed at $3p/N$, where *N* is the number of compounds in the training set and *p* the number of model variables plus one [64,79,80]. This method allows a graphical assessment of the standardized residuals, *SR*, as a function of the leverage values (Williams plot) thus permitting a "better" visualization of the AD and the detection of influential compounds, *i.e.*, compounds structurally distant from training set compounds ($h > h^*$) and of response outliers ($SR > \pm 3\,SD$ units) [79,80]. In this type of plot the AD corresponds to the squared area defined by the range $\pm 3\,SR$ and the threshold leverage value h^* [64,65,80].

3. Results and discussion

3.1. Chemistry

Synthesis of all compounds was successfully achieved using general organic synthetic pathways and their chemical structure (please refer to Fig. 1) was elucidated by extensive spectroscopic analysis (¹H and ¹³C NMR, COSY, DEPT, HMQC, HMBC, NOESY and MS).

Isonicotinoyl hydrazones 1, **2**, **5**−**8** were prepared by functionalization of INH at N-2 using standard methods. Typically, treatment of INH with the appropriate ketone or aldehyde in ethanol at room temperature led to the corresponding Schiff base, which was generally isolated as a crystalline solid after column chromatography. For compounds **1** and **2** a 1:0.5 INH/aldehyde ratio assured the reaction of just one of the aldehyde functions. Compounds **3** and **4** were obtained by reduction of the parent compounds **1** and **2** with NaBH₄ in ethanol, and further recrystallization.

The stereochemistry of the double bond in hydrazones **1**−**4** and **7**−**8** was assigned as synperiplanar *E,* based on ¹H NMR experiments. It is well known that acylhydrazones may exist as E/Z geometric isomers about the C=N double bonds and as synperiplanar/ antiperiplanar amide conformers (Fig. 2). However, it is reported in literature that hydrazones derived from aldehydes and substituted hydrazides are present in solution mainly in the *E* form [81–83]. GC and HPLC analysis of our compounds confirmed that, in all cases, only one isomer was present. NOESY experiments were conducted in DMSO-d_6, and revealed a well-defined cross peak between the CON<u>H</u> proton and the iminic proton, only possible in the *E* geometry of the double bond. Another cross-peak between the CON<u>H</u> proton and the H-2/H-6 protons disclosed a synperiplanar CO−NH conformation. The existence of an equilibrium in solution of the major synperiplanar and the minor antiperiplanar CO−NH conformation was also observed. ¹H NMR experiments at several temperatures (20, 40, 70 and 90 °C) were run in order to confirm that
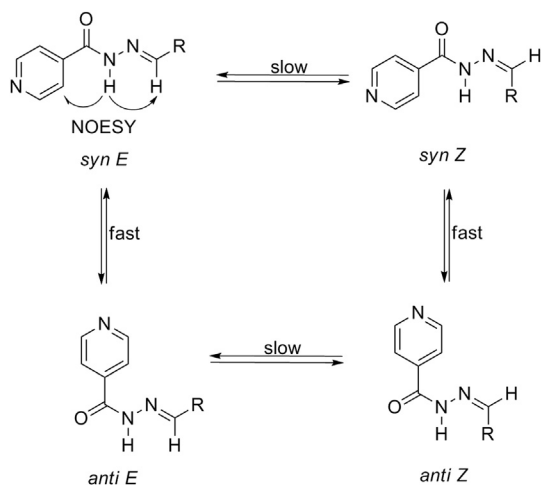
Fig. 2. Stereoisomers of isonicotinoyl hydrazones and NOESY correlation in the E synperiplanar isomer.

the double signs with very small intensity observed in the spectrum, were actually due to the minor antiperiplanar conformer. The coalescence of the signals was complete at 90 °C, being reversible when temperature was returned to 20 °C, thus confirming that the conversion of the two conformers is the only process of isomerization. The presence of conformers can be obtained only by ^{1}H NMR experiments since the interconversion is too fast to be detected by HPLC [81].

The configuration of C=N double bond in the Schiff bases was confirmed by X-ray diffraction crystallography. As an example, the ORTEP diagram of compound **7** is presented in Fig. 3 showing that the asymmetric unit consists of one Schiff base molecule and two water molecules. The structure of the compound reveals *quasi* co-planarity of the whole molecular skeleton, with a dihedral angle of 2.34° between the benzene and the pyridine rings. The *E* configuration of the C(7)=N(3) double bond, is evidenced by the C(6)-N(2)-N(3)-C(7) torsion angle of −178.54(10). The O(1) atom and the hydrazone N(3) atom are *cis* with respect to C(6)-N(2) which confirms the amide synperiplanar arrangement.

In the crystal structure, the water molecules link the molecules into a three-dimensional network by the intermolecular H(2N)···O(1W) and O(2)···H(4W) hydrogen bonds (Fig. 3).

Isonicotinoyl hydrazides **9**, **10** and **11** were synthesized using standard literature methods; compound **13** was obtained according to a specific procedure reported in the literature [84]. Compound **12** was obtained by treatment of the corresponding hydrazone (compound **6**) with NaBH$_3$CN, followed by repeated column chromatography, in order to separate the two diastereoisomers formed in the reduction.

To the best of our knowledge, compounds **1**, **3**, **4**, **6–8**, **10** and **12** are herein reported for the first time. Compounds **5** [85] and **13** [84], although already reported in the literature were spectroscopically characterized for the first time in this work. Compound **9** [86–88] has also been previously synthesized but, along with compound **13**, has no reported MIC values. Likewise, compound **2** [89], used as a precursor of compound **4**, was described and characterized by others, but as far as we know has never been tested against *Mtb*. Compound **11** has a reported MIC value against the BCG [39] strain of 15 μg/mL. Finally, compound **5** has MIC values reported against *Mtb* H37Rv strain (0.05 μg/mL) and against a INH resistant strain (>0.75 μg/mL) [85].

3.2. Information retrieved from QSAR models

3.2.1. RFs and neural networks models

3.2.1.1. ASNN results. In the first stage, ASNN models were developed using the total set of descriptors. The initial number of descriptors was then submitted to an additional reduction procedure in which descriptors with constant values were removed. Then the importance of the descriptors for the anti-TB activity was evaluated by Random Forests and ASNN gradual pruning methods (see Supporting Information, Section 2). New models were developed with the final selection of descriptors. Results are summarized in Table 1. The total accuracy Ac for the training sets ranged from 87.2 to 96.5%. The 45 compounds in the external *test set v* were predicted with an accuracy of 66.6–77.8%.

3.2.1.2. RF results. The results of this method for all data sets were statistically similar to the ASNN results. However, the predictive ability for the test sets was in some cases (models RF1 and RF3) slightly higher than that for ASNNs models (Table 1). The observed Sn and Sp values indicate that the models are balanced.

To visualize the regions of the chemical space covered by the different training sets and the overlap of these regions with the compounds of the virtual set, a principal component analysis was performed. The top 2 principal components were calculated using
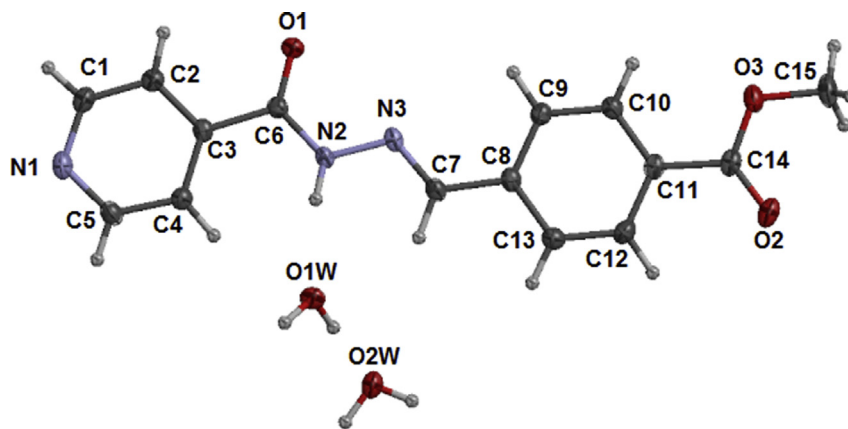

Fig. 3. ORTEP view of compound **7**, determined at 150K, with atom-labeling scheme and thermal ellipsoids drawn at 50% probability level. Selected bond and torsion angles (deg): C(6)-N(2)-N(3) = 118.22(10); C(7)-N(3)-N(2) = 114.85(10); N(3)-C(7)-C(8) = 121.68(11); N(3)-C(7)-H(7) = 120.8(8); O(1)-C(6)-N(2) = 123.42(10); O(1)-C(6)-C(3) = 120.70(10); C(6)-N(2)-N(3)-C(7): −178.54(10); N(2)-N(3)-C(7)-C(8) = −179.57(10); N(3)-N(2)-C(6)-C(3) = 178.06(9); C(10)-C(11)-C(14)-O(2) = 173.93(12); C(10)-C(11)-C(14)-O(3) = −6.13(17); C(15)-O(3)-C(14)-C(11) = 179.15(11).

Table 1
Comparison of classification models built with different RFs and ASNNs.

Model	Sets	N	Number of descriptors	Method							
				RFs			ASNNs				
				Sn	Sp	Correct (Ac,%)[a]	Sn	Sp	Correct (Ac%)[a]	RMSE
1	Training *i*	173	16	0.88	0.85	149 (86.1)	0.88	0.91	155 (89.6)	0.29
	Test *v*	45		0.90	0.71	38 (84.4)	0.81	0.45	30 (66.6)	0.47
2	Training *ii*	372	24	0.84	0.87	316 (84.9)	0.90	0.89	333 (89.0)	0.30
	Test *v*	45		0.78	1.0	36 (80.0)	0.78	0.80	35 (77.8)	0.44
3	Training *iii*	113	8	0.90	0.93	103 (91.2)	0.97	0.96	109 (96.5)	0.2
	Test *v*	45		0.80	1.0	37 (82.2)	0.78	0.56	33 (73.3)	0.48
4	Training *iv*	259	24	0.79	0.82	208 (80.3)	0.86	0.89	226 (87.2)	0.32
	Test *v*	45		0.78	0.80	35 (77.8)	0.78	0.80	35 (77.8)	0.43

[a] Ac − accuracy in %.

the Weka 3.7.3 software [90] with the full set of 3059 initial descriptors, common to all data sets, for the training sets i, iii, iv (training set ii consists in training set iii and iv), the test set and the virtual set. The algorithm to calculate principal components does not take into account any information concerning the activity of the compounds. The visualization of the chemical space based on the projection of the 2 top principal components is also another approach to the applicability domain of the models. In this case it shows that the test set is in the chemical space of the different training sets, and most of the compounds of the virtual set are also in the same chemical space of the training sets (Fig. 4). The virtual set is concentrated in the center of the chemical space, a region with high density of compounds particularly from training set iv. An additional representation is included as Supporting Information (Fig. S1) with active and inactive compounds of training and test sets labeled differently. This representation shows that the region in the chemical space with higher density of active compounds is near the center of the chemical space. Moreover, this region also shows a high density of compounds of the virtual set.

3.2.2. MLR models

Table 2 shows the results for the best found MLR models. The 82 initial compounds were divided in two sets: a training set of 62 compounds and a test set of 20 compounds (∼a 75%:25% splitting, well within recommended limits). A close inspection of the training set matrix immediately pinpointed four outliers: compounds 10, 21, 23 and 39 in Table S3. In fact, these compounds show very high MIC values (between 149.32 and 600 μg/mL) *i.e.*, a very low activity when compared with the remaining 58 compounds which show an average MIC value of 8.35 μg/mL, and were, hence, promptly removed. Models were built with this new training set and 24 new outliers were sequentially detected according to the two criteria mentioned in Section 2.5 (the identified outliers were compounds 1, 2, 9, 11–12, 14, 15, 18, 24, 30–31, 33, 35–36, 40, 51, 62–67, 70 and 79 in Table S3, *i.e.*, 9 hydrazones and 19 hydrazides). These outliers were also eliminated, one-by-one, from the training set, and the models refitted after each removal until we reached the final models for a training set comprising 34 compounds.

The figures of merit for all six models, shown in Table 2, fulfill the statistical criteria set above for both internal and external validation, in particular in the case of the two best models, models A3 and B3, which conform to the most demanding thresholds and are able to predict log (1/MIC) for this set of compounds with an average *SD* of 0.47 log units.

To remove any possibility of attributing the quality of the statistics of these models to a chance correlation between the response variable and the descriptors, a *Y*-randomization was performed as described in Section 2.5 and results are presented in Table 3. As can be observed, there is a significant decrease in the quality of the randomized models when compared to the original non-randomized ones, and therefore there seems to be no chance correlation, as corroborated by the value of $^cR_p^2$, quite above the 0.5 threshold value.

A close analysis of Table 2 shows that the biological behavior of the studied INH derivatives against *Mtb* does not depend on their physicochemical, energetic or electronic characteristics. Descriptors falling in these categories were found not to be important to model log (1/MIC) values. Also, the lipophilicity characteristics measured by Clog*P*, do not appear as relevant to explain the antitubercular activity of these derivatives. A similar result had already been reported by us in a previous study [35] involving the more general class of hydrazide derivatives.

Conversely, steric and geometrical features emerge as very effective in this modeling process. In fact, models A and B both include the steric Verloop's parameters L, B_1, and B_5, where L is a measure of the substituent's length, B_1 is essentially a measure of the size of the first atom in the substituent (largely a steric effect) and B_5 is an attempt to define the effective volume of the whole substituent [91,92]. Concurrently, model A involves an extra geometrical parameter, d_3, which refers to the interatomic distance between the two exocyclic nitrogen atoms, whereas model B contains an additional parameter, d_k, which corresponds to the 2D distance between the pyridinic nitrogen and the terminal nitrogen of the hydrazide functionality (see Fig. 5) [39].

Both models A and B in their complete form, *i.e.*, models A3 and B3, explain virtually the same variability in the response variable (80.1% and 79.3%, respectively) since d_3 and d_k are highly correlated, although no significant collinearity was observed between any pair of parameters in the same equation (r^2 was always below 0.59) or between one parameter against all others (R^2 was always lower than 0.68), as can be seen by the intercorrelation matrices (Table 4). However, the predictive power of model A3 is slightly superior to that of model B3, as suggested by the set of measures used to assess the quality of the predictions for the same test set (Table 2). *CCC* for model B3 is also marginally below the suggested threshold value of 0.85. The scatter plot for model A3 represented in Fig. 6 shows that there are no systematic deviations from the ideal line (represented as a dashed line in the figure); in other words, there are no local, scale or local plus scale shifts in the data. This type of plot allows an easy visual inspection of the quality of the model, thus precluding the possibility of having good statistical validation criteria but a "bad" predictive model [75]. Additionally, as already mentioned, we have tested the predictive power of our models towards a totally external data
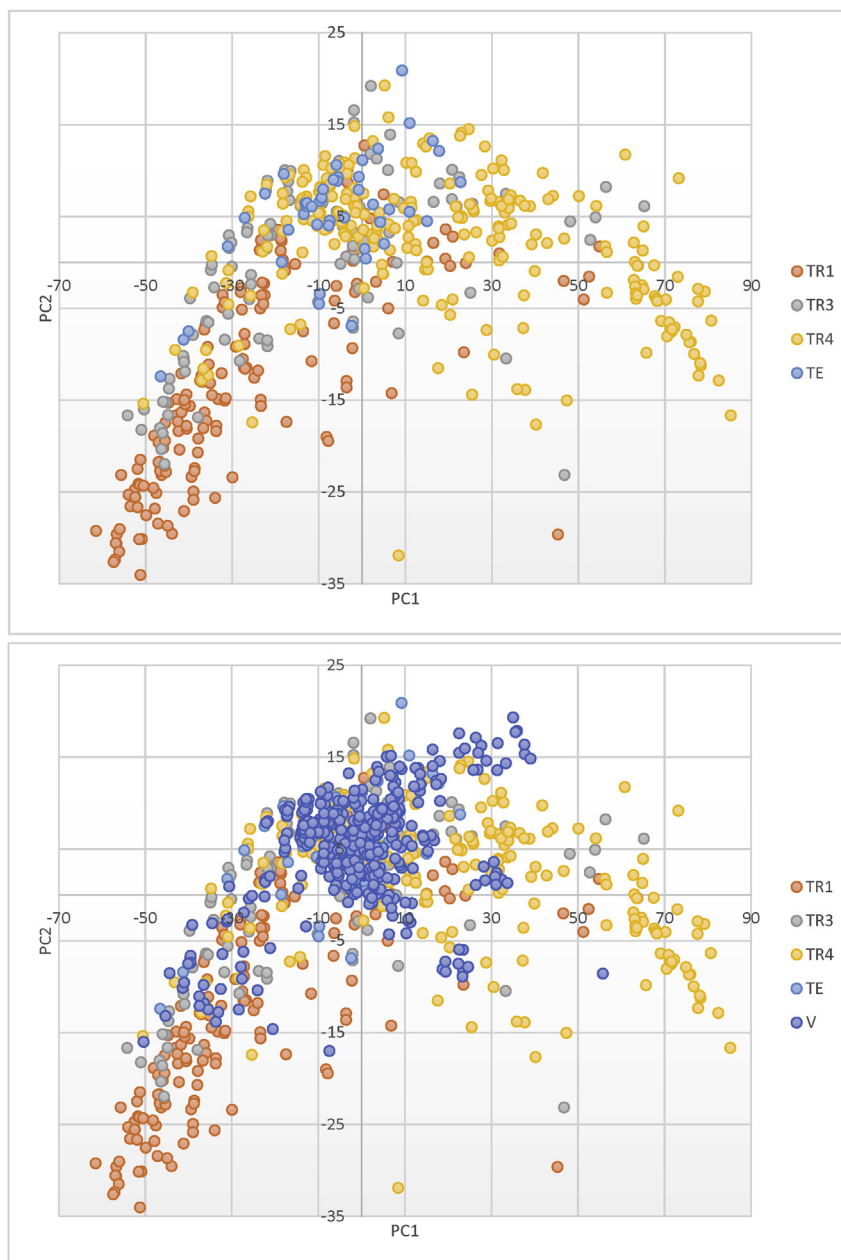
Fig. 4. Representation of the two-top principal components for all training sets used to develop the RF/ASNN classification models, the test set and the virtual set. Training set i consists of 173 compounds (61 actives and 112 inactives), training set iii consists of 113 compounds (68 actives and 45 inactives), training set iv consists of 259 compounds (130 actives and 129 inactives), and test set consists of 45 compounds (32 actives and 13 inactives). The virtual set consists of 430 compounds. Compounds are labeled according to the data set.

set of compounds collected from the Novartis database. These compounds were selected using the INH core structure as a first filter, a similarity level of 40% and a MIC threshold value of less than 400 μg/mL. A sub-set of 13 compounds from those returned from the database was chosen so that the variability of both the dependent and independent variables fell into the same range as that of compounds in the training and test sets of the best models. Model A3 did not perform as well as model B3 and so in Table 2 we present the results retrieved only from model B3 for the Novartis data set. As can be seen from the metrics displayed and also from Fig. 7, model B3 shows a good predictive ability for the Novartis data set which is a further important criterion to assess its quality.

The six MLR models displayed in Table 2 also perform well in terms of applicability domain. In fact, training and test set compounds of all models fall within ±3 SR, signaling the absence of any outlier (see Table S5, in Supporting Information). Moreover, there are also no significant influential compounds in the training sets since when h > h* (models A1, A2 and B2) the distance to h* is quite small [64,79]. Fig. 8 shows as an example the Williams plot for model A3 (see also Table S5). Three compounds from the test set (compounds 7, 13 and 49) have leverage values slightly higher than the cut-off value, h*, being somewhat deviated from the structural centroid of the training set thus falling slightly outside the model AD. However, their residuals are close to zero and so they are accurately predicted by the model. Data points such as these are

Table 2
Best found MLR models.[a]

Sets	$a_0 \pm s(a_0)$[b] (SL)[c]	$a_1 \pm s(a_1)$ (SL)	$a_2 \pm s(a_2)$ (SL)	$a_3 \pm s(a_3)$ (SL)	$a_4 \pm s(a_4)$ (SL)	N[d]	SD[e]	R^2 [f]	F[g]	R^2_0 [h]	AE[i]	AAE[j]	RMSE[k]	Q^2 [l]	$\bar{r}^2_m$ [m]	Δr^2_m [n]	CCC[o]	Models
	$\log(1/MIC) = a_0 + a_1\,d_3 + a_2\,B_1 + a_3\,B_5 + a_4\,L$																**Model A**	
Training	11.562 ± 2.183 (> 99.99 %)	−10.163 ± 1.504 (> 99.99 %)	0.553 ± 0.235 (97.5 %)	—	—	34	0.562	0.732	42	—	—	—	—	0.719	—	—	—	Model A1
Test	—	—	—	—	—	20	0.363	0.833	90	0.757	−0.181	0.377	0.461	0.810	0.638	0.069	0.883	
Training	15.008 ± 2.853 (> 99.99 %)	−12.686 ± 2.023 (> 99.99 %)	0.686 ± 0.238 (99.3 %)	−0.065 ± 0.036 (91.7 %)	—	34	0.543	0.758	31	—	—	—	—	0.758	—	—	—	Model A2
Test	—	—	—	—	—	20	0.529	0.682	39	0.680	0.068	0.431	0.566	0.669	0.629	0.044	0.822	
Training	12.683 ± 2.795 (> 99.99 %)	−11.235 ± 1.956 (> 99.99 %)	0.686 ± 0.221 (99.6 %)	−0.098 ± 0.036 (98.9 %)	0.111 ± 0.044 (98.1 %)	34	0.501	0.801	29	—	—	—	—	0.801	—	—	—	Model A3
Test	—	—	—	—	—	20	0.472	0.741	52	0.738	0.040	0.413	0.504	0.739	0.696	0.006	0.857	
	$\log(1/MIC) = a_0 + a_1\,d_k + a_2\,B_1 + a_3\,B_5 + a_4\,L$																**Model B**	
Training	48.906 ± 10.271 (> 99.99 %)	−8.065 ± 1.626 (> 99.99 %)	—	—	0.094 ± 0.047 (94.2 %)	34	0.598	0.697	36	—	—	—	—	0.697	—	—	—	Model B1
Test	—	—	—	—	—	20	0.443	0.680	38	0.642	−0.003	0.454	0.556	0.682	0.602	0.112	0.799	
Training	41.638 ± 10.116 (99.97 %)	−7.040 ± 1.585 (99.99 %)	0.542 ± 0.234 (97.2 %)	—	0.078 ± 0.045 (90.7 %)	34	0.560	0.743	29	—	—	—	—	0.743	—	—	—	Model B2
Test	—	—	—	—	—	20	0.494	0.648	33	0.607	−0.028	0.500	0.583	0.650	0.573	0.112	0.790	
Training	55.818 ± 10.676 (> 99.99 %)	−9.321 ± 1.684 (> 99.99 %)	0.720 ± 0.224 (99.7 %)	−0.098 ± 0.037 (98.7 %)	0.125 ± 0.045 (99.1 %)	34	0.511	0.793	28	—	—	—	—	0.793	—	—	—	Model B3
Test	—	—	—	—	—	20	0.468	0.717	46	0.714	0.066	0.437	0.526	0.715	0.668	0.016	0.837	
NOVARTIS	—	—	—	—	—	13	0.368	0.629	19	0.534	−0.168	0.349	0.395	0.681	0.475	0.080	0.745	

[a] Values are presented with three decimal places to minimize round-off errors in subsequent calculations.
[b] Standard deviation of coefficient.
[c] Significance level of coefficient.
[d] Number of compounds.
[e] Standard deviation of fit.
[f] Determination coefficient.
[g] Fisher-Snedecor statistics.
[h] Determination coefficient of regression through the origin.
[i] Average error.
[j] Absolute average error.
[k] Root mean squared error.
[l] Cross validation correlation coefficient.
[m] Average value of Roy's metrics.
[n] Absolute difference between r^2_m values.
[o] Concordance correlation coefficient.

Table 3
Results of Y-randomizations for models A3 and B3 in Table 2.

Models	R^2	R_r^2	$^cR_p^2$
A3	0.741	0.147	0.663
B3	0.717	0.125	0.651

sometimes called "good high leverage points or "good influential points" [66].

Looking at the descriptors for the two best models, A3 and B3, one can see that higher values of interatomic distances (d_3 and d_K) lead to a decrease in the activity and the same effect is observed when the volume of the substituent, measured by B_5, increases. On the contrary, the molecular specificities encoded by B_1 and L enhance the activity of these INH derivatives against *Mtb*.

The observation that the chain length of the substituent promotes the activity of these compounds is indeed remarkable. Actually, the favorable effect on the activity of INH derivatives caused by an increase in the chain length of the substituent has already been reported in the literature [41]. The authors suggested that the effect was due to an enhanced ability of the larger compounds to penetrate the cell membrane. However, as already referred in our previous studies [35], and again in this work, the lipophilicity character of the INH derivatives, as measured by ClogP, is not found to be critical to explain the biological behavior of these compounds against *Mtb*. In fact, for the 82 INH derivatives that constitute the object of our study, the findings did not reveal any significant correlation between ClogP and the biological activity expressed in terms of log (1/MIC). It should also be noted that the intercorrelation between ClogP and L is very low (0.248). These observations might thus imply at least one of two hypotheses: (i) the penetration process through the cell membrane is indeed not determinant for the antitubercular activity of these compounds; or (ii) the measure of lipophilicity by an isotropic lipophilicity parameter such as the *n*-octanol–water partition coefficient is not an adequate model to mimic these specific drug–membrane interactions [35]. It is known that there are potential antitubercular compounds with either strong or weak lipophilic/hydrophilic characteristics [88] and the importance of the lipophilicity/hydrophilicity balance on the activity of antitubercular agents has been recognized by various authors [44,93–95]. However, the relationship between lipophilicity and activity is not always evident, making it difficult to define a unique and consensual optimal range for the lipophilicity of a good antitubercular drug candidate.
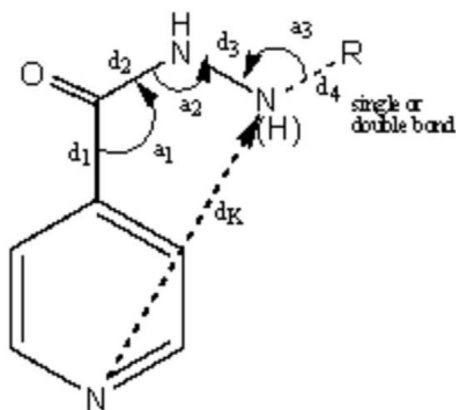
Table 4
Intercorrelation matrices.

1. Between any two descriptors					
Descriptors	d_3	d_k	B_1	B_5	L
d_3	–		0.220	0.593	0.458
dk			0.204	0.578	0.422
B1				0.287	0.156
B5					0.493

2. Between one descriptor and a linear combination of all others				
Model A3 (d_3, B_1, B_5, L)				
Descriptors	B_1, B_5, L	B_5, L, d_3	L, d_3, B_1	d_3, B_1, B_5
d_3	0.634			
B_1		0.295		
B5			0.682	
L				0.538
Model B3 (d_k, B_1, B_5, L)				
Descriptors	B_1, B_5, L	B_5, L, d_k	L, d_k, B_1	d_k, B_1, B_5
d_k	0.607			
B_1		0.292		
B_5			0.684	
L				0.525

3.3. QSAR-oriented design and compound selection

The obtained classification models from RFs and ASNNs methods were used to automatically screen a data set of 430 virtual compounds, potentially active against *Mtb*, and with distinct substitution patterns − see Table S2. These compounds belong to two broad INH derived families, the isonicotinoyl hydrazide family and the isonicotinoyl hydrazone family, and were classified as active/inactive by the eight classification models. A first ranking of these compounds was achieved taking into account their average probability to be active, determined by the probability measure of the RFs. After this initial ranking, two additional pruning operations were performed: i) virtual compounds classified as inactives by a given classification model were deleted; and ii), virtual compounds with an average probability of being active lower than 0.8 were also deleted. As a consequence, the final number of virtual compounds suggested for synthesis was reduced to 118 (see Section 2.1 on *Data sets*). A further refinement of the virtual screening using the web interface OSIRIS
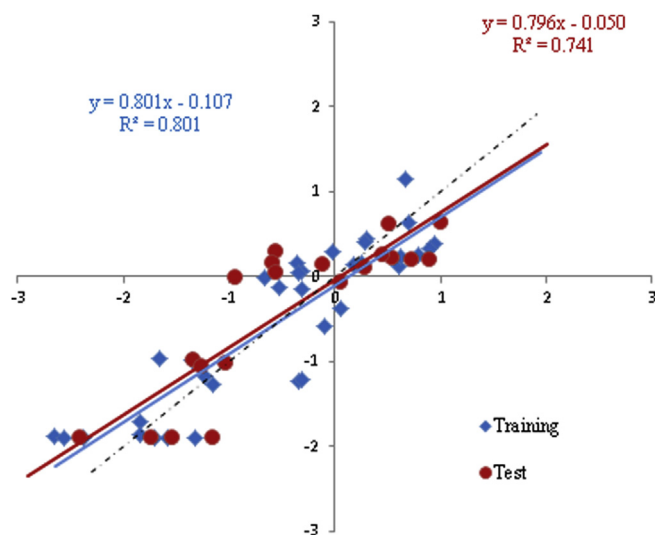

Fig. 5. Representation of the geometrical descriptors considered for isoniazid derivatives.

Fig. 6. Log(1/MIC)$_{pred}$ *vs.* log(1/MIC)$_{exp}$ according to model A3.

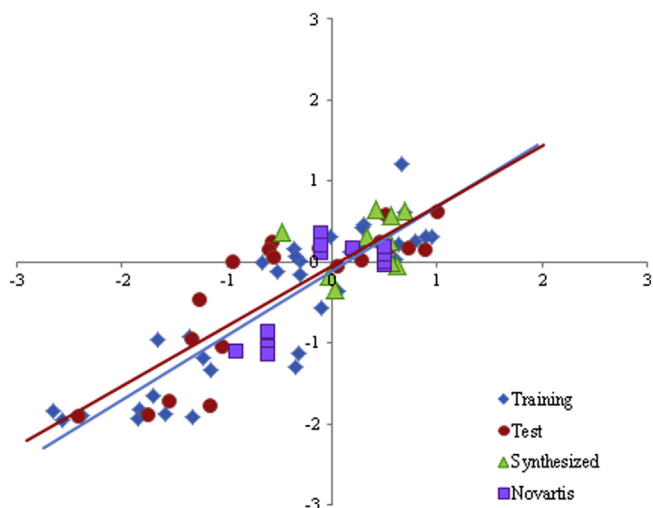**Fig. 7.** Log(1/MIC)$_{pred}$ *vs.* log(1/MIC)$_{exp}$ according to model B3, with predicted values for the synthesized (△) and the Novartis (□) compounds.

[96], allowed an evaluation of these compounds in terms of predicted drug relevant properties such as toxicity (mutagenicity, tumorigenicity, irritation and reproductive effectiveness), drug-likeness, drug score, ClogP, solubility and synthetic accessibility. Based on this information, a subset of 44 compounds was proposed for synthesis. From these, 13 were chosen to be synthesized due their synthetic feasibility, but were first also evaluated in terms of their predicted activity by the MLR models referred to in Section 3.2.2 (estimated MIC values had to be ≤5 μM for a compound to proceed for synthesis).

From the best found MLR models, 17 compounds with suitable molecular characteristics were evaluated, including the 13 proposed by the classification models. Compounds were designed so that they would lead to high activity values as predicted by the majority of the models, *i.e.*, so that their descriptors B_1 and L were higher than the average values in the training set, and B_5 and d_3 (or d_k) were lower than those values. This balance was, however, difficult to attain and therefore we
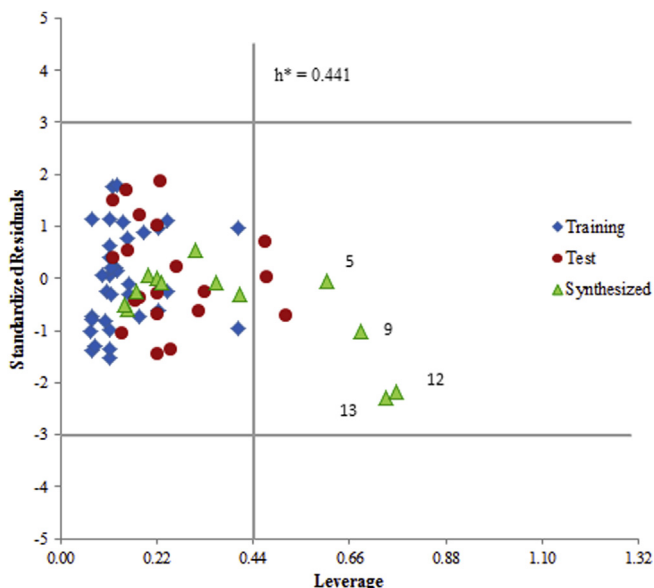

Fig. 8. Williams plot for model A3.

also had to take into consideration the relative weights of the descriptors to explain the antitubercular activity. Table 5 presents the descriptors values computed for the 17 compounds, as well as maximum and minimum values for training and test set descriptors and mean values for training set descriptors. Values in bold correspond to descriptors that fulfill the specified requisites (see 3.2.2). Based on the 6 MLR models presented in Table 2, MIC values (expressed in μM) were subsequently estimated for each of these compounds and results are shown in Table 6. Using the same activity threshold as above, 15 INH derivatives with MIC ≤5 μM, as predicted by at least four MLR models, which concomitantly passed the referred OSIRIS filters, were selected as potential antitubercular agents (in bold in Table 6). From these, 10 compounds (including 6 suggested by the consensus approach) were chosen to be synthesized due to their greater synthetic workability, namely compounds **1**−**10**, in bold and underlined in Table 6.

However, before progressing with the synthesis, we verified the compliance of the designed compounds with the applicability domain of the derived models by introducing them in the Williams plot represented in Fig. 8 for model A3. All the synthesized compounds lie well within ±3 *SD* and the great majority of them are structurally close to training set compounds as their leverage values are lower than the warning leverage value, being therefore within the AD of the models. Obvious exceptions are compounds **5** and **9** which, despite being reliably predicted (low standardized residuals), only display h < h* for two of the models (see Table S5), and hence are farther from the training set structural centroid. Yet, this procedure reinforced the decision to go on with the synthesis, since compounds **5** and **9** fall into the category of "good high leverage points" according to [66] (see Section 3.2.2). The Williams plot for model B3 is presented in Fig. S2 as Supporting Information. Similar trends are observed for all the represented sets, being one of the Novartis compounds structurally more distant from the model centroid than all the others, even if it is the most-well predicted compound from the set.

Although compounds **12** and **13**, shown in Tables 5 and 6, have only been synthesized for comparative purposes, it is easily seen from Fig. 8 (and Fig. S2 in SI) and Table S5 that they have $h >> h^*$, and are therefore structurally remote from the model centroid, besides having relatively high *SD*, close to the lower limit of the AD. This observation is consistent with their behavior in terms of descriptors' range and MIC values (see Tables 5 and 6).

From the 10 selected compounds only two belong to the isonicotinoyl hydrazide family (compounds **9** and **10**) the remaining being isonicotinoyl hydrazones. In terms of the considered molecular descriptors (Table 5), these two compounds differ from all others especially in the interatomic distance, d_3, and in the length of the substituent, *L*.

On the other hand, these 10 compounds have very distinct lipophilicity characteristics, with experimental log$P_{o/w}$ ranging from 0.220 (compound **5**) to 3.700 (compound **8**) [88,97]. This large spectrum of lipophilicity values in compounds with high predicted activity corroborates the aforementioned statement that lipophilicity is not *per se* a good criterion to judge the suitability of a new compound as an antitubercular agent.

In order to further extend the variability of logP, we have also synthesized compound **13**, a high hydrophilic compound just as INH (log$P_{o/w}$ (INH) = −0.851 [88]) even if its descriptors do not follow, in general, the tendency expected to promote activity (please refer to Table 5), as can be confirmed by the high estimated MIC values in Table 6. Furthermore, to be able to compare the effect upon the activity of having a hydrazone or the corresponding hydrazide analogue, we have likewise synthesized compound **12**, the reduced counterpart of compound **6**. Again, estimated MIC values

Table 5
Descriptors and logP values for compounds evaluated by the MLR models.

Compounds	Sub-family/Substituent[a]	Descriptors[b]					$logP_{o/w}$[c] (exp)
		d_3	B_5	B_1	L	d_k	
1	3B/22	**1.2323**	**5.5377**	1.7000	**7.7103**	**6.1382**	1.326
2	3C/22	**1.2324**	7.2014	1.7000	**6.6826**	**6.1377**	1.317
3	3B/41	**1.2322**	**5.5363**	1.7000	**8.2845**	**6.1385**	0.902
4	3C/41	**1.2323**	6.9931	1.7000	**6.7337**	**6.1382**	0.731
5	–	**1.2310**	**3.3452**	**2.2160**	5.4760	**6.1295**	0.220
6	–	**1.2306**	6.9692	**2.0599**	9.0432	**6.1338**	**1.965**[d]
7	3C/17	**1.2322**	8.9992	1.7000	**7.0870**	**6.1393**	1.317
8	3C/37	**1.2319**	10.5646	1.7000	**6.7968**	**6.1360**	3.700
9	–	1.3396	8.7060	1.7644	**12.9649**	6.1374	3.459
10	–	1.3432	**6.0991**	**2.0294**	9.6395	6.1717	**2.090**[d]
12[e]	–	1.3810	9.2267	**2.0798**	4.9223	6.3135	2.833
13[e]	–	1.3797	**1.1700**	1.1700	2.4158	6.2876	−0.541
							ClogP[f]
14	3C/34	**1.2316**	11.1639	1.7000	**7.3171**	**6.1315**	1.211
15	3B/12	**1.2319**	6.9213	1.7000	**6.2980**	**6.1350**	2.117
16	3B/24	**1.2317**	7.9273	1.7000	**6.2874**	**6.1328**	0.570
17	3A/22	**1.2319**	**5.5622**	1.7000	**6.9309**	**6.1347**	0.042
18	3B/17	**1.2316**	**5.5405**	1.7000	**9.1565**	**6.1353**	1.587
19	2A/1	1.3802	**6.1009**	1.7193	4.7461	6.3175	0.469
20	2C/1	1.3806	6.7924	1.7132	5.0700	6.3184	0.519
Mean value in training set	–	1.2917	6.6176	1.8558	6.0279	6.2008	–
Maximum value in training set	–	1.3815	17.7653	3.3884	12.4097	6.3148	–
Minimum value in training set	–	1.2297	1.1700	1.1700	2.1850	6.1259	–
Maximum value in test set	–	1.3814	11.9496	2.4492	8.7913	6.3089	–
Minimum value in test set	–	1.2296	1.1700	1.1700	2.1870	6.1272	–

[a] Please refer to Table S2 for structures.
[b] Values in bold correspond to descriptors whose values promote activity when compared with mean values in training set, according to models in Table 2.
[c] Experimental octanol/water partition coefficient ($logP_{o/w}$) determined by the shake-flask method in refs. [88] and [97].
[d] $logP_{o/w}$ experimentally determined in this work by the shake-flask method.
[e] Compounds **12** and **13** have been synthesized for comparative purposes.
[f] Estimated $logP_{o/w}$ values determined by means of ClogP through ChemDraw, ref. [58].

for compound **12** are much higher than the established threshold value of 5 μM, although its $logP_{o/w}$ value lies within the limits of lipophilicity defined by the other compounds.

3.4. Biological activity

The synthesized compounds were tested *in vitro* against three *Mtb* strains: the H37Rv ATCC27294[T] strain, a *Mtb* clinical strain with

Table 6
MIC values (in μM) predicted by models A1 to B3.

Compounds[a,b]	Model A1	Model A2	Model A3[c]	Model B1	Model B2	Model B3[c]
<u>1</u>	1.0494	0.6627	0.4845	0.7509	1.1229	0.5626
<u>2</u>	1.0502	0.8516	0.9188	0.9293	1.3409	1.0899
<u>3</u>	1.0457	0.6597	0.4167	0.6670	1.0175	0.4796
<u>4</u>	1.0496	0.8248	0.8645	0.9266	1.3381	1.0345
<u>5</u>	0.5271	0.2030	0.2230	1.0339	0.7654	0.2307
<u>6</u>	0.6372	0.4424	0.2582	0.5187	0.5248	0.2646
<u>7</u>	1.0453	1.1089	1.2386	0.8772	1.2794	1.5040
<u>8</u>	1.0389	1.3918	1.8884	0.8774	1.2766	2.1645
<u>9</u>	11.9046	22.1251	3.7643	0.2380	0.3973	0.2225
<u>10</u>	9.2605	10.8655	3.4943	0.9175	0.9096	0.4364
12[d]	20.9857	48.7145	58.4770	35.5198	19.7877	66.1828
13[d]	64.8971	58.9162	72.8658	37.6827	63.4557	57.6342
14	1.0308	1.5081	1.8776	0.7219	1.0814	1.9369
15	1.0397	0.8064	0.9408	0.9603	1.3753	1.0790
16	1.0352	0.9328	1.1790	0.9239	1.3296	1.2946
17	1.0398	0.6576	0.5884	0.8329	1.2212	0.6575
18	1.0319	0.6493	0.3290	0.5203	0.8253	0.3481
19	32.6245	52.6503	52.2265	39.7614	34.1821	68.2964
20	33.1831	59.6566	57.3596	37.6674	32.9455	74.7712

[a] In bold compounds predicted as active (MIC ≤ 5 μM) by at least 4 MLR models.
[b] In bold and underlined compounds selected for synthesis.
[c] Best MLR models.
[d] Compounds **12** and **13** have been synthesized for comparative purposes.

a *katG* S315T mutation and the INH resistant strain H37Rv ATCC27294[T] (H37RvINH) with a full deletion in the *katG* gene, as referred in Section 5.2.

Results are shown in Table 7 where INH, the reference compound, is introduced for comparative purposes.

3.4.1. Activities against the wt strain

The ten designed compounds show experimental MIC values (expressed in μM units) against the wild type strain of *Mtb* very close to values predicted by all MLR models which give predictions for these compounds < 5 μM (column 3, Table 7) − *i.e.*, by a consensus prediction [36,38] − and by the two best MLR models, models A3 and B3 (column 5, Table 7). The mean average deviations (*AE*) are, respectively, 0.25 and 0.36. Within experimental uncertainty, compounds are more active than predicted in all cases, with the exception of compound **10**.

A very relevant result is that five out of the 10 compounds (compounds **1**, **2**, **4**, **5** and **6**) for which the majority of descriptors clearly contribute to enhance the activity (see Table 5), show a measured activity against H37Rv higher than that of INH, measured with the same protocol, BACTEC 960, in the same laboratory (please refer to Section 5.2) - Fig. 9. These findings clearly refute the arguments of some authors who claim that INH analogous backbones never achieve the (*in vitro*) efficiency of INH of which they would be nothing else than mere precursors, thus postulating that INH is not a lead compound [94]. If the action of these derivatives would be to simply produce, completely and readily, INH, as referred [94], a higher activity than INH would never be observed.

Compared to the previous compounds, compound **3**, rather unexpectedly based solely on its descriptor values, is approximately twice less active than INH. Additionally, compounds **7** to **9** have a very unfavorable B_5 (taken as the effective volume of the whole

Table 7
Predicted and experimental MIC values (μM) against wild-type *Mtb* and experimental MIC (μM) values against mutant *Mtb*.

| Compounds | MIC/μM | | | | | MIC/μM | |
| | H37Rv | | | | | Clinical mutation *katG* (S315T) | H37Rv$_{INH}$ (Δ*katG*) |
	Exp.	Pred.[a]	AE[b]	Pred.[c]	AE[b]	Exp.	Exp.
INH[d]	0.29	–	–	–	–	43.8	>72.9
1	0.28	0.77	0.49	0.52	0.24	>39.5	>39.5
2	0.24	1.03	0.79	1.00	0.76	>39.5	>39.5
3	0.47	0.71	0.24	0.45	−0.02	39.2	>39.2
4	0.27	1.01	0.74	0.95	0.68	39.2	>39.2
5	0.20	0.50	0.30	0.23	0.03	39.4	>49.2
6	0.27	0.44	0.17	0.26	−0.01	34.1	>34.1
7	1.06	1.18	0.12	1.37	0.31	21.2	n.d.[e]
8	0.95	1.44	0.49	2.03	1.08	18.9	n.d.
9	0.38	1.16	0.78	1.99	1.61	6.9	>34.3
10	3.04	1.44	−1.60	1.97	−1.07	>38.0	n.d.
11	44.65	–	–	–	–	>55.8	>55.8.
12	0.44	–	–	–	–	>33.9	>33.9
13	0.42	–	–	–	–	47.9	n.d.

[a] Mean value of at least four MLR models.
[b] Average error (pred.−exp.).
[c] Mean value of two best MLR models.
[d] Reference compound.
[e] Not determined.

substituent) which seems to cancel out the positive effects of the other descriptors. Compound **10**, on the other hand, has a consensus predicted activity of 1.44 μM but is the less active compound of this set. These observations cannot be rationalized in terms of simple electronic effects, which in any case did not appear as relevant descriptors in any of the best models. The behavior of these compounds is most probably related with the steric characteristics of the R substituents and their ability or incapability to promote the formation of the isonicotinoyl radical postulated to be formed in the generally accepted mechanism of action of INH [18,24,98,99]. In fact, it is reported in the literature that access to the heme active site of KatG poses steric constraints and therefore any factor influencing the stereochemistry of these compounds should be relevant [100,101]. Some authors [102] suggest that the formation of the INH analogues adducts with NAD$^+$ requires acyl radicals sufficiently long-lived, and that, for aromatic hydrazides, the nature of substituents should play an important role in the stabilization of radical intermediates involved in the overall process of INH activation.
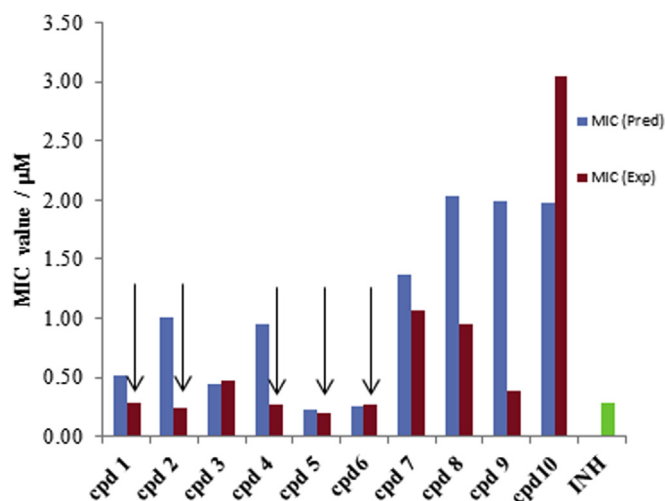
Fig. 9. Predicted and experimental MIC values for the 10 selected compounds against H37Rv (MIC values given by an average prediction based on the two best MLR models).

A comparison between the hydrazone compound **6** and the corresponding hydrazide compound **12**, shows that the hydrazone is more active than its analogue (0.27 μM *vs.* 0.44 μM), as already perceived by some authors [39] and contested by others [94], based on the assumption that MIC measurements in the two situations were not standardized, which is clearly not the case here. The higher activity of hydrazones is possibly due to a higher efficacy in the formation, *in loco*, of the prodrug INH, as a result of hydrazone hydrolysis, and the subsequent production of the acyl radical, as a result of INH activation by KatG.

The biological behavior of the isonicotinoyl hydrazide compounds **12** and **13**, is very similar, with reasonably low experimental MIC values (0.44 and 0.42 μM, respectively) in spite of their (not surprisingly) high estimated MIC values (41.61 and 59.24 μM, respectively). These compounds show, however, very different lipophilic characteristics as given by their log$P_{oct/w}$ (Table 5), which, again, further supports the inference that there is no clear relationship between lipophilicity and *in vitro* activity.

To summarize, the information in Tables 5 and 6 shows that N–N distances (d_3 and d_k) and the length of the substituent (*L*) are the properties with the most impact in the activity of these compounds. Shorter distances and lengthy substituents appear to favor the activity by allegedly facilitating the production of the isonicotinoyl radical.

3.4.2. Activities against resistant strains

Given the results obtained for the *Mtb wt* strain, we carried out the testing of these compounds against INH mono-resistant strains whose mechanism of resistance is related only to the diminished effectiveness of KatG. Values in Table 7 reveal that seven of the synthesized compounds are more effective than INH against the mutated S315T resistant strain, *i.e.*, compounds **3**–**9**, with compounds **7**–**9** showing a significant decrease in MIC − Table 7 and Fig. 10. Among these seven INH derivatives, six are isonicotinoyl hydrazones, being compound **9** the only isonicotinoyl hydrazide. However, it is precisely this compound that presents the most significant difference in activity showing a six fold increase in activity by comparison with INH.

It has been reported in the literature that *katG* (S315T) confers resistance to INH through subtle changes in the INH binding site [103], but without losing its ability to bind INH [98,104]. Moreover,
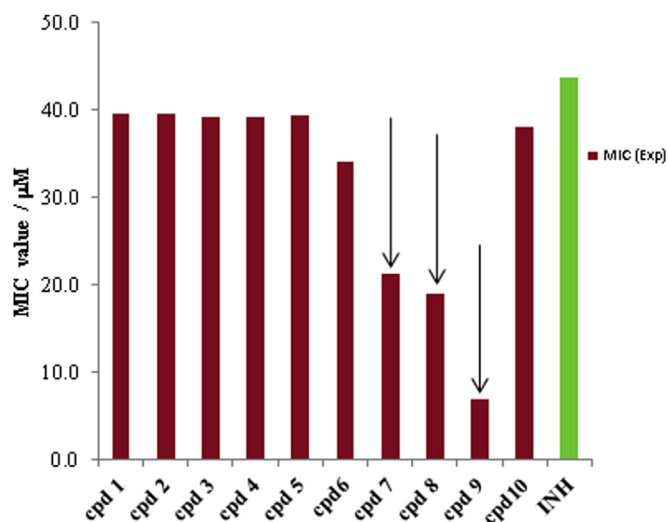
Fig. 10. Experimental MIC values for the 10 selected compounds against mutated *katG* (S315T). MIC values for compounds **1**, **2** and **10** represent the lower limit obtained.

it has been described that the addition of a methyl group in the mutated *katG* actually constricts the accessibility to the heme active site by closing down the dimensions of the narrowest part of the access channel from 6 Å in the *wt katG* to 4.7 Å in the S315T mutated *katG* [100–102]. This steric effect is taken as interfering with the interaction between *katG* and INH, leading some authors to suggest that INH binding within the heme pocket is most probably a prerequisite for INH activation [101]. Furthermore, these steric effects on INH binding allegedly have an impact upon INH oxidation, thus decreasing the conversion of INH to a biologically active form [105,106]. However, the effect of any perturbation on INH binding upon drug activation needs to be further analyzed [102].

Preliminary docking studies of INH with both *wt* and mutated KatG carried out by us have shown that the S315T mutation subtly narrows the accessibility channel to the catalytic centre of KatG but the *in silico* binding constants of INH to both strains turned out to be similar within the method uncertainty. Although the exact binding site of INH to the CP-KatG *wt* enzyme (and for that matter, also to the KatG (S315T)) is still controversial [102] and needs further investigation, recent studies [100,103,107] indicate that it might be close to the KatG heme pocket, at about 12 Å from the heme iron for both enzymes. Therefore, any factor affecting the stereochemistry of a compound might be important to rationalize its activity, especially in the mutated strain.

These evidences make the performance of compound **9** in the mutated strain even more peculiar because its molecular structure, with a long alkyl chain, does anticipate large steric constrictions. To have a measure of the importance of the alkyl group upon the antitubercular activity, we have synthesized the acetyl analogue, compound **11** (*N'*-acetylisonicotinohydrazide), which showed a very low activity both in the *wt* strain and in the mutated strain. Hence, the C10 chain of compound **9** seems indeed crucial to enhance its mycobacterial activity. Some authors have also reported improved activities for other long chain hydrazides but there seems to be no consistency between the length of the chain and the corresponding MIC value [41,87,108]. Regarding the effectiveness of compound **9**, our hypothesis is that the active site in the mutated strain might be located in a more peripheral area of the KatG enzyme and/or that the compound–enzyme interaction may be more efficient.

Since we do not expect that a substitution of a serine by a threonine at amino acid residue 315 shall introduce very significant

modifications in the active site of KatG, molecular dynamics studies to model the behavior of both compounds, *i.e.*, INH and compound **9**, *vis-à-vis* the wild type and the mutated strain, as well experimental determinations of binding constants in these four situations are underway and will hopefully give us further insights into the real molecular picture of these drug–target interactions.

Results presented in this work seem to suggest that redesign of the INH molecule in order to improve drug binding may indeed be a feasible approach to overcome resistance in *katG* (S315T), the most frequent and important cause of INH resistance worldwide [109]. In support of this hypothesis is the total ineffectiveness of the new INH derivatives against the H37Rv$_{INH}$ (Δ*katG*) that suffered a full deletion of the *katG* gene after prolonged exposure to INH (Table 7) which confirms the importance of the activation role of KatG upon INH and its derivatives.

New INH derivatives redesigned to take advantage of the diminished but still effective activation of INH by the mutated *katG*, would thus allow the maintenance in the therapeutic regimen of one of the two most active and effective anti-TB drug known so far, *i.e.*, isoniazid [28,110].

3.5. Cytotoxicity assays

In parallel with the antitubercular activity evaluation, all compounds with MIC values against the H37Rv strain <0.5 μM were screened towards African green monkey kidney cells (VERO cells) using the MTT assay, in order to determine their *in vitro* cytotoxicity (IC$_{50}$) and to establish the selectivity index (SI). Results showed that these compounds were non-toxic to the host cells in the concentration range studied (1 MIC–500 MIC, with MIC in μg/mL) — Table 8. Isoniazid was also analyzed and had a IC$_{50}$ value >200. The SI was determined as the ratio of the measured IC$_{50}$ to the measured IC$_{99}$ (H37Rv *Mtb* MIC value) and was in all cases > 400 (for INH, SI > 500). SI is commonly used to estimate the therapeutic window of a drug and to identify drug candidates for further studies. According to literature [49,111,112] candidates for new drugs must have a SI ≥ 10, with MIC values lower than 6.25 μg/mL and a low cytotoxicity, as is indeed the case for all tested compounds (compounds **1**–**6** and **9**).

4. Conclusions

The use of a rational design approach based on well-validated classification and regression QSAR models has allowed a successful synthesis of several potent antitubercular isoniazid derivatives.

Table 8
Cytotoxicity (IC$_{50}$) and selectivity index (SI) for the synthesized compounds.[a]

Compound	IC$_{50}$ (Vero cells)		SI IC$_{50}$/MIC
	μg/mL	μM	
INH[b]	>27.43	>200	>686
1	>50.65	>200	>724
2	>50.65	>200	>844
3	>51.05	>200	>425
4	>51.05	>200	>729
5	>38.62	>190	>965
6	>58.67	>200	>733
7	n.d.	n.d.	n.d.
8	n.d.	n.d.	n.d.
9	>58.28	>75	>530
10	n.d.	n.d.	n.d.
11	n.d.	n.d.	n.d.
12	>59.08	>200	>454
13	>33.43	>200	>478

[a] Evaluation for cytotoxicity in VERO cells at concentrations up to 500 times the MIC for *Mtb* H37Rv. The activity/cytotoxicity criterion is an SI > 10.
[b] Reference compound.

Compounds **1**, **2**, **4**, **5** and **6** showed an activity against H37Rv higher (MIC ≤ 0.28 μM) than the one exhibited by the reference compound, INH (MIC = 0.29 μM). It is postulated that the behavior of these compounds is probably dependent on stereochemical aspects associated with the nature of the substituents and on their ability to promote the formation (and stabilization) of the iso-nicotinoyl radical in the overall process of activation by KatG. As suggested in our previous works, lipophilicity was again not identified as an important factor to describe the activity of these compounds, which was now further confirmed by the absence of correlation between measured $logP_{o/w}$ and MIC values.

On the other hand, compounds **3** to **9** showed to be more effective than INH against a *Mtb* clinical strain carrying only a *katG* S315T mutation. In particular, compound **9**, with an alkyl chain in C10, showed a six fold increase in activity against this resistant strain. This finding seems to question the recently proposed relationship between the increased resistance of *katG* (S315T) to INH and a putative larger steric constraint in the access channel to the heme active site in the mutated strain. Experimental and *in silico* studies to evaluate the interaction of compound **9** with the biological target in both strains are now being carried out and will be reported in due course.

Finally, the ineffectiveness of all the new INH derivatives against the INH high-level resistant strain H37Rv$_{INH}$ (Δ*katG*) with a full deletion of the *katG* gene, indubitably attests the crucial role of KatG in the activation process of these compounds.

In summary, our results clearly demonstrate that using a properly validated QSAR methodology it is possible to design more active compounds departing from the INH core structure and overcome the decreased susceptibility of the *katG* mutated strains towards INH.

5. Experimental section

5.1. Chemistry

5.1.1. General

All solvents and reagents were obtained from commercial suppliers and were used without further purification. The reactions were monitored by thin layer chromatography performed on silica gel 60 F254 aluminum sheets and visualized with UV light (254 nm). Flash column chromatographies were performed on silica-gel 60, 40–63 μm. Melting points were taken in open capillary tubes using an apparatus Stuart® SMP 30 or in a Reichert Thermovar Microscope and are uncorrected. IR spectra were obtained using a Mattson Satellite FTIR and only the diagnostic absorption bands were reported, in cm^{-1}. ^{1}H NMR (400 MHz) and ^{13}C NMR spectra (100.6 MHz) were recorded on a Bruker AMX spectrometer; chemical shifts were expressed as δ values and referenced to the residual solvent peak (DMSO-d_6, δ_H = 2.50); coupling constants were reported in units of Hertz (Hz). HRMS were recorded in a mass spectrometer Apex Ultra FTICR−MS. The purity of all compounds (≥98%) was assessed by GC−MS analysis using a ThermoQuest 8000 series chromatograph coupled with a MS Trio 1000-Fisions [DB-1 coated capillary column (5 m × 0.53 mm × 2.65 μm) using N$_2$ as carrier gas; the GC oven was programmed at 120 °C (2 min), heated to 200 °C at 10 °C min^{-1}, held 2min at 200 °C, further heated to 280 °C at 20 °C min^{-1} and held at final temperature for 30 min] or a Surveyor Plus HPLC system (Thermo Finnigan) equipped with a Surveyor Autosampler Plus with a 100 μL loop, a quaternary pump Surveyor LC Pump Plus, and a diode-array detector Surveyor PDA Plus Merck [column Lichrospher®100 RP-18 column(250 × 4 mm, particle size 5 μm), gradient elution (40% acetonitrile to 100% 0.1% TFA over 20 min and return to the initial conditions within 13 min)].

5.1.2. General procedure for the synthesis of hydrazones 1 and 2

Isoniazid (1 equiv.) was added in portions over a period of three hours to a solution of the dialdehyde (1 equiv.) in absolute ethanol. The mixture was stirred at room temperature for 24 h, the resulting precipitate was filtered off and the mother liquor concentrated to afford a crude product which was further purified by flash chromatography and recrystallization.

5.1.2.1. N′-(E)-(3-methanoylbenzylidene)isonicotinohydrazide (compound 1). White crystals (flash column chromatography with EtOAc/MeOH 93:7 and recrystallization from ethanol/water 1:1), yield: 41% (1.027 g), mp 210.1−211.5 °C. IR (KBr): v (cm^{-1}) = 3300 (NH amide), 3045−2854 (Ar C−H), 1676 (C=O amide and aldehyde), 1550 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 12.39 (s, 1H, N$\underline{H}$), 10.01 (s, 1H, C$\underline{H}$O), 8.81 (AA′ part of a AA′XX′ system, 2H, J = 4.6, 1.4 Hz; H-2, H-$\underline{6}$), 8.56(s, 1H, C$\underline{H}$=N), 8.29 (s, 1H, H-2′), 8.07 (d, 1H, J = 7.8 Hz), 8.00 (d, 1H, J = 7.6, H-4′), 7.85 (XX′ part of a AA′XX′ system, 2H, J = 4.6, 1.5 Hz, H-3, H-5), 7.72 (t, 1H, J = 7.7 Hz, H-5′). ^{13}C NMR (DMSO-d_6): δ = 193.0 (C$\underline{H}$O) 161.8 (NH$\underline{C}$O), 150.3 (C-2, C-6), 147.6 (C$\underline{H}$N), 140.3 (C-4), 136.6 (C-3′), 135.0 (C-$\underline{1}$′), 132.9 (C-6′), 130.8 (C-4′), 129.8 (C-5′), 127.8 (C-2′), 121.5(C-3, C-5). HRMS: *m/z* 254.09231 [M+H]$^+$(calcd. for C$_{14}$H$_{11}$N$_3$ O$_2$+H 254.09240).

5.1.2.2. N′-(E)-(4-methanoylbenzylidene)isonicotinohydrazide (compound 2). White crystals, mp 222.1−223.3 °C [lit. 219−220 °C] [89]. HRMS: *m/z* 254.09214 [M+H]$^+$(calcd. for C$_{14}$H$_{11}$O$_2$N$_3$+H 254.09240).

5.1.3. General procedure for the synthesis of hydrazones 3 and 4

1equiv of NaBH$_4$ was added, portion wise, to a solution of compounds **1** or **2** dissolved in 96% ethanol. After stirring the reaction mixture for 2 h at room temperature, water was added and the resulting precipitate was filtered off to afford a crude product which was further purified by recrystallization.

5.1.3.1. N′-(E)-3-(hydroxymethyl)benzylidene)isonicotinohydrazide (compound 3). White crystals (recrystallization from water), yield: 43% (0.108 g), mp 230.4−232.3 °C. IR (KBr): v (cm^{-1}) = 3373 (OH), 3202 (NH amide), 3034−2897 (Ar C−H), 1671 (C=O amide), 1568 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 12.08 (s, 1H, N$\underline{H}$), 8.80 (AA′ part of a AA′XX′ system), 2H, J = 5.6 Hz, H-2, H-6), 8.47 (s, 1H, C$\underline{H}$N), 7.84 (XX′ part of a AA′XX′ system, 2H, J = 5.8 Hz, H-3, H-5), 7.76 (s, 1H, H-2′), 7.60 (d, 1H, J = 7.1 Hz, H-6′), 7.41 (m, 1H, H-5′), 4.56 (d, 2H, J = 5.3 Hz C$\underline{H_2}$OH). ^{13}C NMR (DMSO-d_6): δ = 161.6 (NH$\underline{C}$O), 150.3 (C-2, C-6), 149.1 (C$\underline{H}$N), 143.3 (C-3′), 140.4 (C-4), 133.8 (C-$\underline{1}$′), 128.6 (C-4′), 126.4 (C-5′), 126.0 (C-6′), 124.7 (C-2′), 121.5 (C-3, C-5), 62.5 (C$\underline{H_2}$OH). HRESIMS: *m/z* 256.10806 [M+H]$^+$(calcd. for C$_{14}$H$_{13}$N$_3$O$_2$ + H 256.10805).

5.1.3.2. N′-(E)-4-(hydroxymethyl)benzylidene)isonicotinohydrazide (compound 4). White crystals (recrystallization from EtOH/H$_2$O 1:1), yield: 29% (0.119 g), mp 220.1−221.3 °C. IR (KBr): v (cm^{-1}) = 3399 (OH), 3300 (NH amide), 3045−2854 (Ar C−H), 1667 (C=O amide), 1567 (C=N imine). ^{1}H NMR (CD$_3$OD): δ = 8.77 (AA′ part of a AA′XX′ system, 2H, J = 4.6, 1.5 Hz, H-2, H-6), 8.38 (s, 1H, C$\underline{H}$N), 7.91 (XX′ part of a AA′XX′ system, 2H, J = 4.4, 1.6 Hz, H-3, H-5), 7.84 (AA′ part of a AA′XX′ system, 2H, J = 8.2 Hz, H-2′, H-6′), 7.46 (XX″ part of a AA′XX′ system, 2H, J = 8.0, H-3′, H-5′), 4.67 (s, 2H, C$\underline{H_2}$OH). ^{13}C NMR (CD$_3$OD): δ = 163.2 (NH$\underline{C}$O), 150.3 (C$\underline{H}$N), 149.7 (C-2, C-6), 144.6 (C-4′), 141.1 (C-4), 132.8 (C-$\underline{1}$′), 127.7 (C-2′, C-6′), 126.8 (C-3′, C-5′), 121.8 (C-3, C-5), 63.3 (C$\underline{H_2}$OH). HRESIMS: *m/z* 256.10805 [M+H]$^+$(calcd. for C$_{14}$H$_{13}$N$_3$O$_2$ + H 256.10805).

5.1.4. General procedure for the synthesis of hydrazones 5–8

A solution of the aldehyde or ketone (1 equiv.) in absolute ethanol was added dropwise to a solution of isoniazid (1 equiv.) in water. The mixture was stirred at room temperature for 24 h and the resulting solid collected and washed with cold water. Flash column chromatography over silica-gel and/or recrystallization afforded pure compounds.

5.1.4.1. N′-cyclopentylideneisonicotinohydrazide (compound 5).

White crystals (recrystallization from water), yield: 63% (1.877 g), mp 182.3–183.2 °C. IR (KBr): v (cm^{-1}) = 3200 (NH amide), 3084–3013 (Ar C–H), 2976–2816 (aliph. C–H), 1658, (C=O amide), 1645 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 10.58 (s, 1H, N$\underline{H}$), 8.73 (AA′ part of a AA′XX′ system, 2H, J = 5.6 Hz, H-2, H-6), 7.72 (XX′ part of a AA′XX′ system, 2H, J = 5.6 Hz, H-3, H-5), 2.44 (t, 2H, J = 6.8 Hz, H-5′), 2.41 (t, 2H, J = 6.8 Hz, H-2′), 1.72 (m, 4H, H-3′, H-4′). ^{13}C NMR (DMSO-d_6): δ = 172.5 (C-1′), 162.2 (NH$\underline{C}$O), 150.5 (C-2; C-6), 141.7 (C-4), 122.2 (C-3; C-5), 33.6 (C-2′), 29.4 (C-5′), 24.8 (C-3′*), 24.7 (C-4′*). HRESIMS: m/z 204.11308 [M+H]$^+$(calcd. for C$_{11}$H$_{13}$N$_3$O + H 204.11314).

5.1.4.2. N′-(4-phenylcyclohexylidene)isonicotinohydrazide (compound 6).

White crystals (recrystallization from ethanol/water 1:2), yield: 74% (3.252 g), mp 178.3–180.3 °C (dec.). IR (KBr): v (cm^{-1}) = 3443 (NH amide), 2935 (Ar C–H), 1665, (C=O amide), 1636 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 11.0 (s, 1H, N$\underline{H}$), 8.74 (AA′ part of a AA′XX′ system, 2H, J = 4.6; 1.2 Hz, H-2, H-6), 7.76 (XX′ part of a AA′XX′ system, 2H, J = 4.6, J = 1.2 Hz, H-3, H-5), 7.29 (m, 4H, H-2″, H-3″, H-5″, H-6″), 7.19 (m, 1H, H-4″), 3.04 (d, 1H, J = 13.7 Hz, H-6$_{eq}$′), 2.89 (t, 1H, J = 12.08 Hz, H-4′), 2.56 (d, 1H, J = 13.6 Hz, H-2$_{eq}$′), 2.46 (ddd, 1H, J = 13.6, 13.6, 4.86 Hz, H-2$_{ax}$′), 2.09 (ddd, 1H, J = 13.7, 13.7, 5.24 Hz, H-6$_{ax}$′), 1.96 (dt, 1H, J = 12.8 Hz, H-5$_{eq}$′), 1.71 (dddd, 1H, J = 12.7, 12.7, 4.24 Hz, H-3$_{ax}$′), 1.61 (dddd, 1H, J = 12.8, 12.8, 3.6 Hz, H-5$_{ax}$′), 2.03 (m, 1H, H-3$_{eq}$′). ^{13}C NMR (DMSO-d_6): δ = 166.7 (C-1′), 161.8 (NH$\underline{C}$O), 150.1 (C-2; C-6), 145.7 (C-1″), 141.1 (C-4), 128.4* (C-2″, C-6″), 126.7* (C-3″, C-5″), 126.1 (C-4″), 121.7 (C-3; C-5), 42.4 (C-4′), 34.8 (C-2′), 33.9 (C-3′), 33.0 (C-5′), 27.3 (C-6′). HRESIMS: m/z 294.15969 [M+H]$^+$(calcd. for C$_{18}$H$_{19}$N$_3$O + H 294.16009).

5.1.4.3. Methyl 4-[(E)-(isonicotinoylhydrazone)methyl] benzoate (compound 7).

White crystals (recrystallization from acetone), yield: 59% (1.808 g), mp 216.2–217.1 °C. IR (KBr): v (cm^{-1}) = 3262 (NH amide), 3000–3126 (Ar C–H), 2956–23848 (aliph. C–H), 1716 (C=O ester), 1665 (C=O amide), 1604 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 12.26 (s, 1H, N$\underline{H}$), 8.81 (AA′ part of a AA′XX′ system, 2H, J = 4.4, 1.5 Hz, H-2′, H-6′), 8.52 (s, 1H, C$\underline{H}$N), 8.05 (AA′ part of a AA′BB′ system, 2H, J = 8.3 Hz, H-2, H-6), 7.90 (BB″ part of a AA′BB′ system, 2H, J = 8.3 Hz, H-3, H-5), 7.84 (XX′ part of a AA′XX′ system, 2H, J = 4.4, 1.5 Hz, H-3′, H-5′), 3.88 (s, 3H, OC$\underline{H}_3$). ^{13}C NMR (DMSO-d_6): δ = 166.3 (CH$_3$O$\underline{C}$O), 162.3 (NH$\underline{C}$O), 150.9 (C-2′; C-6′), 148.0 ($\underline{C}$HN), 140.7 (C-4′), 138.9 (C-4), 131.2 (C-1), 130.2 (C-2, C-6), 127.9 (C-3, C-5), 122,0 (C-3′, C-5′), 52.8 (O$\underline{C}$H$_3$). HRESIMS: m/z 284.10687 [M+H]$^+$(calcd. for C$_{15}$H$_{14}$O$_3$N$_3$+H 284.10297).

5.1.4.4. N′-(E)-(4-phenoxybenzylidene)isonicotinohydrazide (compound 8).

White crystals (flash column chromatography with EtOAc and recrystallization from ethanol/water), yield: 77% (1.552 g), mp 175.8–176.8 °C. IR (KBr): v (cm^{-1}) = 3421 (NH amide), 3103–3038 (Ar C–H), 1659, (C=O amide), 1607 (C=N imine). ^{1}H NMR (DMSO-d_6): δ = 12.65 (s, 1H, N$\underline{H}$), 8.79 (AA′ part of a AA′XX′ system, 2H, J = 5.9 Hz, H-2, H-6), 8.45 (s, 1H, C$\underline{H}$N), 7.83 (XX′ part of a AA′XX′ system, 2H, J = 5.9 Hz, H-3, H-5), 7.77 (AA′ part of a AA′XX′ system, 2H, J = 8.7 Hz, H-2′, H-6′), 7.45 (AA′ part of a AA′BB′C system, 2H, J = 7.9 Hz, H-3″, H-5″), 7.21 (C part of a AA′BB′C system,

1H, J = 7.4 Hz, H-4″), 7.10 (BB′ part of a AA′BB′C system, 2H, J = 7.7 Hz, H-2″), 7.07 (XX′ part of a AA′XX′ system, 2H, J = 8.6 Hz, H-3′, H-5′). ^{13}C NMR (DMSO-d_6): δ = 162.0 (NH$\underline{C}$O), 159.3 (C-4′), 156.1 (C-1″), 150.8 (C-2, C-6), 148.9 ($\underline{C}$HN), 141.0 (C-4), 130.7 (C-3″, C-5″), 129.7 (C-2′, C-6′), 129.4 (C-1′), 124.7 (C-4″), 122.0 (C-3, C-5), 120.0 (C-2″, C-6″), 118.7 (C-3′, C-5′). HRESIMS: m/z 318.12909 [M+H]$^+$(calcd. for C$_{19}$H$_{16}$O$_2$N$_3$+H 318.12370).

5.1.5. N′-decanoylisonicotinohydrazide (compound 9)

A solution of isoniazid (4 g, 29.2 mmol) and N-methylmorpholine (3.2 mL, 29.1 mmol) in acetonitrile (65 mL) was stirred and heated to reflux. Then, a solution of decanoyl chloride (6.0 mL, 28.9 mmol) in acetonitrile (15 mL) was added dropwise, and the reaction mixture kept at reflux for 2 h. After cooling to room temperature, the precipitate was filtered off and washed with water.

Several recrystallizations from ethanol/water 2:8 and n-hexane/ethanol 10:2.5 afforded compound **9**.

White crystals, yield: 35% (2.977 g), mp 118.4–119.2 °C [lit. 110–112 °C] [87]. HRESIMS: m/z 292.20190 [M+H]$^+$(calcd. for C$_{16}$H$_{25}$N$_3$O$_2$ + H 292.20195).

5.1.6. N′-(2-methylheptanoyl)isonicotinohydrazide (compound 10)

2-methylheptanoic acid (10 mL, 61 mmol) was refluxed with 35.8 mL (487 mmol) of thionyl chloride for two hours. The excess of thionyl chloride was removed under reduced pressure and the crude 2-methylheptanoyl chloride was used in the next step without further purification. A solution of isoniazid (8.36 g, 60.9 mmol) in acetonitrile (136 mL) was stirred and heated to reflux. Then, a solution of 2-methylheptanoyl chloride (8.96 g, 60.31 mmol) in acetonitrile (30 mL) was added dropwise, and the reaction mixture kept at reflux for 1 h and left overnight at room temperature. The reaction mixture was neutralized with a 25% solution of ammonium hydroxide and the formed precipitate washed with water and filtered off.

White solid, yield: 30.6% (4.866 g), mp 122.2–123.5 °C. IR (KBr): v (cm^{-1}) = 3194 (N–H amide), 3027 (ArC–H), 2954–2930 (aliph. C–H), 1602 (C=O amide). ^{1}H NMR (DMSO-d_6): δ = 10.79 (s, 1H, pyCON$\underline{H}$), 9.57 (s, 1H, N$\underline{H}$COCH$_2$), 8.66 (d, 2H, J = 5.0 Hz, H-2, H-6), 7.63 (d, 2H, J = 5.7 Hz, H-3, H-5), 2.47 (ddd, 1H, J = 13.8; 6.8 Hz, H-2′), 1.65 (m, 1H, H-3$_a$′), 1.42 (m, 1H, H-3$_b$′), 1.25 (s, 6H, H-4′, H-5′, H-6′), 1.17 (d, 3H, J = 6.80 Hz, H-8′), 0.86 (t, 3H, J = 5.9 Hz, H-7′); ^{13}C NMR (DMSO-d_6): δ = 175.1 (C-1′), 162.6 (py$\underline{C}$ONH), 150.4 (C-2, C-6), 138.3 (C-4), 121.1 (C-3, C-5), 39.1 (C-2′), 34.0 (C-3′), 31.7 (C-5′), 27.0 (C-4′), 22.5 (C-6′), 17,61 (C-8′), 14.0 (C-7′). HRESIMS: m/z 264.107748 [M+H]+(calcd. for C14H21N3O2+H 264.107653).

5.1.7. N′-acetylisonicotinohydrazide (compound 11)

Acetic anhydride (1.8 mL, 19.0 mmol) was added dropwise to a solution of isoniazid (2 g, 14.6 mmol) in water (20 mL). The reaction mixture was stirred at room temperature for 30 min and the resulting solid filtered off, washed with dichloromethane and recrystallized from water.

White crystals, yield: 53.5% (1.399 g), mp 163.0–163.9 °C, lit. [162.0–163.0 °C] [113]. EIMS (probe) 70 eV: m/z 179 [M]$^{+\cdot}$, 137 [C$_6$H$_7$N$_3$O]$^+$, 106 [C$_6$H$_4$NO]$^+$, 78 [C$_5$H$_4$N]$^+$, 51 [C$_4$H$_3$]$^+$.

5.1.8. N′-(4-phenylcyclohexyl)isonicotinohydrazide (compound 12)

A solution of methanolic 5 M HCl was added, dropwise, to a solution of compound **6** (1.004 g, 3.4 mmol) and NaBH$_3$CN (143 mg, 2.3 mmol) in 50 mL of methanol until pH3 was reached. The solution was stirred for 3 h, and methanol was removed under reduced pressure. The residue was flash chromatographed over silica-gel (acetone/CH$_2$Cl$_2$ 3:7). Fractions containing the reduced

compound were further chromatographed using hexane/acetone 6:4 as eluent.

White crystals, yield: 10% (0.107 g), mp 114—115 °C (dec.). IR (KBr): v (cm^{-1}) = 3296 (NH amide), 2922 (Ar C—H), 1674, (C=O amide). 1H NMR (CDCl$_3$): δ = 8.73 (AA′ part of a AA′XX′ system, 2H, J = 4.6; 1.4 Hz, H-2, H-6), 7.65 (XX′ part of a AA′XX′ system, 2H, J = 4.6, J = 1.2 Hz, H-3, H-5), 7.30* (m, 2H, H-3″, H-5″), 7.28* (m, 2H, H-2″, H-6″), 7.20 (tt, 1H, J = 7.1; 1.5 Hz H-4″), 3.36 (m, 1H, H-1′), 2.60 (m, 1H, H-4′), 1.94** (m, 4H, H-2′, H-6′), 1.69 (m, 4H, H-3′, H-5′). ^{13}C NMR (CDCl$_3$): δ = 165.6 (NH$\underline{C}$O), 150.6 (C-2; C-6), 147.1 (C-1″), 140.2 (C-4), 128.4* (C-2″, C-6″), 126.9* (C-3″, C-5″), 125.9 (C-4″), 120.9 (C-3; C-5), 53.9 (C-1′), 43.7 (C-4′), 29.1** (C-3′), 28.0** (C-2′). HRESIMS: m/z 296.17583 [M+H]$^+$(calcd. for C$_{18}$H$_{21}$N$_3$O + H 296.175739).

5.1.9. 2-Hydrazinoisonicotinohydrazide (compound 13)

Compound **13** was prepared according to a procedure reported in literature [84]. Yellow crystals (recrystallization from water), yield: 25% (0.270 g), mp 215.8—216.9 °C, [lit. 216.5—217.0 °C]. IR (KBr): v (cm^{-1}) = 3312—3123 (NH amide and amine), 3020—2920 (Ar C—H), 1653 (C=O amide). ^{1}H NMR (DMSO d_6): δ = 9.84 (s, 1H, N$\underline{H}$NH$_2$), 8.03 (dd, 1H, J = 5.2, 0.4 Hz, H-6), 7.64 (s, 1H, NH$\underline{NH}_2$), 7.08 (s, 1H, H-3), 6.85 (dd, 1H, J = 5.2, 1.6 Hz, H-5). ^{13}C NMR (DMSO d_6): δ = 164.8 (NH$\underline{C}$O), 162.4 (C-2), 147.8 (C-6) 141.4 (C-4), 109.3 (C-5), 104.2 (C-3). HRESIMS: m/z 168.08797[M+H]$^+$(calcd. for C$_6$H$_9$N$_5$O + H 168.08799).

5.1.10. Crystallographic structure determination

Single crystal X-ray analysis of a suitable quality crystal of compound **7** was performed with a Bruker AXS-KAPPA APEX II diffractometer using graphite-monochromated Mo-Kα radiation (λ = 0.71069 Å) and operating at 50 kV and 30 mA. A single crystal was mounted loop with protective oil and placed under a flow of nitrogen gas at 150 K. All measurements were conducted on the Crystallography Laboratory at Instituto Superior Técnico. Cell parameters were retrieved using Bruker SMART (Bruker 2005) software and refined using Bruker SAINT (Bruker 2005) on all observed reflections. Absorption corrections were applied using SADABS (Bruker 2005). Structure solution and refinement were performed using direct methods with program SHELXS [114]and SHELXL-2013 [114], both included in the package of programs WINGX-Version 2013.3 [115]. Non-hydrogen atoms were refined anisotropically. A full matrix least-squares refinement on F^2 was used for the non-hydrogen atoms with anisotropic thermal parameters. All hydrogens were located in difference Fourier maps, and their coordinates were refined with isotropic displacement parameters. Table 9 summarizes the crystallographic data for compound **7**.

5.2. Microbiology studies

For the biological evaluation of the synthesized compounds, the *Mtb* strain H37Rv ATCC27294^T experimentally assayed in this work, pan-susceptible to the first-line antibiotics (MIC for INH = 0.04 μg/mL), was obtained from the culture collection of Grupo de Micobactérias, Unidade de Microbiologia Médica, Instituto de Higiene e Medicina Tropical (IHMT, UNL). For the resistance studies, we have used a *Mtb* clinical strain carrying only a *katG* S315T mutation (MIC for INH = 6 μg/mL) and the INH high-level resistant strain H37Rv ATCC27294^T (H37Rv$_{INH}$) that suffered a full deletion of the *katG* gene after prolonged exposure to INH (MIC for INH = 128 μg/mL) [28]. Cultures and minimum inhibitory concentration (MIC) determination were conducted using the BACTEC™ MGIT™ 960 system (BACTEC 960) and the Epicenter V5.53A software equipped with the TB eXIST module (Becton Dickinson Diagnostic Systems, Sparks, MD, USA), as previously described [26,28]. Stock-solutions of the compounds were prepared in methanol (Pronalab, Lisbon,

Table 9
Crystal data and structure refinement for Compound **7**.

Crystal data	
Empirical formula	C$_{15}$H$_{17}$N$_3$O$_5$
M_r	319.32
Temperature	150(2) K
Wavelength	0.71073 Å
Crystal system	Triclinic
Space group	P −1
Unit cell dimensions	a = 6.4597(7) Å
	b = 9.8034(11) Å
	c = 12.4981(14) Å
	α = 103.411(7) °.
	β = 97.331(6) °.
	γ = 94.030(7) °.
Volume	759.49(15) Å^3
Z	2
Calculated density	1.396 Mg/m^3
Absorption coefficient	0.106 mm^{-1}
F(000)	336
Crystal size	0.430 × 0.250 × 0.190 mm
θ range for data collection	3.197–28.368°.
Limiting indices	−8 ≤ h ≤ 8
	−13 ≤ k ≤ 13
	−16 ≤ l ≤ 16
Reflections collected/unique	13890/3751 [R(int) = 0.0354]
Completeness to θ = 25.242[deg.]	99.1%
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	3751/0/276
Goodness-of-fit on F^2	1.014
Final R indices [I > 2sigma(I)]	R$_1$ = 0.0387;
	wR$_2$ = 0.0956
R indices (all data)	R$_1$ = 0.0561;
	wR$_2$ = 0.1054
Extinction coefficient	n/a
Largest diff. peak and hole	0.315 and −0.215 e Å^{-3}

Portugal). The stock solutions were aliquoted and stored at −20 °C and the working solutions freshly prepared with sterile deionized water on the day of the experiment. The desired concentrations of the compounds were prepared with sequential dilutions in sterile deionized water and ranged from 0.01 to 12.0 μg/mL. MICs of compounds were evaluated in accordance to the procedures issued by the manufacturer of the BACTEC 960 system, revised by Springer et al. and slightly adapted by us [116,117]. Briefly, each compound-containing tube was inoculated with 0.8 mL of SIRE supplement, 0.1 mL of each compound in the appropriated concentration and 0.5 mL of the culture. For the preparation of the compound-free growth control tube (proportional control), the culture was diluted 1:100 with a sterile saline solution and 0.5 mL transferred into a new MGIT tube. Additionally, a second compound-free growth control, inoculated with 0.5 mL of the undiluted suspension of the strain, was prepared and served as absolute control for inoculums errors. The tubes were inserted in the BACTEC 960 system and the growth was continuously monitored with the EpiCenter TB eXIST software. The interpretation of the results was performed as follows: when the compound-free proportional control tube reached a growth unit (GU) value of 400 the test was considered complete. At that time, if the GU of the compound-containing tube to be compared with was ≥100, the strain was considered resistant; otherwise, if the GU of the compound-containing tube was <100, the strain was considered susceptible to that concentration. The MIC was defined as being the lowest concentration necessary to inhibit 99% of the bacterial population, corresponding to the compound-containing tube first recorded as susceptible [26,28]. Each MIC value determined with the BACTEC™ MGIT™ 960 system corresponds to three independent assays with a nil variance. MIC values reported for INH were also determined with this protocol for each strain.

5.3. Cytotoxicity evaluation by the MTT assay

The cytotoxicity of compounds with MIC values <0.5 was evaluated in Vero (African green monkey kidney) cells. Exponentially growing cells were seeded, in a 96-well flat-bottomed microplate, at a density of approximately 4×10^5 cells/mL and incubated for 48 h in a humidified atmosphere of 95% air/5% CO_2 at 37 ± 0.1 °C. Cells were then treated with the selected compounds, dissolved in DMSO, in concentrations ranging from 1 to 500 MIC, and incubated at the conditions described above. After 48 h, the MTT dye solution (100 μL, 0.5 mg/mL) was added to each well and cells were further incubated for 2 h at 37 ± 0.1 °C. The cellular medium was removed and DMSO was added to dissolve the formazan crystals. The absorbance was measured at 570 nm using a 96-well multiscanner autoreader (Tecan Sunrise). DMSO was used as control.

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejmech.2014.04.077.

References

[1] World Health Organization, Global Tuberculosis Report 2013, World Health Organization, Geneva, Switzerland, 2013. http://www.who.int/tb/publications/global_report/en/index.html (last accessed 06.02.14.).

[2] http://www.who.int/tb/publications/factsheets/en/ (last accessed 06.02.14.).

[3] G.B. Migliori, G. De Iaco, G. Besozzi, R. Centis, D.M. Cirillo, First tuberculosis cases in Italy resistant to all tested drugs, Eurosurveillance 12 (2007).

[4] A.A. Velayati, M.R. Masjedi, P. Farnia, P. Tabarsi, J. Ghanavi, A.H. ZiaZarifi, S.E. Hoffner, Emergence of new forms of totally drug-resistant tuberculosis bacilli super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran, Chest 136 (2009) 420–425.

[5] Z.F. Udwadia, R.A. Amale, K.K. Ajbani, C. Rodrigues, Totally drug-resistant tuberculosis in India, Clinical Infectious Diseases 54 (2012), 579–U156.

[6] WHO The Stop TB Strategy, 2010. http://www.who.int/tb/strategy/stop_tb_strategy/en/index.html (last accessed 11.04.13.).

[7] Global Plan to Stop TB 2011-2015, Stop TB Partnership and World Health Organization, World Health Organization, Geneva, 2010 (last accessed 11.04.13.), http://www.stoptb.org/global/plan/.

[8] A. Koul, E. Arnoult, N. Lounis, J. Guillemont, K. Andries, The challenge of new drug discovery for tuberculosis, Nature 469 (2011) 483–490.

[9] S. Munack, V. Leroux, K. Roderer, M. Ökvist, A. van Eerde, L.L. Gundersen, U. Krengel, P. Kast, When inhibitors do not inhibit: critical evaluation of rational drug design targeting chorismate mutase from *Mycobacterium tuberculosis*, Chemistry & Biodiversity 9 (2012) 2507–2527.

[10] C. Lienhardt, M. Raviglione, M. Spigelman, R. Hafner, E. Jaramillo, M. Hoelscher, A. Zumla, J. Gheuens, New drugs for the treatment of tuberculosis: needs, challenges, promise, and prospects for the future, The Journal of Infectious Diseases 205 (2012) S241–S249.

[11] H. Tomioka, Y. Tatano, K. Yasumoto, T. Shimizu, Recent advances in antituberculous drug. Development and novel drug target, Expert Review of Respiratory Medicine 2 (2008) 455–471.

[12] B. Villemagne, C. Crauste, M. Flipo, A.R. Baulard, B. Déprez, N. Willand, Tuberculosis: the drug development pipeline at a glance, European Journal of Medicinal Chemistry 51 (2012) 1–16.

[13] A. Zumla, P. Nahid, S.T. Cole, Advances in the development of new tuberculosis drugs and treatment regimens, Nature Reviews. Drug Discovery 12 (2013) 388–404.

[14] FDA, U.S. Foods and Drugs Administration, Press Release of Dec. 31, 2012.

[15] HIV-HCV-TB Pipeline Report, Drugs, Diagnostics, Vaccines & Preventive Technologies, Update February 2013. http://www.pipelinereport.org/browse/tb-treatments/bedaquiline-tmc207 (last accessed 13.02.13.).

[16] Briefing document, Anti-Infective Drugs Advisory Committee Meeting, Jansson Pharmaceutical Company of Johnson & Johnson, November 2013, http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/Anti-InfectiveDrugsAdvisoryCommittee/UCM329260.pdf (last accessed 11.04.13.).

[17] J. Bernstein, W.A. Lott, B.A. Steinberg, H.L. Yale, Chemotherapy of experimental tuberculosis .5. Isonicotinic acid hydrazide (Nydrazid) and related compounds, American Review of Tuberculosis 65 (1952) 357–364.

[18] C. Vilchèze, W.R. Jacobs Jr., The mechanism of isoniazid killing: clarity through the scope of genetics, Annual Review of Microbiology 61 (2007) 35–50.

[19] F. Bardou, C. Raynaud, C. Ramos, M.A. Lanéelle, G. Lanéelle, Mechanism of isoniazid uptake in *Mycobacterium tuberculosis*, Microbiology (Reading, U.K.) 144 (1998) 2539–2544.

[20] D.A. Mitchison, J.B. Selkon, The bactericidal activities of antituberculous drugs, American Review of Tuberculosis 74 (1956) 109–116.

[21] F.G. Winder, P.B. Collins, Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*, Journal of General Microbiology 63 (1970) 41–48.

[22] K. Johnsson, P.G. Schultz, Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*, Journal of the American Chemical Society 116 (1994) 7425–7426.

[23] B.F. Lei, C.J. Wei, S.C. Tu, Action mechanism of antitubercular isoniazid − activation *Mycobacterium tuberculosis* KatG, isolation, and characterization of InhA inhibitor, The Journal of Biological Chemistry 275 (2000) 2520–2526.

[24] R. Rawat, A. Whitty, P.J. Tonge, The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: adduct affinity and drug resistance, Proceedings of the National Academy of Sciences of the United States of America 100 (2003) 13881–13886.

[25] Y. Zhang, B. Heym, B. Allen, D. Young, S. Cole, The catalase peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*, Nature 358 (1992) 591–593.

[26] D. Machado, J. Perdigão, J. Ramos, I. Couto, I. Portugal, C. Ritter, E.C. Boettger, M. Viveiros, High level resistance to isoniazid and ethionamide in multidrug resistant *Mycobacterium tuberculosis* of the Lisboa family is associated with InhA double mutations, Journal of Antimicrobial Chemotherapy 68 (2013) 1728–1732.

[27] M. Viveiros, M. Martins, L. Rodrigues, D. Machado, I. Couto, J. Ainsa, L. Amaral, Inhibitors of mycobacterial efflux pumps as potential boosters for anti-tubercular drugs, Expert Review of Anti-infective Therapy 10 (2012) 983–998.

[28] D. Machado, I. Couto, J. Perdigão, L. Rodrigues, I. Portugal, P. Baptista, B. Veigas, L. Amaral, M. Viveiros, Contribution of efflux to the emergence of isoniazid and multidrug resistance in *Mycobacterium tuberculosis*, Plos One 7 (2012) e34538.

[29] R.Z. Pellicani, A. Stefanachi, M. Niso, A. Carotti, F. Leonetti, O. Nicolotti, R. Perrone, F. Berardi, S. Cellamare, N.A. Colabufo, Potent galloyl-based selective modulators targeting multidrug resistance associated protein 1 and P-glycoprotein, Journal of Medicinal Chemistry 55 (2012) 424–436.

[30] M.H. Hazbón, M. Brimacombe, M.B. del Valle, M. Cavatore, M.I. Guerrero, M. Varma-Basil, H. Billman-Jacobe, C. Lavender, J. Fyfe, L. García-García, C.I. León, M. Bose, F. Chaves, M. Murray, K.D. Eisenach, J. Sifuentes-Osornio, M.D. Cave, A.P. de León, D. Alland, Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobactetium tuberculosis*, Antimicrobial Agents and Chemotherapy 50 (2006) 2640–2649.

[31] H. van de Waterbeemd, E. Gifford, ADMET *in silico* modelling: towards prediction paradise? Nature Reviews. Drug Discovery 2 (2003) 192–204.

[32] L.C. Yee, Y.C. Wei, Current modeling methods used in QSAR/QSPR, in: M. Dehmer, K. Varmuza, D. Bonchev (Eds.), Statistical Modeling of Molecular Descriptors in QSAR/QSPR, Wiley-VCH Verlag GmbH & Co., KGaA Weinhein, 2012.

[33] R.V.C. Guido, G. Oliva, A.D. Andricopulo, Virtual screening and its integration with modern drug design technologies, Current Medicinal Chemistry 15 (2008) 37–46.

[34] K.K. Chohan, S.W. Paine, N.J. Waters, Advancements in predictive in silico models for ADME, Current Chemical Biology 2 (2008) 215–228.

[35] C. Ventura, F. Martins, Application of quantitative structure-activity relationships to the modeling of antitubercular compounds. 1. The hydrazide family, Journal of Medicinal Chemistry 51 (2008) 612–624.

[36] V. Kovalishyn, J. Aires-de-Sousa, C. Ventura, R.E. Leitão, F. Martins, QSAR modeling of antitubercular activity of diverse organic compounds, Chemometrics and Intelligent Laboratory Systems 107 (2011) 69–74.

[37] C. Ventura, D. Latino, F. Martins, Comparison of multiple linear regressions and neural networks based QSAR models for the design of new antitubercular compounds, European Journal of Medicinal Chemistry 70 (2013) 831–845.

[38] P. Gramatica, P. Pilutti, E. Papa, Validated QSAR prediction of OH tropospheric degradation of VOCs: splitting into training-test sets and consensus modeling, Journal of Chemical Information and Computer Science 44 (2004) 1794–1802.

[39] G. Klopman, D. Fercu, J. Jacob, Computer-aided study of the relationship between structure and antituberculosis activity of a series of isoniazid derivatives, Chemical Physics 204 (1996) 181–193.

[40] J.K. Seydel, K.J. Schaper, E. Wempe, H.P. Cordes, Mode of action and quantitative structure-activity correlations of tuberculostatic drugs of isonicotinic-acid hydrazide type, Journal of Medicinal Chemistry 19 (1976) 483–492.

[41] S. Mohamad, P. Ibrahim, A. Sadikun, Susceptibility of Mycobacterium tuberculosis to isoniazid and its derivative, 1-isonicotinyl-2-nonanoyl hydrazine: investigation at cellular level, Tuberculosis 84 (2004) 56–62.

[42] A. De Logu, V. Onnis, B. Saddi, C. Congiu, M.L. Schivo, M.T. Cocco, Activity of a new class of isonicotinoylhydrazones used alone and in combination with isoniazid, rifampicin, ethambutol, para-aminosalicylic acid and clofazimine against Mycobacterium tuberculosis, The Journal of Antimicrobial Chemotherapy 49 (2002) 275–282.

[43] D. Sriram, P. Yogeeswari, K. Madhu, Synthesis and in vitro and in vivo antimycobacterial activity of isonicotinoyl hydrazones, Bioorganic & Medicinal Chemistry Letters 15 (2005) 4502–4505.

[44] P. Prathipati, N.L. Ma, T.H. Keller, Global Bayesian models for the prioritization of antitubercular agents, Journal of Chemical Information and Modeling 48 (2008) 2362–2370.

[45] B. Bottari, R. Maccari, F. Monforte, R. Ottana, E. Rotondo, M.G. Vigorita, Isoniazid-related copper(II) and nickel(TI) complexes with antimycobacterial in vitro activity. Part 9, Bioorganic & Medicinal Chemistry Letters 10 (2000) 657–660.

[46] J. Bernstein, W.P. Jambor, W.A. Lott, F. Pansy, B.A. Steinberg, H.L. Yale, Chemotherapy of experimental tuberculosis .6. Derivatives of isoniazid, American Review of Tuberculosis 67 (1953) 354–365.

[47] N. Georgieva, V. Gadjeva, Isonicotinoylhydrazone analogs of isoniazid: relationship between superoxide scavenging and tuberculostatic activities, Biochemistry (Moscow) 67 (2002) 588–591.

[48] M.C.D. Lourenço, M.D. Ferreira, M.V.N. de Souza, M.A. Peralta, T.R.A. Vasconcelos, M.D.M.O. Henriques, Synthesis and anti-mycobacterial activity of (E)-N'-(monosubstituted-benzylidene)isonicotinohydrazide derivatives, European Journal of Medicinal Chemistry 43 (2008) 1344–1347.

[49] M.J. Hearn, M.H. Cynamon, M.F. Chen, R. Coppins, J. Davis, H.J.O. Kang, A. Noble, B. Tu-Sekine, M.S. Terrot, D. Trombino, M. Thai, E.R. Webster, R. Wilson, Preparation and antitubercular activities in vitro and in vivo of novel schiff bases of isoniazid, European Journal of Medicinal Chemistry 44 (2009) 4169–4178.

[50] C.H. Andrade, L.D. Salum, M.S. Castilho, K.F.M. Pasqualoto, E.I. Ferreira, A.D. Andricopulo, Fragment-Based, Classical Quantitative, Structure-activity relationships for a series of hydrazides as antituberculosis agents, Molecular Diversity 12 (2008) 47–59.

[51] https://www.collaborativedrug.com/pages/public_access (last accessed 13.02.13.).

[52] http://www.talete.mi.it/products/dragon_description.htm (last accessed 20.05.13.).

[53] http://www.molecular-networks.com/products/adrianacode (last accessed 12.04.13.).

[54] J.A. Platts, D. Butina, M.H. Abraham, A. Hersey, Estimation of molecular linear free energy relation descriptors using a group contribution approach, Journal of Chemical Information and Computer Science 39 (1999) 835–845.

[55] J.A. Platts, M.H. Abraham, D. Butina, A. Hersey, Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients, Journal of Chemical Information and Computer Science 40 (2000) 71–80.

[56] M.H. Abraham, A. Ibrahim, A.M. Zissimos, Y.H. Zhao, J. Comer, D.P. Reynolds, Application of hydrogen bonding calculations in property based drug design, Drug Discovery Today 7 (2002) 1056–1063.

[57] Molecular Modeling Pro Plus, version 6.2.5. www.chemistry-software.com.

[58] ChemDraw Ultra, Version 11.0.1, ©, CambridgeSoft, 1986–2007.

[59] L. Breiman, Random forests, Machine Learning 45 (2001) 5–32.

[60] R Development Core Team, R: a Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2004, ISBN 3-900051-07-0. URL, http://www.R-project.org (last accessed 20.05.13.). Fortran Original by Leo Breiman, Adele Cutler, R port by Andy Liaw and Matthew Wiener. (2004).

[61] I.V. Tetko, Neural network studies. 4. Introduction to associative neural networks, Journal of Chemical Information and Computer Science 42 (2002) 717–728.

[62] Organization for Economic Co-operation and Development (OECD), Environment Health and Safety Publications Series on Testing and Assessment, 69, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models, 2007. http://www.oecd.org/fr/env/ess/risques/guidancedocumentsandreportsrelatedtoqsars.htm (last accessed 11.04.13.).

[63] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, QSAR & Combinatorial Science 22 (2003) 69–77.

[64] T.I. Netzeva, A.P. Worth, T. Aldenberg, R. Benigni, M.T.D. Cronin, P. Gramatica, J.S. Jaworska, S. Kahn, G. Klopman, C.A. Marchant, G. Myatt, N. Nikolova-Jeliazkova, G.Y. Patlewicz, R. Perkins, D.W. Roberts, T.W. Schultz, D.T. Stanton, J.J.M. van de Sandt, W.D. Tong, G. Veith, C.H. Yang, Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships – the report and recommendations of ECVAM workshop 52, Atla-Altern, Laboratory Animals 33 (2005) 155–173.

[65] A. Gissi, D. Gadaleta, M. Floris, S. Olla, A. Carotti, E. Novellino, E. Benfenati, O. Nicolotti, An alternative QSAR-based approach for predicting the bioconcentration factor for regulatory purposes, Altex-Alternatives to Animal Experimentation 31 (2014) 23–36.

[66] J. Jaworska, N. Nikolova-Jeliazkova, T. Aldenberg, QSAR applicability domain estimation by projection of the training set in descriptor space: a review, Altex-Alternatives to Animal Experimentation 33 (2005) 445–459.

[67] A. Tropsha, Best practices for QSAR model development, validation, and exploitation, Molecular Informatics 29 (2010) 476–488.

[68] P. Gramatica, Principles of QSAR models validation: internal and external, QSAR & Combinatorial Science 26 (2007) 694–701.

[69] D. Livingstone, A Practical Guide to Scientific Data Analysis, Wiley & Sons Ltd, Chichester, 2009.

[70] C. Kim, Y. Lee, B.U. Park, Cook's distance in local polynomial regression, Statistics & Probability Letters 54 (2001) 33–40.

[71] J.A. Díaz-García, G. González-Farías, A note on the Cook's distance, Journal of Statistical Planning and Inference 120 (2004) 119–136.

[72] A. Golbraikh, A. Tropsha, Beware of q2!, Journal of Molecular Graphics and Modelling 20 (2002) 269–276.

[73] P.P. Roy, S. Paul, I. Mitra, K. Roy, On two novel parameters for validation of predictive QSAR models, Molecules 14 (2009) 1660–1701.

[74] N. Chirico, P. Gramatica, Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient, Journal of Chemical Information and Modeling 51 (2011) 2320–2335.

[75] N. Chirico, P. Gramatica, Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection, Journal of Chemical Information and Modeling 52 (2012) 2044–2058.

[76] K. Roy, I. Mitra, S. Kar, P.K. Ojha, R.N. Das, H. Kabir, Comparative studies on some metrics for external validation of QSPR models, Journal of Chemical Information and Modeling 52 (2012) 396–408.

[77] P.K. Ojha, I. Mitra, R.N. Das, K. Roy, Further exploring r^2m metrics for validation of QSPR models, Chemometrics and Intelligent Laboratory Systems 107 (2011) 194–205.

[78] I. Mitra, A. Saha, K. Roy, Exploring quantitative structure-activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants, Molecular Simulation 36 (2010) 1067–1079.

[79] L. Eriksson, J. Jaworska, A.P. Worth, M.T.D. Cronin, R.M. McDowell, P. Gramatica, Methods for reliability and uncertainty assessment and for applicability evaluation of classification- and regression-based QSARs, Environmental Health Perspectives 111 (2003) 1361–1375.

[80] N. Minovski, S. Zuperl, V. Drgan, M. Novic, Assessment of applicability domain for multivariate counter-propagation artificial neural network predictive models by minimum Euclidean distance space analysis: a case study, Analytica Chimica Acta 759 (2013) 28–42.

[81] G. Palla, G. Predieri, P. Domiano, C. Vignali, W. Turner, Conformational behavior and E/Z isomerization of N-acyl and N-aroylhydrazones, Tetrahedron 42 (1986) 3649–3654.

[82] E. Wyrzykiewicz, D. Prukala, Isomeric N-substituted hydrazones of 2-, 3- and 4-pyridinecarboxaldehydes, Journal of Heterocyclic Chemistry 35 (1998) 381–387.

[83] E. Wyrzykiewicz, A. Blaszczyk, New isomeric N-substituted hydrazones of 2-, 3-and 4-pyridinecarboxaldehydes and methyl-3-pyridylketone, Journal of Heterocyclic Chemistry 37 (2000) 975–981.

[84] A.M. Municio, A. Ribera, Sintesis De 2-Hidracino-Isonicotinoilhidracida Y Su N-Oxido, Anales de la Real Sociedad Española de Física y Química B 59 (1963) 179–184.

[85] M.J. Hearn, Antimycobacterial Compounds and Method for Making the Same. USA Patent n° US 6,846,933 B1, 2005.

[86] A. Rieche, G. Hilgetag, C. Bischoff, H. Mucke, Die Tuberkulostatische Wirksamkeit Von Inh-Acyl-Verbindungen, Archiv der Pharmazie Bericht 295 (1962) 707–714.

[87] L.I. Giannola, G. Giammona, R. Alotta, Pro-drugs of isoniazid – synthesis and diffusion characteristics of acyl derivatives, Pharmazie 47 (1992) 423–425.

[88] C. Ràfols, E. Bosch, R. Ruiz, K.J. Box, M. Reis, C. Ventura, S. Santos, M.E. Araújo, F. Martins, Acidity and hydrophobicity of several new potential antitubercular drugs: isoniazid and benzimidazole derivatives, Journal of Chemical & Engineering Data 57 (2012) 330–338.

[89] Y. Benadie, M. Deysel, D.G.R. Siko, V.V. Roberts, S. Van Wyngaardt, S.T. Thanyani, G. Sekanka, A.M.C. Ten Bokum, L.A. Collett, J. Grooten, M.S. Baird, J.A. Verschoor, Cholesteroid nature of free mycolic acids from M. tuberculosis, Chemistry and Physics of Lipids 152 (2008) 95–103.

[90] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, SIGKDD Explorations 11 (2009) 10–18.

[91] A. Verloop, J. Tipker, Pharmaco chemistry library, in: D. Hadzi, B. Jorman-Blazic (Eds.), QSAR in Drug Design and Toxicology, Elsevier Science Publishers BV, Amsterdam, 1987, pp. 97–125.

[92] R.P. Verma, A. Kurup, S.B. Mekapati, C. Hansch, Chemical-biological interactions in human, Bioorganic & Medicinal Chemistry 13 (2005) 933–948.

[93] C.E. Barry, R.A. Slayden, A.E. Sampson, R.E. Lee, Use of genomics and combinatorial chemistry in the development of new antimycobacterial drugs, Biochemical Pharmacology 59 (2000) 221–231.

[94] T. Scior, S.J. Garces-Eisele, Isoniazid is not a lead compound for its pyridyl ring derivatives, isonicotinoyl amides, hydrazides, and hydrazones: a critical review, Current Medicinal Chemistry 13 (2006) 2205–2219.

[95] A. Punkvang, P. Saparpakorn, S. Hannongbua, P. Wolschann, A. Beyer, P. Pungpo, Investigating the structural basis of arylamides to improve potency against *M. tuberculosis* strain through molecular dynamics simulations, European Journal of Medicinal Chemistry 45 (2010) 5585–5593.

[96] http://www.organic-chemistry.org/prog/peo/.

[97] A.M. Matos, M. Reis, V. Miranda, L. Santos, S. Santos, F. Martins, C. Ventura, M.S. Santos, Lipophilicity of some active *Mycobacterium tuberculosis* isoniazid derivatives: a comparative study between octanol-water (log Po/w) and micelle-water (log Kp) partition coefficients, in: Abstracts of the XXII International Symposium on Medicinal Chemistry, Berlin, Germany, 2012, ChemMedChem, Weinheim, Germany, 2012, p. 328. Abstract P521.

[98] N.L. Wengenack, F. Rusnak, Evidence for isoniazid-dependent free radical generation catalyzed by *Mycobacterium tuberculosis* KatG and the isoniazid-resistant mutant KatG(S315T), Biochemistry-Us 40 (2001) 8990–8996.

[99] R. Pierattelli, L. Banci, N.A.J. Eady, J. Bodiguel, J.N. Jones, P.C.E. Moody, E.L. Raven, B. Jamart-Gregoire, K.A. Brown, Enzyme-catalyzed mechanism of isoniazid activation in class I and class III peroxidases, The Journal of Biological Chemistry 279 (2004) 39000–39009.

[100] T. Deemagarn, X. Carpena, R. Singh, B. Wiseman, I. Fita, P.C. Loewen, Structural characterization of the Ser324Thr variant of the catalase-peroxidase (KatG) from *Burkholderia pseudomallei*, Journal of Molecular Biology 345 (2005) 21–28.

[101] X.B. Zhao, H. Yu, S.W. Yu, F. Wang, J.C. Sacchettini, R.S. Magliozzo, Hydrogen peroxide-mediated isoniazid activation catalyzed by *Mycobacterium tuberculosis* catalase-peroxidase (KatG) and its S315T mutant, Biochemistry-Us 45 (2006) 4131–4140.

[102] X.B. Zhao, S.W. Yu, R.S. Magliozzo, Characterization of the binding of isoniazid and analogues to *Mycobacterium tuberculosis* catalase-peroxidase, Biochemistry-Us 46 (2007) 3161–3170.

[103] N.L. Wengenack, S. Todorovic, L. Yu, F. Rusnak, Evidence for differential binding of isoniazid by *Mycobacterium tuberculosis* KatG and the isoniazid-resistant mutant KatG(S315T), Biochemistry-Us 37 (1998) 15825–15834.

[104] B. Heym, P.M. Alzari, N. Honore, S.T. Cole, Missense mutations in the catalase-peroxidase gene, Katg, are associated with isoniazid resistance in *Mycobacterium tuberculosis*, Molecular Microbiology 15 (1995) 235–245.

[105] J.M. Musser, V. Kapur, D.L. Williams, B.N. Kreiswirth, D. vanSoolingen, J.D.A. vanEmbden, Characterization of the catalase-peroxidase gene (katG) and InhA locus in isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance, The Journal of Infectious Diseases 173 (1996) 196–202.

[106] R.A. Slayden, C.E. Barry, The genetics and biochemistry of isoniazid resistance in *Mycobacterium tuberculosis*, Microbes and Infection 2 (2000) 659–669.

[107] L. Powers, A. Hillar, P.C. Loewen, Active site structure of the catalase-peroxidases from *Mycobacterium tuberculosis* and *Escherichia coli* by extended X-ray absorption fine structure analysis, Biochimica et Biophysica Acta 1546 (2001) 44–54.

[108] Y.G. Jin, S.F. Chen, R. Xin, Y.S. Zhou, Monolayers of the lipid derivatives of isoniazid at the air/water interface and the formation of self-assembled nanostructures in water, Colloid Surface B 64 (2008) 229–235.

[109] S.W. Yu, S. Girotto, C. Lee, R.S. Magliozzo, Reduced affinity for isoniazid in the S315T mutant of *Mycobacterium tuberculosis* KatG is a key factor in antibiotic resistance, The Journal of Biological Chemistry 278 (2003) 14769–14775.

[110] K.E. Dooley, C.D. Mitnick, M.A. DeGroote, E. Obuku, V. Belitsky, C.D. Hamilton, M. Makhene, S. Shah, J.C.M. Brust, N. Durakovic, E. Nuermberger, Resist-Tb, old drugs, new purpose: retooling existing drugs for optimized treatment of resistant tuberculosis, Clinical Infectious Diseases 55 (2012) 572–581.

[111] I. Orme, J. Secrist, S. Anathan, C. Kwong, J. Maddry, R. Reynolds, A. Poffenberger, M. Michael, L. Miller, J. Krahenbuh, L. Adams, A. Biswas, S. Franzblau, D. Rouse, D. Winfield, J. Brooks, T.D.S. Progra, Search for new drugs for treatment of tuberculosis, Antimicrobial Agents and Chemotherapy 45 (2001) 1943–1946.

[112] F.R. Pavan, P.I.D. Maia, S.R.A. Leite, V.M. Deflon, A.A. Batista, D.N. Sato, S.G. Franzblau, C.Q.F. Leite, Thiosemicarbazones, semicarbazones, dithio-carbazates and hydrazide/hydrazones: anti-*Mycobacterium tuberculosis* activity and cytotoxicity, European Journal of Medicinal Chemistry 45 (2010) 1898–1905.

[113] W. Vonsassen, M. Castroparra, E. Musch, M. Eichelbaum, Determination of isoniazid, acetylisoniazid, acetylhydrazine and diacetylhydrazine in biological-fluids by high-performance liquid-chromatography, Journal of Chromatography 338 (1985) 113–122.

[114] G.M. Sheldrick, A short history of SHELX, Acta Crystallographica A 64 (2008) 112–122.

[115] L.J. Farrugia, WinGX suite for small-molecule single-crystal crystallography, Journal of Applied Crystallography 32 (1999) 837–838.

[116] S. Rusch-Gerdes, G.E. Pfyffer, M. Casal, M. Chadwick, S. Siddiqi, Multicenter laboratory validation of the BACTEC MGIT 960 technique for testing susceptibilities of *Mycobacterium tuberculosis* to classical second-line drugs and newer antimicrobials, Journal of Clinical Microbiology 44 (2006) 688–692.

[117] B. Springer, K. Lucke, R. Calligaris-Maibach, C. Ritter, E.C. Bottger, Quantitative drug susceptibility testing of *Mycobacterium tuberculosis* by use of MGIT 960 and EpiCenter instrumentation, Journal of Clinical Microbiology 47 (2009) 1773–1780.