Discovery of New Inhibitors of *Mycobacterium tuberculosis* InhA Enzyme Using Virtual Screening and a 3D-Pharmacophore-Based Approach

Ivani Pauli,[‡,+,§,†] Ricardo N. dos Santos,[○] Diana C. Rostirolla,[+] Leonardo K. Martinelli,[+,§]
Rodrigo G. Ducati,[+] Luís F. S. M. Timmers,[‡,+,§] Luiz A. Basso,[+,§] Diógenes S. Santos,[+,§]
Rafael V. C. Guido,[○] Adriano D. Andricopulo,*,[○] and Osmar Norberto de Souza*,[‡,+,§,•]

[‡]Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas - LABIO, Faculdade de Informática, PUCRS, Brazil

[+]Centro de Pesquisas em Biologia Molecular e Funcional - CPBMF, Instituto Nacional de Ciência e Tecnologia em Tuberculose - INCT-TB, PUCRS, Brazil

[§]Programa de Pós-Graduação em Biologia Celular e Molecular - PPGBCM, Brazil

[•]Programa de Pós-Graduação em Ciência da Computação - PPGCC, Brazil

[○]Laboratório de Química Medicinal e Computacional - LQMC, Instituto de Física de São Carlos — IFSC-USP, Brazil

ABSTRACT: *Mycobacterium tuberculosis* InhA (*Mt*InhA) is an attractive enzyme to drug discovery efforts due to its validation as an effective biological target for tuberculosis therapy. In this work, two different virtual-ligand-screening approaches were applied in order to identify new InhA inhibitors' candidates from a library of ligands selected from the ZINC database. First, a 3-D pharmacophore model was built based on 36 available *Mt*InhA crystal structures. By combining structure-based and ligand-based information, four pharmacophoric points were designed to select molecules able to satisfy the binding features of *Mt*InhA substrate-binding cavity. The second approach consisted of using four well established docking programs, with different search algorithms, to compare the binding mode and score of the selected molecules from the aforementioned library. After detailed analyses of the results, six ligands were selected for *in vitro* analysis. Three of these molecules presented a satisfactory inhibitory activity with IC_{50} values ranging from 24 ($\pm$2) μM to 83 ($\pm$5) μM. The best compound presented an uncompetitive inhibition mode to NADH and 2-*trans*-dodecenoyl-CoA substrates, with K_i values of 24 ($\pm$3) μM and 20 ($\pm$2) μM, respectively. These molecules were not yet described as antituberculars or as InhA inhibitors, making its novelty interesting to start efforts on ligand optimization in order to identify new effective drugs against tuberculosis having InhA as a target. More studies are underway to dissect the discovered uncompetitive inhibitor interactions with *Mt*InhA.
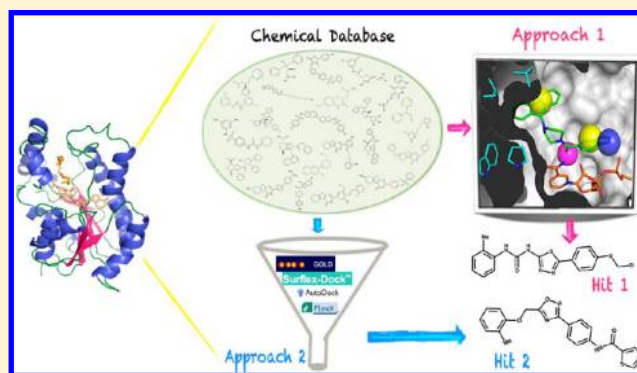
■ INTRODUCTION

Tuberculosis (TB) is more prevalent in the world today than at any other time in human history.[1] The resurgence of TB as a public health threat was mainly due to the arising of multi (MDR-TB), extensively (XDR-TB),[2] and recently, totally (TDR-TB)[3] drug resistant *Mycobacterium tuberculosis* strains. TB coinfections in HIV/AIDS patients, also contributed significantly to increase this problem.[4,5] Besides, *M. tuberculosis*, the TB etiological agent, is known to use diverse strategies to survive in a variety of host lesions and to evade immune surveillance.[1] In this context the urgency to pursue novel alternative treatments to TB infected patients is clear. To achieve global control of this epidemic, there is a need for new TB drugs, which can (1) shorten treatment duration; (2) target drug resistant strains; (3) simplify treatment by reducing the daily pill burden; (4) lower dosing frequency; and (5) be coadministered with HIV medications.[1]

Among the most attractive molecular targets to the design of novel antibacterial agents are the Fatty Acid Synthase (FAS) pathway enzymes.[6,7] The *Mycobacterium tuberculosis* InhA (*Mt*InhA) or 2-*trans*-enoyl-ACP (CoA) reductase (E.C.1.3.1.9), the fourth enzyme of the type II fatty acid synthase system (FAS II), is one of the key enzymes involved in the elongation cycle of fatty acids in *M. tuberculosis*. Its biological role includes the preferentially reduction of long chain enoyl thioester substrates (*e.g.*, containing 16 or more carbon atoms) yielding the long carbon chain of the meromycolate branch of mycolic acids (C_{40-60}), α-branched fatty acids, the hallmark of mycobacteria.[8] Previously, it has been shown that InhA is essential to the mycolic acid biosynthesis in *Mycobacterium*.[9] Its inactivation induces the

accumulation of saturated fatty acids, leading to cell wall alterations, lysis, and, consequently, to the pathogen dead. Therefore, owing to its biological importance as well as structural differences between *M. tuberculosis* and the human homologue, the InhA enzyme has been investigated as an attractive target for drug design.[6,10] In the literature there are many studies describing efforts in finding InhA inhibitors, as very well documented in a review published by Lu and co-workers,[11] and other recently published works.[12−14]

In the present work our aim is to identify new chemotypes that could serve as a starting point for the design of chemically novel and patentable drug candidates in drug discovery programs using very well established methodologies.[15] To achieve this goal, we describe here the use of two distinct *in silico* approaches: i) the development of a 3-D pharmacophore model for ligands of *Mt*InhA followed by a pharmacophore-based virtual screening based on the 3D model and ii) a virtual screening using a combination of four docking programs that employ distinct search algorithms. The use of these methodologies resulted in the identification of compounds with *in vitro* inhibitory activity against *Mt*InhA.

■ MATERIALS AND METHODS

Methodology Validation. Before starting the two approaches, we tested their capacity to correctly predict the pose and rank the ligands from the available crystallographic structures. We have used as the receptor, the crystallographic structure 1P44.[16] The ligand list included, besides the ligands from the crystallographic structures, 100 random molecules from ZINC. The Unity Chemical Information software[14] was able to retrieve the crystallographic molecules used to generate the pharmacophore model, and the four docking programs reproduced the crystallographic poses (the higher RMSD value was 0.5 when comparing with the respective crystallographic structures).

Molecule Library Selection. Based on the structural and physicochemical analysis of *Mt*InhA, a preselection of compounds to be screened was carried out from the ZINC Database.[17] Molecular 2D filters identified molecules with favorable features to interact with InhA binding cavity: 4 < LogP < 7; rotatable bonds < 6; hydrogen bond acceptors < 8; Polar Surface Area < 40 Å^2; and 250 Da < molecular weight < 400 Da. The final set of compounds gathering those properties was comprised by 999,853 molecules and was used in both, first and second approaches, which are described below.

■ APPROACH: 1 STRUCTURE-BASED PHARMACOPHORE GENERATION

Computational Approach. The pharmacophore search was performed using the SYBYL 8.0 package (Tripos Inc., St. Louis, MO) running on Red Hat Enterprise Linux workstations. The 3-D structures of the inhibitors were generated using standard geometric parameters of the molecular modeling software package SYBYL 8.0. Each single optimized conformation of a molecule in the data set was energetically minimized employing the Tripos force field[18] and the Powell conjugate gradient algorithm[19] with a convergence criterion of 0.005 kcal/mol·Å. Partial atomic charges were calculated by the Gasteiger−Hückel method.[20] The analyses, calculations, and visualizations were performed using the programs SYBYL 8.0 and PyMol.[21]

Pharmacophore Model Generation. Currently, 36 *Mt*InhA crystal structures are available at the Protein Data Bank (PDB).[22,23] Among them, we found InhA wild type and its mutants related to drug resistance, in apo form (without a ligand) and in complex with NADH (coenzyme), substrate analog, and several ligands. The 36 crystal structures were superposed by least-squares fit of the Cα atoms with TopMatch[24,25] available at (http://topmatch.services.came.sbg.ac.at/). The binding cavity volumes were calculated with CASTp[26] (http://sts.bioengr.uic.edu/castp/). Hydrogen bond and hydrophobic contacts were also analyzed with LIGPLOT[27] to obtain a more robust profile of the interaction pattern between *Mt*InhA and its crystallographic ligands. The structural analysis of *Mt*InhA in complex with the ligands bound to the substrate-binding cavity indicated several key amino acid residues that might be involved in the molecular recognition process. On the basis of that, a four-point 3-D pharmacophore model was developed.

3-D Pharmacophore Generation and Database Searching. The molecule library 3-D structures were transferred into an UNITY database in SYBYL 8.0. The 3-D search with UNITY was based on the generated pharmacophore model. In order to employ the pharmacophore model as 3-D constraints, we used the crystal spatial coordinates of the ligand 5-[4-(9H-fluoren-9-yl)piperazin-1-yl]carbonyl-1H-indole (Genz-10850) bound to *Mt*InhA (PDB ID: 1P44). The donor and acceptor atoms were defined by connecting the donor and the acceptor atoms of the ligand via the partial match utility. The hydrophobic feature was partially characterized based on two rings present in the ligand structure. The spheres diameters were adjusted in such a way that the test sample, composed of known *Mt*InhA crystallographic ligands, could be retrieved. To screen the molecules library we performed a UNITY 3-D flexible search.

Molecular Modeling. UNITY algorithm for flexible searches assigns high score to the molecules that match the pharmacophore points, because of that, the optimized geometry of the compounds under evaluation may be affected. Therefore, the molecular structures of the selected set of compounds were minimized to avoid false positive results. To energetically minimize the molecules we used the same protocol as described in the Computational Approach section. Subsequently, the selected ligand candidates were docked into *Mt*InhA substrate-binding cavity. The GOLD suite of programs[28] was used to perform the molecular docking simulations. The *Mt*InhA crystal structure used as the receptor (PDB ID: 1P44) was chosen among all the 36 available structures, because of the best relation between binding cavity volume (3,007.8 Å^3) and protein resolution (2.7 Å). For the calculations, hydrogen atoms were added to the protein molecule in standard geometry, using the Biopolymer module as implemented in SYBYL 8.0. The active site was defined incorporating all amino acid residues within a radius sphere of 8.0 Å centered on the crystallographic bound ligand (Genz-10850). The NADH coenzyme was treated as part of the protein in all docking simulations. Molecular modeling studies were carried out using GOLD default parameters. The GoldScore fitness function was employed to select the representative pose for each compound. The solutions were visually checked to find if the poses obtained by docking still matched the original pharmacophore hypothesis. The visual inspection is a common practice important to avoid testing false positives due to assumptions and shortcomings in docking methods and scoring functions.[29] In the visual inspection we critically assessed the suggested binding conformation, the mutual surface complementarity of ligand and protein, and the

possible presence of unfilled space along the protein ligand interface. Only those molecules that matched all the four pharmacophore points after the docking simulations were selected to the next step (*in silico* toxicity prediction).

APPROACH 2: EXHAUSTIVE DOCKING SIMULATIONS USING DIFFERENT ALGORITHMS

Docking Algorithms. Parallel to the structure-based pharmacophore approach, molecular docking simulations were performed using four different and well established docking programs: GOLD,[28] AutoDock,[30,31] FlexX,[32] and Surflex-Dock.[33]

GOLD,[28] the Genetic Optimization for Ligand Docking, is a program that employs a genetic algorithm for docking flexible ligands into protein binding sites, using the Genetic Algorithm. The program gives some options in relation to the scoring function, and among them we elected GoldScore to be used in our simulations. AutoDock[30,31] uses a semiempirical free energy force field to evaluate conformations during docking simulations and provides several methods to do the conformational search. The Lamarckian Genetic Algorithm, a modification of the traditional Genetic Algorithm, was our choice for this study. FlexX[32] has implemented an Incremental Construction Algorithm where the base fragment (the ligand core) is automatically selected and placed into the active site using an algorithmic approach based on a pattern recognition technique called pose clustering. After a good set of placements has been obtained for the base, the remaining portions of the ligand are divided into small fragments and incrementally "grown" onto the base alternatives. The Surflex-Dock algorithm[33] employs the empirical Hammerhead[34] scoring function and uses an idealized active site ligand, or protomol, to generate ligand poses by incremental construction and a crossover procedure that combines pieces from distinct poses. Protomols were computed using the position of the cognate ligand to define the binding site.

Protein Receptor. The protein receptor chosen for the simulations was *Mt*InhA with PDB access code 1P44, and NADH coenzyme was treated as part of the protein. In all docking simulations the receptor structure was kept rigid, while the ligands were treated as totally flexible.

Molecule Data Set. The same used for the pharmacophore based approach.

Docking Simulations. First, all the molecules in the initial data set were docked into *Mt*InhA using GOLD.[28] GOLD was chosen for this first step because it was the one, among the four docking programs, which achieved the best results in predicting the crystallographic poses of the ligands in the validation step.

The active site was defined incorporating all amino acid residues within a radius sphere of 8.0 Å centered on the 1P44 crystallographic bound ligand (Genz-10850). The 100 molecules top-scored by GOLD were selected to be submitted to the other three programs, AutoDock, FlexX, and Surflex-Dock, using the same parameters as in GOLD, being different only in the algorithms for searching and scoring the best ligand poses, which are characteristic of each program. The ligands that achieved a very similar conformation, in terms of functional groups positioning, by at least three of the four program algorithms, were selected to the next step (*in silico* toxicity prediction).

***In Silico* Predictive Toxicology.** Unacceptable toxicity is still a major bottleneck in the drug discovery process.[35] Aiming at addressing this problem, *in silico* predictive toxicology techniques are fast and cost efficient alternatives (or supplements) to bioassays for the identification of toxic effects at an early stage of product development. Therefore, as an additional *in silico* analysis for all the molecules selected both by the first approach as well as for the second one, we used the OSIRIS drug discovery informatics system[36] to predict toxicity features. The OSIRIS Property Explorer (available at: http://www.organic-chemistry. org/prog/peo/) is an integral part of Actelion's in house substance registration system. It calculates on-the-fly, while the user draws a molecule, various drug-relevant properties. The molecules selected from both approaches used in this work were submitted to OSIRIS, and all those that presented unfavorable features were removed from the final set of molecules to be tested *in vitro*.

Similarity Ensemble Approach (SEA). For the ligands that passed the toxicity analysis we did a search at the SEA online tool[37] (http://sea.bkslab.org/search/) to assess the relationships between our hits and molecules with described properties and targets.

***Mt*InhA Expression and Purification.** The enzyme was expressed and purified as previously described.[38,39]

***In Vitro Mt*InhA Enzyme Inhibition Measurements.** To evaluate the relative potency of each selected ligand, measurements of steady-state velocities in the presence of ligands were carried out using a UV-2550 UV/visible spectrophotometer (Shimadzu), monitoring the NADH oxidation at 340 nm $(\varepsilon_{\beta\text{-NADH}} = 6.22\ \text{M}^{-1}\ \text{cm}^{-1})$ due to reduction of 2-*trans*-dodecenoyl-CoA substrate. Experiments were performed at 25 °C, in 100 mM NaH_2PO_4 pH 7.5, and measurement of enzyme-catalyzed chemical reaction started with the addition of *Mt*InhA in a 3 μM concentration to the assay mixture. In short, different concentrations of the selected compounds (1−20 μM) were added to the reaction mixture, and the enzyme velocity was analyzed as the % of inhibition. As the negative control, the maximal rate for the enzyme reaction (100% of *Mt*InhA activity) was determined in the absence of inhibitor and in the presence of fixed nonsaturating concentrations of NADH (60 μM $\cong K_m$) and 2-*trans*-dodecenoyl-CoA (45 μM $\cong K_m$).[38,39] We have recently shown that a pentacyano(isoniazid)ferrateII complex (IQG607) inhibits the enzyme activity of both wild-type InhA and I21V mutant InhA identified in isoniazid-resistant clinical isolates of *M. tuberculosis*.[40] This compound is a slow-onset inhibitor of InhA with an overall inhibition constant value of 70 nM[41] and was used as the positive control of the assays.

To determine the mode of inhibition and overall inhibition constant (K_i), initial rates were measured as a function of NADH concentration at fixed nonsaturating 2-*trans*-dodecenoyl-CoA concentration (45 μM) and fixed-varied inhibitor concentrations (0, 1, 10, and 20 μM). Inhibition studies were also carried out in the presence of fixed nonsaturating concentration of NADH (60 μM) and fixed-varied inhibitor concentrations (0, 6, 10, and 20 μM) and 2-*trans*-dodecenoyl-CoA as the variable substrate. The K_i values for both substrates were calculated fitting the data to an equation describing uncompetitive inhibition (eq 1), in which $[I]$ is the inhibitor concentration, $[S]$ is the substrate concentration, K_m is the Michaelis−Menten constant, V_{max} is the maximal velocity, and K_i is the overall inhibition constant.

$$v = \frac{V_{max}[S]}{[S]\left(1 + \frac{[I]}{K_i}\right) + K_m} \tag{1}$$

RESULTS AND DISCUSSION

Approaches 1 and 2. Two distinct methodologies were used in order to identify the most promising inhibitors for the
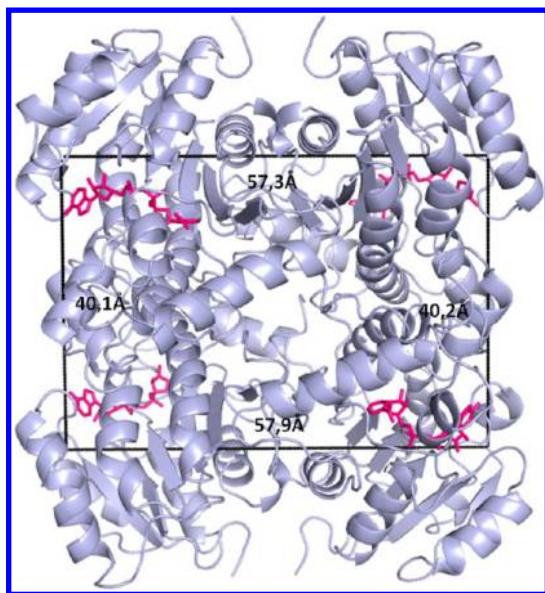
Figure 1. Distance between monomers' active sites. Ribbons representation of *Mt*InhA main chain. Image generated with PyMol.[21] The distances between adjacent monomers were measured with reference to the PA atom of NADH (magenta).

enzyme InhA from *M. tuberculosis*. As a first approach, a 3-D pharmacophore model was built based on the features of all the

36 *Mt*InhA crystal structures available at PDB and on information from the literature. In the second approach we used four molecular docking programs (described in the Materials and Methods section), each one with its particular algorithm to search and score the best pose for each ligand.

Protein Target Features. InhA is encoded by the *inh*A gene and is composed of 269 amino acids with a molecular mass of ∼29 kDa. Size exclusion chromatography analysis demonstrates that InhA is a homotetramer in solution, and this is the biologically active structure.[38] The active sites of the four monomers, far apart more than 40.0 Å considering neighboring subunits, are facing opposite sides in the InhA quaternary structure (Figure 1), justifying the use of the monomer in computational simulations.[38]

Overall, the *Mt*InhA monomer structure (Figure 2) seems like a chair and is composed of a seven strands β-sheet and eight α-helices.[42] Each monomer has two binding sites, one for the NADH coenzyme and another for the substrate (2-*trans*-dodecenoyl-CoA). Several α-helices and the β-sheet of the *Rossmann fold* extend over the NADH binding site, creating a crevice where the substrate binds to. Three loops, the substrate-binding loop, loop A, and loop B, are important components of the substrate-binding site.

Modeling studies aim to identify competitive inhibitors with the substrate, avoiding molecules that bind to the NADH cavity, since NADH is also a coenzyme for many other proteins and has important roles in other metabolic pathways as well. In this way, a
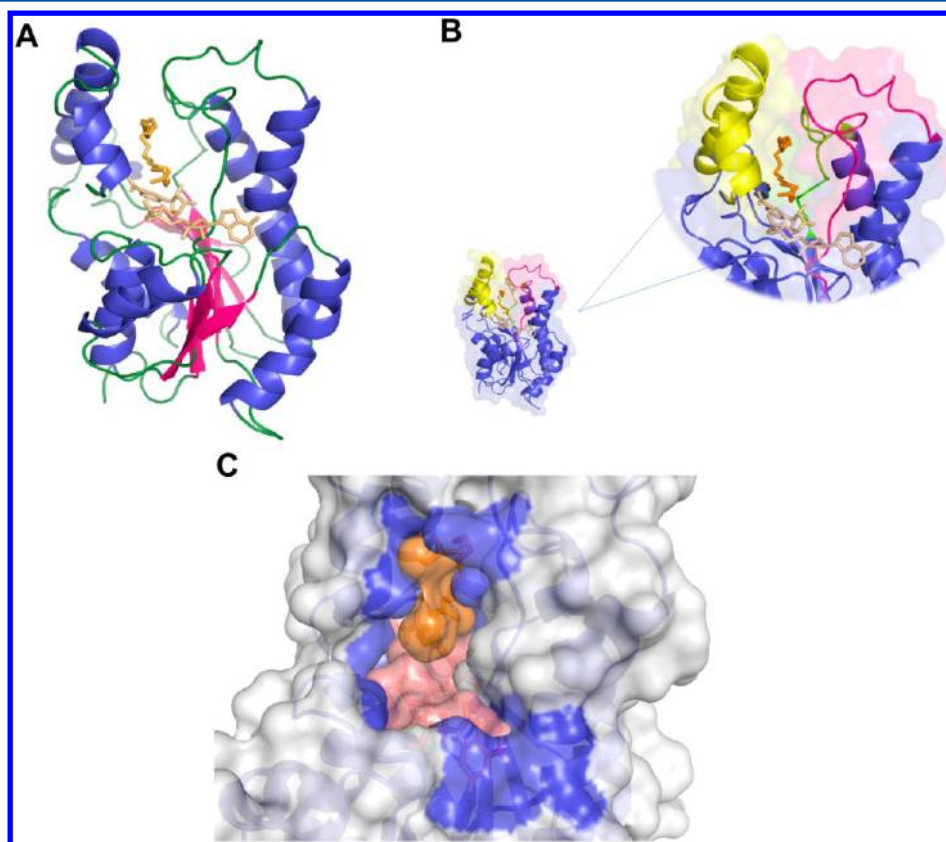

Figure 2. Main structural features of *Mt*InhA (PDB ID: 1BVR). (A) Monomer structure of the InhA enzyme in complex with NADH and the C16 substrate analog. Ribbons representation of subunit C: α-helices in blue, β-sheet in magenta (characterizing the Rossmann fold topology), and loops in green. NADH (salmon) is positioned over the Rossmann fold, and the substrate analog (orange) is found in its binding cavity, right above NADH. (B) InhA active site. Highlighted in yellow is the substrate-binding loop. Loops A and B are shown in magenta and green, respectively. (C) Molecular surface detail of the substrate-binding cavity. In blue are the binding site residues that interact with NADH and C16. The images were generated with PyMol.[41]

D

dx.doi.org/10.1021/ci400202t | *J. Chem. Inf. Model.* XXXX, XXX, XXX−XXX

molecule that binds to the substrate-binding cavity would be much more selective for InhA. Therefore, in the present study, only the substrate-binding cavity was explored as the target site to identify new inhibitors.

■ APPROACH 1

Construction of the Pharmacophore Model. One of the most modern concepts defines pharmacophore as "an ensemble
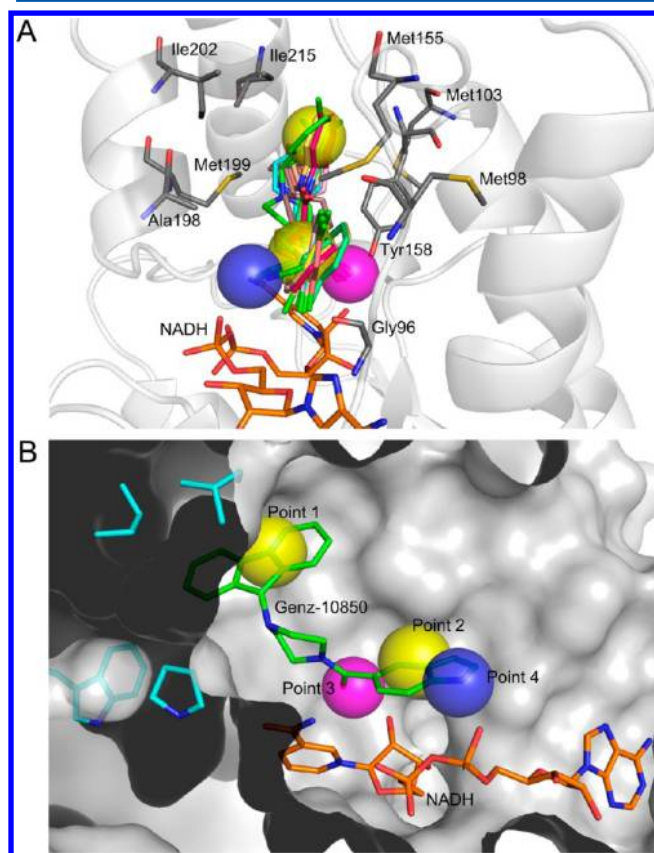

Figure 3. (A) Superposition of all ligands present in the InhA crystal structures' substrate-binding cavity. Highlighted are the important amino acids that interact with different ligands. (B) 3-D pharmacophore model obtained from the receptor structures and the known ligands features. Yellow = regions favorable to hydrophobic groups; Magenta = favorable region to hydrogen bond donor or acceptor groups; Blue = region propitious to hydrogen bond donor groups. The images were generated by PyMol.[41]

of steric and electronic features that is necessary to ensure optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response".[43] To achieve success using a structure-based pharmacophore approach, it is essential to have available a great amount of information about the protein target of interest, since the molecular recognition phenomenon relies on the properties and features of the binding pocket, which are determined by the amino acids comprising the binding cavity. Therefore, our goal was to design a robust 3-D pharmacophore-based virtual screening approach in order to identify molecules able to inhibit *Mt*InhA. In this way, a detailed structural analysis was conducted on the 36 *Mt*InhA crystal structures to generate a receptor-based pharmacophore model.

The superposition of the Cα atoms of the 36 monomers (in an all against all manner) revealed that the overall structure topology root-mean-square deviation (RMSD) among different

complexes ranges from 0.1 Å to 1.4 Å, indicating that the overall structure topology of the different complexes is significantly different. This difference is also evident by analyzing the binding cavity volumes, ranging from 1,597.3 Å^3 to 3,046.7 Å^3 for the whole cavity (NADH + substrate-binding cavity). We also demonstrate that the major contribution to flexibility comes from the substrate-binding cavity, with volumes ranging from 327.6 Å^3 to 2,109.8 Å^3, while the NADH binding site volume is more conserved (ranging from 1,292.8 Å^3 to 1,525.7 Å^3).

No crystal water molecule was found to be conserved in the substrate-binding cavity, maybe due to its high hydrophobicity level (66% of all amino acids that established at least one contact with a ligand in the *Mt*InhA crystal structures have a hydrophobic side chain).

Among the protein residues making hydrophobic contacts with the ligands in the crystal structures, Gly96, Phe97, Met103, Phe149, Met155, Ile215, and the three residues from the substrate-binding loop described in the literature as important for slow binding inhibition (Ala198, Met199, and Ile202) are the ones that were more recurrent.[44] Therefore, we chose two pharmacophore points and placed them in regions favorable to hydrophobic interactions, one (point 1) nearby the substrate-binding loop and another (point 2) near loops A and B. The spheres diameters were adjusted in order to cover the interaction surface of these residues (Figure 3).

In relation to the hydrogen bond pattern into the substrate-binding cavity, the role of a particular residue, Tyr158, deserves to be highlighted. Parikh and collaborators[45] described Tyr158 as an electrophilic catalyst, stabilizing the transition state for hydride transfer by hydrogen bonding to the substrate carbonyl. In addition, Rozwarski et al.[46] reported a structure of a C16 fatty acid substrate analog bound to InhA that shows Tyr158 hydrogen bonded to the substrate carbonyl group and rotated from the position it occupies in the InhA-NADH binary complex. We also found in our analyses that all the crystallographic ligands in the InhA substrate-binding cavity are hydrogen bonded to Tyr158. Considering its importance, we placed a third pharmacophore point (point 3) next to the Tyr158 side chain (Figure 3). This same point is also nearby the O2D atom of the NADH nicotinamide ring, another important recurrent hydrogen bonding site to ligands in the substrate-binding cavity.

The fourth pharmacophore point (point 4) was placed in another region favorable to hydrogen bond interactions, close to the phosphate groups of the NADH molecule (Figure 3), specifically enclosing the favorable region for interaction with atoms O1A, O3, and O3D of the phosphate groups.

Structure-Based Virtual Screening. To highlight the spatial arrangement of chemical features that represent the essential interactions of small molecule ligands and *Mt*InhA, the 3-D pharmacophore model was translated into a flexible search query appropriate for a 3-D search with UNITY.[47] The tolerance spheres for the four pharmacophore points were placed by picking the appropriate atoms of the ligand Genz-10850 bound to the InhA active site (PDB ID: 1P44), and their diameters were manually adapted in order to cover the favorable areas indicated by the structural analysis.

A database containing 999,853 was screened, and all four pharmacophore points were required in the 3-D searches. From the initial data set, a total of 108,705 (10.87%) molecules were retrieved by satisfying the pharmacophore restrictions. These molecules were docked into the *Mt*InhA substrate-binding cavity using GOLD.[28] The coenzyme NADH was considered as part of the protein.
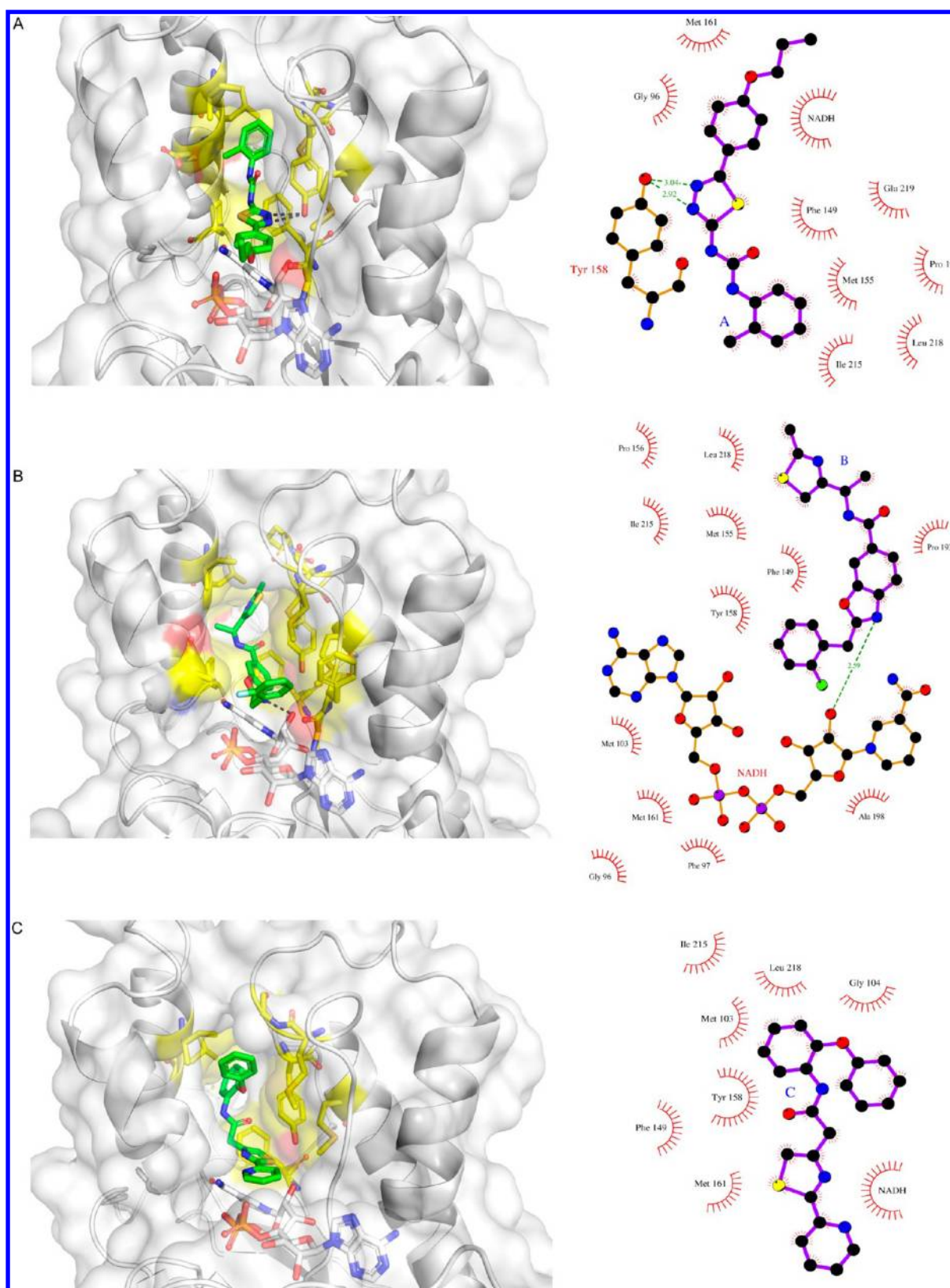
Figure 4. InhA-ligand interaction analyses for Approach 1. This figure shows the binding modes of the molecules selected using the first approach described in this work, in which a pharmacophore model was used to search new potential inhibitor molecules: (A) ZINC09137707, (B) ZINC12759934, (C) ZINC22559057. Ligands are represented in green, and the protein residues making hydrophobic contacts with them are shown in yellow. The equivalent LIGPLOT diagrams of the interactions are shown on the right column. Green dashed lines are hydrogen bonds. Single spoked arcs represent hydrophobic contacts.[20] Images were prepared with PyMol.[21]

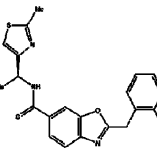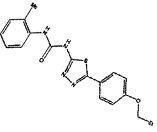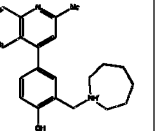**Table 1. Information about the Molecules Selected from Approach 1**
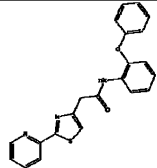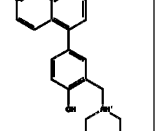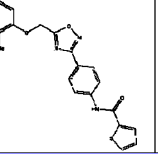
ZINC code	Structure	IUPAC Name	xLogP	HBD	HBA	Charge	Mwt	NRB
APPROACH 1								
ZINC 22559057		N-(2-phenoxyphenyl)-2-[2-(2-pyridyl)thiazol-4-yl] acetamide	4.57	1	5	0	387.464	6
ZINC 12759934		2-[(2-fluorophenyl)methyl]-N-[(1S)-1-(2-methylthiazol-4-yl)ethyl]-1,3-benzoxazole-6-carboxamide	4.06	1	5	0	395.459	5
ZINC 09137707		3-(o-tolyl)-1-[5-(4-propoxyphenyl)-1,3,4-thiadiazol-2-yl]urea	4.74	1	6	-1	367.454	7

Table 2. Information about the Molecules Selected from Approach 2

ZINC code	Structure	IUPAC Name	xLogP	HBD	HBA	Charge	Mwt	NRB
APPROACH 2								
ZINC 12242826		2-(azepan-1-ylmethyl)-4-(2-methyl-4-quinolyl)phenol	5.54	2	3	1	347.482	3
ZINC 12509636		4-(7-chloro-4-quinolyl)-2-(diethylaminomethyl)phenol	5.1	3	3	2	342.87	5
ZINC 02931014		N-(4-{5-[(2-methylphenoxy)methyl]-1,2,4-oxadiazol-3-yl}phenyl)thiophene-2-carboxamide	4.89	1	6	0	391.452	6

The predicted binding modes into the *Mt*InhA substrate-binding site were ranked based on their relative affinity using the GoldScore scoring function implemented in GOLD. The top 300 ranked molecules were selected for visual inspections to select the final set of compounds. The criteria used in the visual inspection included the following: (i) matching to the 4 points of the pharmacophore model after docking; (ii) overall matching to the hydrogen-bonding network, with emphasis on the formation of hydrogen bonds to Tyr158; (iii) overall matching of hydrophobic contacts to the regions assigned by the pharmacophore model; and (iv) complementarity between ligands and protein surfaces in terms of spatial occupancy. After this analysis the final subset of compounds consisted of 34 molecules.

■ APPROACH 2

Docking Programs. Molecular docking programs are commonly used to orient small molecules within a three-dimensional representation of the protein structure. There is a large number of docking programs currently available, with a wide diversity of algorithms to generate ligand conformations and scoring functions to rank the pose solutions obtained. Therewith our aim was to access whether the different algorithms implemented in each of the well established and widely used programs GOLD,[28] AutoDock,[30,31] FlexX,[32] and Surflex-Dock[33] could be able to generate the similar equivalent poses for a particular ligand, indicating that this may be the favorable mode of interaction between the protein and the ligand.

Docking Simulations. After docking the whole data set using GOLD, the 100 top ranked molecules were resubmitted to molecular docking simulations to AutoDock, FlexX, and Surflex, using the same parameters as used for the docking with GOLD. In this approach, the main difference relied on the implemented algorithms for searching and scoring the best ligand poses.

Molecule Conformation Analysis. The best pose obtained by each program for each of the 100 GOLD top-ranked
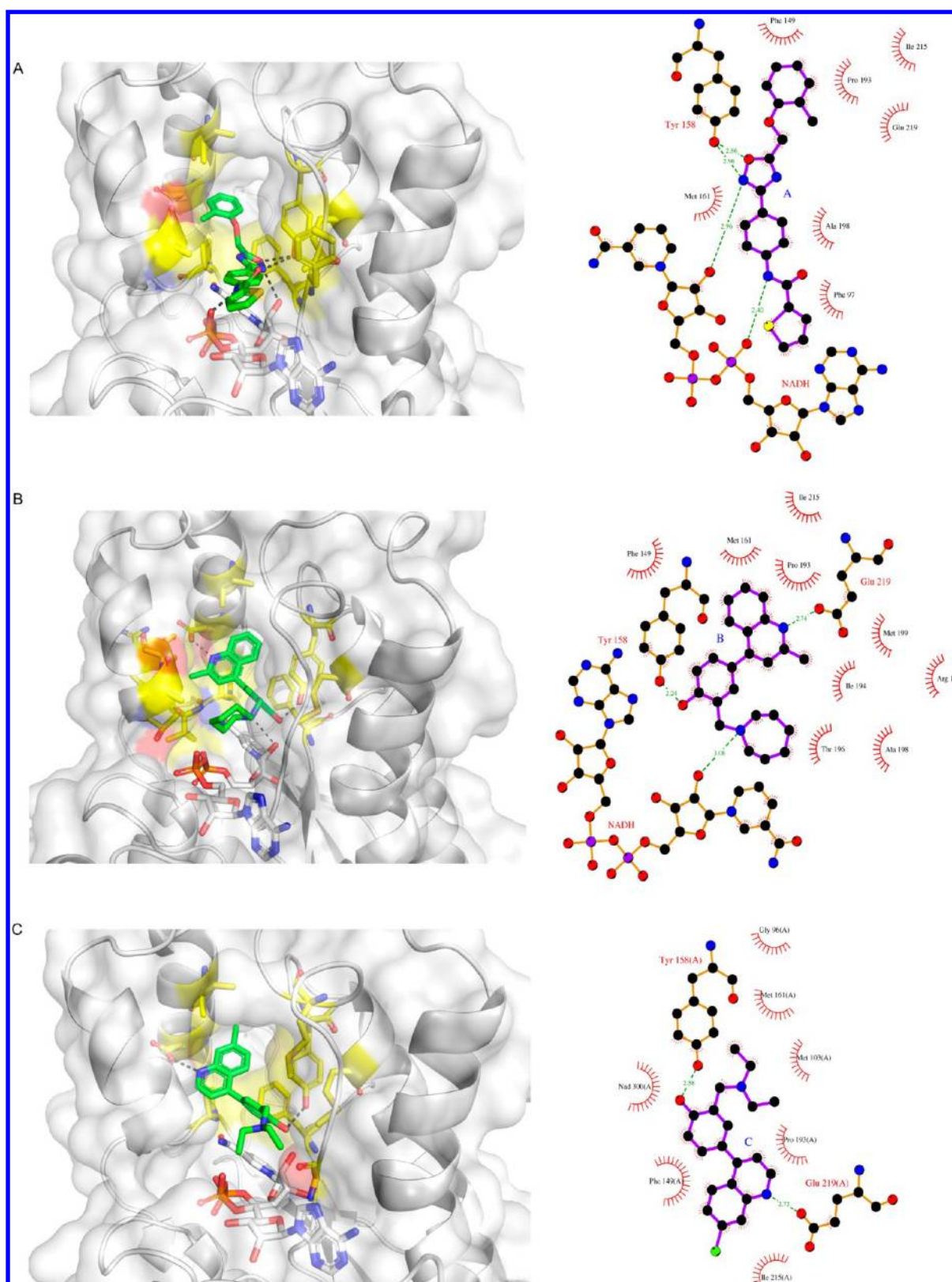
Figure 5. InhA-ligand interaction analyses for Approach 2. This figure shows the binding modes of the molecules selected using the second approach described in this work, in which four docking programs were used to identify the most promising ligands: (A) ZINC02931014, (B) ZINC12242826, (C) ZINC12509636. Ligands are represented in green, and the protein residues making hydrophobic contacts with them are shown in yellow. The equivalent LIGPLOT diagrams of the interactions are shown on the right column. Green dashed lines are hydrogen bonds. Single spoked arcs represent hydrophobic contacts.[20] Images were prepared with PyMol.[21]

molecules was compared in terms of conformational geometry and positioning in the substrate-binding cavity. We visually

analyzed the results obtained by the four docking programs and selected the molecules for which at least three of the four
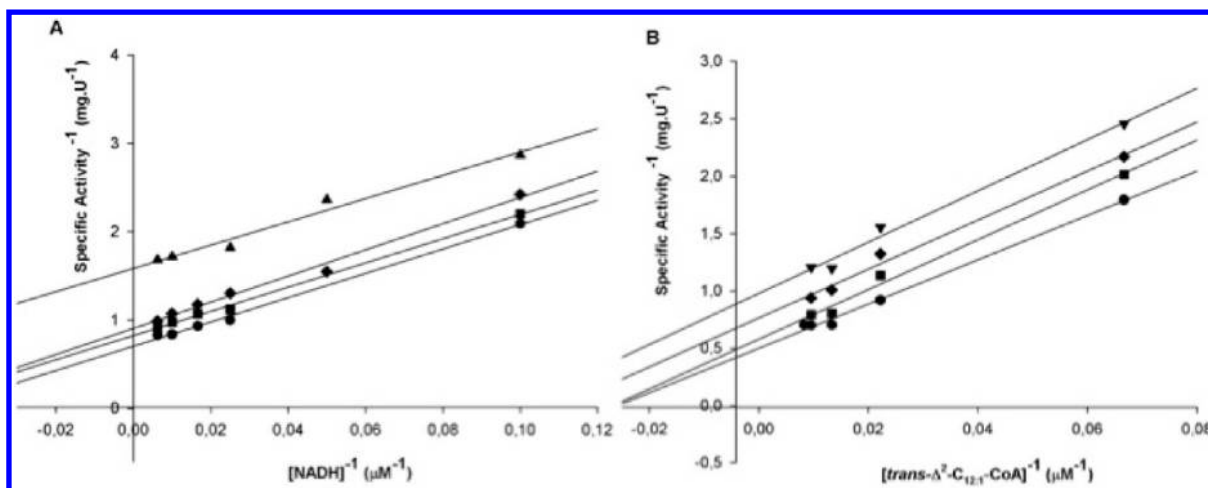
Figure 6. Steady-state kinetics analysis of *Mt*InhA-ZINC09137707. (A) Double-reciprocal plot of specific activity^{-1} (mg U^{-1}) versus [NADH]$^{-1}$ (μM^{-1}) at 0, 1, 10, and 20 μM ZINC09137707. (B) Double-reciprocal plot of specific activity^{-1} (mg U^{-1}) versus [2-*trans*-dodecenoyl-CoA]$^{-1}$ (μM^{-1}) at 0, 6, 10, and 20 μM ZINC09137707.
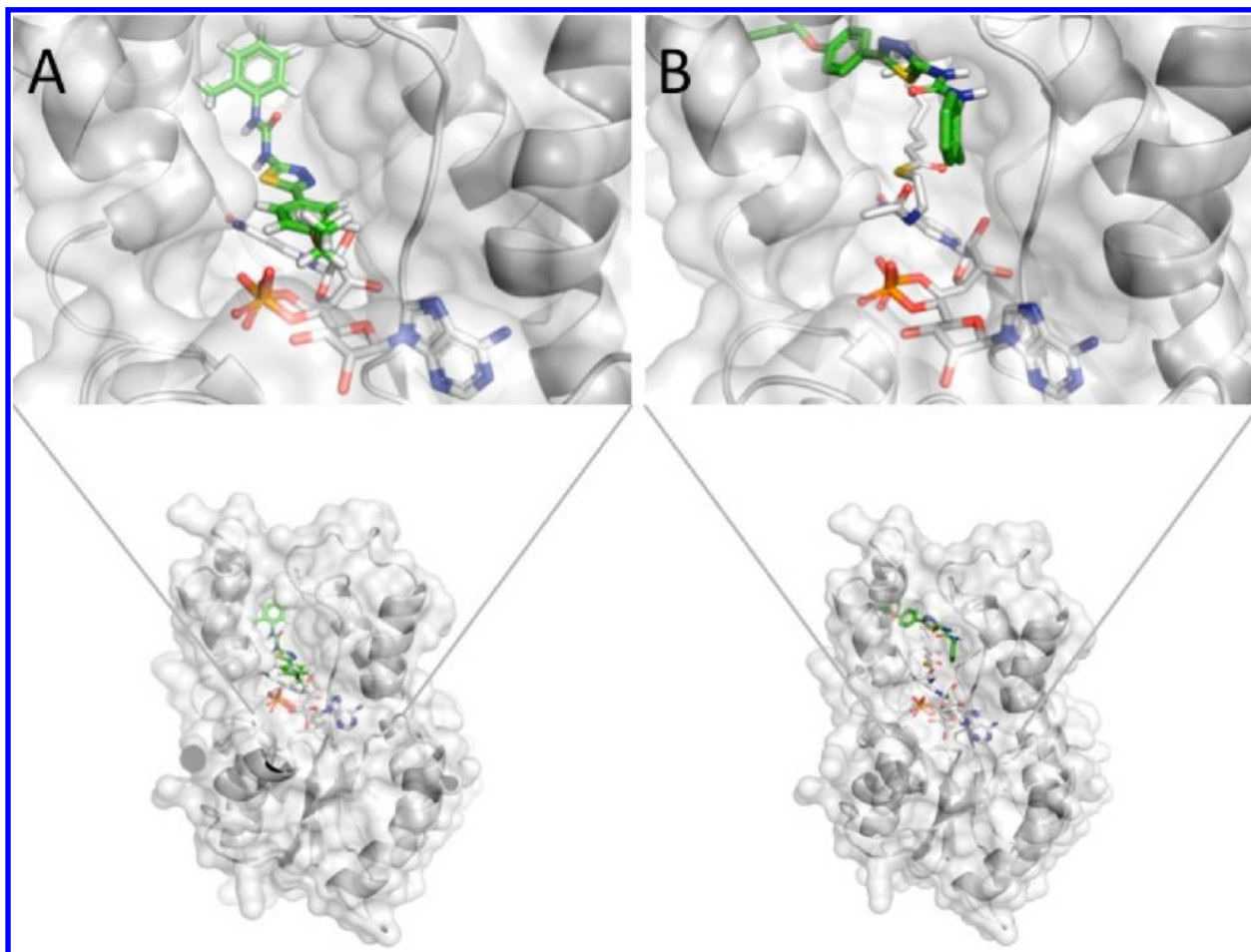

Figure 7. Docking simulations of compound ZINC09137707 to the binary and ternary complexes *Mt*InhA-NADH (A) and *Mt*InhA-NADH-C16 (B), respectively. See Figure 4 for coloring schema. Images were generated with PyMol.[21]

programs assigned the same conformation when bound to *Mt*InhA substrate-binding cavity. The results of these analyses yielded a final subset of 21 molecules.

Toxicity Evaluation. The final set of 55 compounds (34 from Approach 1 and 21 from Approach 2) was submitted to OSIRIS toxicity prediction server.[36] For the final set of hit candidates we selected only the molecules with no toxicity risks

(mutagenicity, tumorigenicity, irritability, reproductive effectiveness) and with the most favorable cLogP solubility, molecular weight, drug-likeness, and drug-score features. According to these criteria, 19 molecules were selected.

Similarity Ensemble Approach (SEA). Chemically similar drugs often bind biologically diverse protein targets, and proteins with similar sequences or structures do not always recognize the

same ligands. SEA considers proteins from a chemocentric point of view, relating them through the chemical similarity of their ligands. The idea is that similar molecules have similar biological profiles and bind similar targets.[37] In general, the authors classified similarity according to the E-value. On one hand, for E-value $< 1 \times 10^{-10}$ the similarity is significant, on the other hand for E-value > 1.0 it is insignificant. The most interesting and significant results for our set of molecules were obtained for molecules ZINC12509636 and ZINC14986320, referenced as antimalarials with E-values of 9.38×10^{-29} and 4.61×10^{-19}, respectively.

Lack of Correlation between Results from Both Approaches. Comparing the set of molecules resultant from both approaches, we could not identify a superposition of compounds. In other words, these approaches were unable to reach a common end result. One explanation could be the large size of the initial data set (about one million molecules) and the selection of just the 100 best scored molecules by GOLD in Approach 2, while all the other molecules were eliminated from the subsequent steps. Notwithstanding, it is interesting because it allows an enlargement of the chemical space occupied by these compounds.

Selection of the Molecules to *In Vitro* Evaluation. The final selection of the molecules to be tested *in vitro* for its inhibitory activity against *Mt*InhA was based on all the analyses described above. From the 19 remaining molecules six were chosen to be tested *in vitro* aiming to cover different modes of interaction with the *Mt*InhA substrate-binding cavity. From the first pharmacophore-based approach, we chose ZINC22559057 which interacts with the binding cavity mainly through hydrophobic contacts. ZINC12759934 interacts via a hydrogen bond to NADH and several hydrophobic interactions with the protein, including Ala198, important to slow tight binding. ZINC09137707 is not hydrogen bonded to NADH; however, it makes two hydrogen bonds with Tyr158 and several hydrophobic interactions (Figure 4 and Table 1, Figure 5 and Table 2).

Inhibition Studies Assessed by Steady-State Kinetics. A preliminary screen showed that three of the six selected compound presented a satisfactory percentage (%) of inhibition, thus requiring further analysis by a different approach. Compound ZINC02931014 at a final concentration of 2 μM showed 5.5% inhibition of *Mt*InhA enzymatic activity, whereas ZINC12509636 at a final concentration of 4 μM showed 2.5% inhibition. The linear decrease in *Mt*InhA enzyme activity as a function of increasing ZINC02931014 and ZINC12509636 concentrations allowed a rough estimate, although not canonical, of IC_{50} values of 24 ($\pm$2) μM and 83 ($\pm$5) μM, respectively. Compound ZINC09137707 at a final concentration of 1 μM showed 15% of *Mt*InhA enzyme inhibition. Accordingly, the ZINC09137707 chemical compound was selected for further studies. The mode of inhibition and overall inhibition constant (K_i) values were derived from initial rate measurements as described in the Materials and Methods section. Double-reciprocal plots at different ZINC09137707 concentrations displayed a pattern of parallel lines, suggesting that this compound acts as an uncompetitive inhibitor toward NADH (Figure 6A) and 2-*trans*-dodecenoyl-CoA (Figure 6B), in which V_{max} and K_m values were simultaneously reduced.

Data fitting to eq 1 for uncompetitive inhibition yielded K_i values of 24 ($\pm$3) μM and 20 ($\pm$2) μM toward, respectively, NADH and 2-*trans*-dodecenoyl-CoA. This profile of inhibition indicates that the inhibitor binds exclusively to the enzyme–substrate (*ES*) complex yielding an inactive enzyme–substrate–

inhibitor (*ESI*) complex, and inhibitor affinity is greatest at saturating concentrations of substrate.[40] Interestingly, for an uncompetitive inhibitor, enzyme inhibition cannot be overcome by high substrate concentrations, which is a weakness of competitive inhibitors. These wet inhibition studies provide support to the "dry", *in silico* screening method here described as a tool for mining chemical compounds with inhibitory activity toward *Mt*InhA enzyme and may prove useful to further drug discovery efforts.

These findings suggest that a conformational change is needed for a proper inhibitor association, and since this compound is uncompetitive to both substrates, the structural change triggered by NADH or 2-*trans*-dodecenoyl-CoA could be similar, allowing the inhibitor to bind. In order to investigate these hypotheses, molecular docking simulations were performed to test the ability of compound ZINC09137707 to accommodate in the substrate binding site, as proposed by the pharmacophore model. Figure 7 presents compound ZINC09137707 bound into the substrate-binding cavity in the presence of NADH (Figure 7A) and 2-*trans*-dodecenoyl-CoA and NADH (Figure 7B).

Taking these results into account, we can suggest why compound ZINC09137707 is not a competitive inhibitor to 2-*trans*-dodecenoyl-CoA. The most probable reason is that the *Mt*InhA substrate-binding cavity is large enough to accommodate, simultaneously, both substrate and inhibitor, as shown in Figure 7, where two association possibilities were tested. From the molecular docking experiments we can observe that the best pose to ZINC09137707 is situated in the 2-*trans*-dodecenoyl-CoA cavity providing further evidence that the pharmacophore model was precise enough to find inhibitors that bind exclusively in this pocket.

■ **AUTHOR INFORMATION**

Corresponding Author
*E-mail: osmar.norberto@pucrs.br (O.N.S.), aandrico@ifsc.usp.br (A.D.A.).

Present Address
†Ivani Pauli - Laboratório de Química Medicinal e Computacional - LQMC, Instituto de Física de São Carlos — IFSC-USP, Brazil.

Notes
The authors declare no competing financial interest.

■ **REFERENCES**

(1) Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. The Challenge of New Drug Discovery for Tuberculosis. *Nature* **2011**, *469*, 483−490.

(2) Jain, A.; Mondal, R. Extensively Drug-Resistant Tuberculosis: Current Challenges and Threats. *FEMS Immunol. Med. Microbiol.* **2008**, *53*, 145−150.

(3) Velayati, A. A.; Farnia, P.; Masjedi, M. R.; Ibrahim, T. A.; Tabarsi, P.; Haroun, R. Z.; Kuan, H. O.; Ghanavi, J.; Varahram, M. Totally Drug-Resistant Tuberculosis Strains: Evidence of Adaptation at the Cellular Level. *Eur. Respir. J.* **2009**, *34*, 1202−1203.

(4) Di Perri, G.; Aguilar Marucco, D.; Mondo, A.; Gonzalez de Requena, D.; Audagnotto, S.; Gobbi, F.; Bonora, S. Drug-Drug Interactions and Tolerance in Combining Antituberculosis and Antiretroviral Therapy. *Expert Opin. Drug Saf.* **2005**, *4*, 821−836.

(5) Morrison, J.; Pai, M.; Hopewell, P. C. Tuberculosis and Latent Tuberculosis Infection in Close Contacts of People with Pulmonary Tuberculosis in Low-Income and Middle-Income Countries: a Systematic Review and Meta-Analysis. *Lancet Infect. Dis.* **2008**, *8*, 359−368.

(6) Agueero, F.; Al-Lazikani, B.; Aslett, M.; Berriman, M.; Buckner, F. S.; Campbell, R. K.; Carmona, S.; Carruthers, I. M.; Chan, A. W. E.; Chen, F.; Crowther, G. J.; Doyle, M. A.; Hertz-Fowler, C.; Hopkins, A. L.; McAllister, G.; Nwaka, S.; Overington, J. P.; Pain, A.; Paolini, G. V.; Pieper, U.; Ralph, S. A.; Riechers, A.; Roos, D. S.; Sali, A.; Shanmugam, D.; Suzuki, T.; Van Voorhis, W. C.; Verlinde, C. L. M. J. Innovation Genomic-Scale Prioritization of Drug Targets: the TDR Targets Database. *Nat. Rev. Drug Discovery* **2008**, *7*, 900−907.

(7) Campbell, J. W.; Cronan, J. E. Bacterial Fatty Acid Biosynthesis: Targets for Antibacterial Drug Discovery. *Annu. Rev. Microbiol.* **2001**, *55*, 305−332.

(8) Schroeder, E. K.; de Souza, N.; Santos, D. S.; Blanchard, J. S.; Basso, L. A. Drugs that Inhibit Mycolic Acid Biosynthesis in Mycobacterium tuberculosis. *Curr. Pharm. Biotechnol.* **2002**, *3*, 197−225.

(9) Vilchèze, C.; Morbidoni, H. R.; Weisbrod, T. R.; Iwamoto, H.; Kuo, M.; Sacchettini, J. C.; Jacobs, W. R. Inactivation of the inhA-Encoded Fatty Acid Synthase II (FASII) Enoyl-Acyl Carrier Protein Reductase Induces Accumulation of the FASI End Products and Cell Lysis of Mycobacterium smegmatis. *J. Bacteriol.* **2000**, *182*, 4059−4067.

(10) Kumar, A.; Siddiqi, M. I. CoMFA Based de Novo Design of Pyrrolidine Carboxamides as Inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis. *J. Mol. Model.* **2008**, *14*, 923−935.

(11) Lu, X. Y.; You, Q. D.; Chen, Y. D. Recent Progress in the Identification and Development of InhA Direct Inhibitors of Mycobacterium tuberculosis. *Mini-Rev. Med. Chem.* **2010**, *10*, 181−192.

(12) Hartkoorn, R. C.; Sala, C.; Neres, J.; Pojer, F.; Magnet, S.; Mukherjee, R.; Uplekar, S.; Boy-Röttger, S.; Altmann, K. H.; Cole, S. T. Towards a New Tuberculosis Drug: Pyridomycin - Nature's Isoniazid. *EMBO Mol. Med.* **2012**, *4*, 1032−1042.

(13) Menendez, C.; Chollet, A.; Rodriguez, F.; Inard, C.; Pasca, M. R.; Lherbet, C.; Baltas, M. Chemical Synthesis and Biological Evaluation of Triazole Derivatives as Inhibitors of InhA and Antituberculosis Agents. *Eur. J. Med. Chem.* **2012**, *52*, 275−283.

(14) Deraeve, C.; Dorobantu, I. M.; Rebbah, F.; Le Quéméner, F.; Constant, P.; Quémard, A.; Bernardes-Génisson, V.; Bernadou, J.; Pratviel, G. Chemical Synthesis, Biological Evaluation and Structure-Activity Relationship Analysis of Azaisoindolinones, a Novel Class of Direct Enoyl-ACP Reductase Inhibitors as Potential Antimycobacterial Agents. *Bioorg. Med. Chem.* **2011**, *19*, 6225−6232.

(15) Postigo, M. P.; Guido, R. V.; Oliva, G.; Castilho, M. S.; da R Pitta, I.; de Albuquerque, J. F.; Andricopulo, A. D. Discovery of New Inhibitors of Schistosoma mansoni PNP by Pharmacophore-Based Virtual Screening. *J. Chem. Inf. Model.* **2010**, *50*, 1693−1705.

(16) Kuo, M. R.; Morbidoni, H. R.; Alland, D.; Sneddon, S. F.; Gourlie, B. B.; Staveski, M. M.; Leonard, M.; Gregory, J. S.; Janjigian, A. D.; Yee, C.; Musser, J. M.; Kreiswirth, B.; Iwamoto, H.; Perozzo, R.; Jacobs, W. R.; Sacchettini, J. C.; Fidock, D. A. Targeting Tuberculosis and Malaria Through Inhibition of Enoyl Reductase: Compound Activity and Structural Data. *J. Biol. Chem.* **2003**, *278*, 20851−20859.

(17) Irwin, J. J.; Shoichet, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−182.

(18) Clark, M.; Cramer, R. D.; Vanopdenbosch, N. Validation of the General-Purpose Tripos 5.2 Force-Field. *J. Comput. Chem.* **1989**, *10*, 982−1012.

(19) Powell, M. J. D. Restart Procedures for Conjugate Gradient Method. *Math. Program.* **1977**, *12*, 241−254.

(20) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Eletronegativity - a Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219−3228.

(21) DeLano, W. L. *PyMol*; DeLano Scientific: San Carlos, CA, USA, 2002.

(22) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(23) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. Protein Data Bank - Computer-Based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535−542.

(24) Sippl, M. J.; Wiederstein, M. A Note on Difficult Structure Alignment Problems. *Bioinformatics* **2008**, *24*, 426−427.

(25) Sippl, M. J. On Distance and Similarity in Fold Space. *Bioinformatics* **2008**, *24*, 872−873.

(26) Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: Computed Atlas of Surface Topography of Proteins with Structural and Topographical Mapping of Functionally Annotated Residues. *Nucleic Acids Res.* **2006**, *34*, W116−118.

(27) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a Program to Generate Schematic Diagrams of Protein-Ligand Interactions. *Protein Eng.* **1995**, *8*, 127−134.

(28) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.

(29) Klebe, G. *Virtual Screening in Drug Discovery*, 1st ed.; CRC Press: Boca Raton, FL, 2005; Vol. *1*.

(30) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed Automated Docking of Flexible Ligands to Proteins: Parallel Applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293−304.

(31) Goodsell, D. S.; Olson, A. J. Automated Docking of Substrates to Proteins by Simulated Annealing. *Proteins* **1990**, *8*, 195−202.

(32) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(33) Jain, A. N. Surflex: Fully Automatic Flexible Molecular Docking Using a Molecular Similarity-Based Search Engine. *J. Med. Chem.* **2003**, *46*, 499−511.

(34) Jain, A. N. Scoring Noncovalent Protein-Ligand Interactions: a Continuous Differentiable Function Tuned to Compute Binding Affinities. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 427−440.

(35) Helma, C. In Silico Predictive Toxicology: the State-of-the-Art and Strategies to Predict Human Health Effects. *Curr. Opin. Drug Discovery Dev.* **2005**, *8*, 27−31.

(36) Sander, T.; Freyss, J.; von Korff, M.; Reich, J. R.; Rufener, C. OSIRIS, an Entirely In-House Developed Drug Discovery Informatics System. *J. Chem. Inf. Model.* **2009**, *49*, 232−246.

(37) Keiser, M. J.; Roth, B. L.; Armbruster, B. N.; Ernsberger, P.; Irwin, J. J.; Shoichet, B. K. Relating Protein Pharmacology by Ligand Chemistry. *Nat. Biotechnol.* **2007**, *25*, 197−206.

(38) Quémard, A.; Sacchettini, J. C.; Dessen, A.; Vilcheze, C.; Bittman, R.; Jacobs, W. R.; Blanchard, J. S. Enzymatic Characterization of the Target for Isoniazid in Mycobacterium tuberculosis. *Biochemistry* **1995**, *34*, 8235−8241.

(39) Basso, L. A.; Zheng, R. J.; Musser, J. M.; Jacobs, W. R.; Blanchard, J. S. Mechanisms of Isoniazid Resistance in Mycobacterium tuberculosis: Enzymatic Characterization of Enoyl Reductase Mutants Identified in Isoniazid-Resistant Clinical Isolates. *J. Infect. Dis.* **1998**, *178*, 769−775.

(40) Oliveira, J. S.; Sousa, E. H.; Basso, L. A.; Palaci, M.; Dietze, R.; Santos, D. S.; Moreira, I. S. An Inorganic Iron Complex that Inhibits Wild-Type and an Isoniazid-Resistant Mutant 2-Trans-Enoyl-ACP

(CoA) Reductase from Mycobacterium tuberculosis. *Chem. Commun. (Cambridge, U. K.)* **2004**, 312−313.

(41) Oliveira, J. S.; de Sousa, E. H.; de Souza, O. N.; Moreira, I. S.; Santos, D. S.; Basso, L. A. Slow-Onset Inhibition of 2-Trans-enoyl-ACP (CoA) Reductase from Mycobacterium tuberculosis by an Inorganic Complex. *Curr. Pharm. Des.* **2006**, *12*, 2409−2424.

(42) Dessen, A.; Quemard, A.; Blanchard, J. S.; Jacobs, W. R.; Sacchettini, J. C. Crystal-Structure and Function of the Isoniazid Target of Mycobacterium tuberculosis. *Science* **1995**, *267*, 1638−1641.

(43) Wermuth, C.; Ganellin, C.; Lindberg, P.; Mitscher, L. In *Chapter Pure and Applied Chemistry*; 1998; Vol. *70*, pp 1129−1143.

(44) Luckner, S. R.; Liu, N.; am Ende, C. W.; Tonge, P. J.; Kisker, C. A Slow, Tight Binding Inhibitor of InhA, the Enoyl-Acyl Carrier Protein Reductase from Mycobacterium tuberculosis. *J. Biol. Chem.* **2010**, *285*, 14330−14337.

(45) Parikh, S.; Moynihan, D. P.; Xiao, G.; Tonge, P. J. Roles of Tyrosine 158 and Lysine 165 in the Catalytic Mechanism of InhA, the Enoyl-ACP Reductase from Mycobacterium tuberculosis. *Biochemistry* **1999**, *38*, 13623−13634.

(46) Rozwarski, D. A.; Vilchèze, C.; Sugantino, M.; Bittman, R.; Sacchettini, J. C. Crystal Structure of the Mycobacterium tuberculosis Enoyl-ACP Reductase, InhA, in Complex with NAD+ and a C16 Fatty Acyl Substrate. *J. Biol. Chem.* **1999**, *274*, 15582−15589.

(47) *UNITY Chemical Information Software*, 4.1; Tripos Inc.: St. Louis, MO, 2006.