J Mol Model (2012) 18:4013–4024
DOI 10.1007/s00894-012-1383-6

ORIGINAL PAPER

Interaction of wild type, G68R and L125M isoforms of the arylamine-*N*-acetyltransferase from *Mycobacterium tuberculosis* with isoniazid: a computational study on a new possible mechanism of resistance

Ricardo Martins Ramos · Janaína Menezes Perez · Luis André Baptista · Hermes Luís Neubauer de Amorim

Received: 24 October 2011 / Accepted: 14 February 2012 / Published online: 30 March 2012

Abstract Isoniazid (INH) is a front-line drug used in the treatment of tuberculosis (TB), a disease that remains a major cause of death worldwide. Isoniazid is a prodrug, requiring activation in the mycobacterial cell by the catalase-peroxidase (CP) enzyme. Recent studies have suggested that acetylation of INH by the arylamine-N-acetyltransferase from *Mycobacterium tuberculosis* (TBNAT) may be a possible cause of inactivation of the drug thus resulting in resistant strains. In this study, computational techniques were applied to investigate the binding of isoniazid to three TBNAT isoforms: wild type, G68R and L125M. Since there is no experimental structure available, molecular dynamics (MD) simulations were initially used for the refinement of TBNAT homology models. Distinct conformations of the models were selected during the production stage of MD simulations for molecular docking experiments with the drug. Finally, each mode of binding was refined by new molecular MD simulations. Essential dynamics (ED) analysis and linear interaction energy calculations (LIE) were used to evaluate the impact of amino acid substitutions on the structural and binding properties of the enzymes. The results suggest that the wild type and the G68R TBNATs have a similar pattern of affinity to INH. On the other hand, the calculated enzyme-INH dissociation constant (KD) was estimated 33 times lower for L125M isoform in comparison with wild type enzyme. This last finding is consistent with the hypothesis that isolated mutations in the *tbnat* gene can produce *M. tuberculosis* strains resistant to isoniazid.

Keywords Arylamine-N-acetyltransferase from *Mycobacterium tuberculosis* · Docking · Essential dynamics · Isoniazid · Linear interaction energy · Molecular dynamics simulations

Abbreviations

TB	Tuberculosis
INH	Isoniazid
CP	Catalase-peroxidase
NAT	Arylamine N-acetyltransferase
TBNAT	Arylamine-N-acetyltransferase from Mycobacterium tuberculosis
MD	Molecular dynamics
ED	Essential dynamics
LIE	Linear interaction energy
SNPs	Single nucleotide polymorphisms
K_D	Dissociation constant
NCBI	National Center for Biotechnology Information
InChI™	The IUPAC International Chemical Identifier
AM1	Austin model 1
MEP	Molecular electrostatic potential
CHELPG	Charges from electrostatic grid based
SPC	Simple point charge

R. M. Ramos · J. M. Perez · L. A. Baptista · H. L. N. de Amorim (✉)
Laboratório de Bioinformática Estrutural (LaBiE), Universidade Luterana do Brasil (ULBRA), Av. Farroupilha 1001, Prédio 01, Sala 122, Canoas, RS, Brazil 92450-900
e-mail: hermes.amorim@ulbra.br

R. M. Ramos
Instituto Federal do Piauí (IFPI), Praça da Liberdade 1597, Centro, Teresina, PI, Brazil 64000-040

ADT	AutoDock tools
RMSD	Root mean square deviation
LGA	Lamarckian genetic algorithm
LS	Local search
vdW	van der Waals
el	Electrostatic
Rg	Radius of gyration
NHb	Number of intramolecular hydrogen bonds
SASA	Surface accessible solvent area
RMSF	Root mean square fluctuation
MC	Monte Carlo

Introduction

Since its introduction in 1952 isoniazid, INH (Fig. 1), became one of the front-line drugs for tuberculosis (TB) treatment [1–3]. TB is responsible for more deaths throughout the world than any other bacterial infectious disease [4]. *Mycobacterium tuberculosis* is the etiologic agent of tuberculosis in humans. The cell wall of *M. tuberculosis* has as its main component a class of long fatty acids called mycolic acids. They constitute the inner leaflet of the lipid bilayer of microbial cell and form an effective barrier that allows the survival of the microorganism [5]. The role of isoniazid is to inhibit the formation of the cell wall of mycolic acids, resulting in cell death [6]. The mechanism of action of INH is complex and depends on its activation by the catalase-peroxidase (CP) enzyme (encoded by *katG* gene) to generate a reactive compound that has lethal effect for intracellular targets [7, 8].

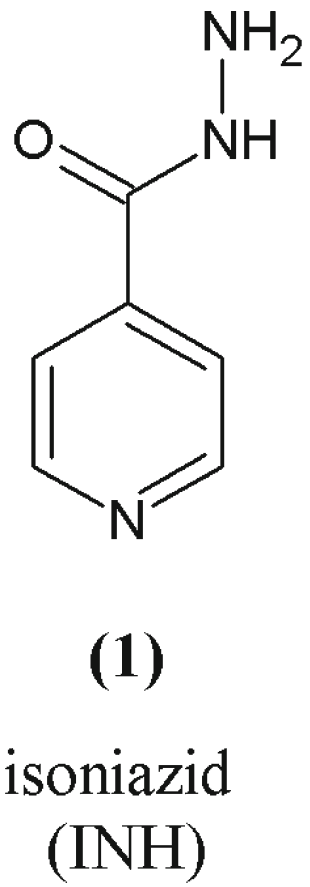

Fig. 1 Isoniazid molecular structure

The resistance to INH drug has been associated with mutations in several genes including *katG*, *inhA* and *ahpC* [9–14]. On the other hand, 20–30% of the INH resistant isolates do not show mutation in these genes, suggesting the involvement of other genes or mechanisms. A possible route of INH inactivation has been attributed to the *tbnat* gene [15–18], which expresses an enzyme of the arylamine N-acetyltransferase (NAT) family.

Arylamine N-acetyltransferases (EC 2.3.1.5) catalyze the N-acetylation of arylamines and hydrazines and O-acetylation of N-hydroxy-arylamines and heterocyclic amines [19]. The enzyme mechanism involves the transfer of an acetyl group to the terminal acceptor atom of the substrate, using acetyl coenzyme A (acetyl-CoA) as cofactor. These reactions are important for the activation and deactivation of exocyclic amine-containing procarcinogens, and for the metabolism of some pharmaceutical drugs [20, 21].

Studies have indicated that NAT of *M. tuberculosis* (TBNAT) is able to acetylate INH, representing an additional route of drug metabolism [22]. In this case, the acetylation of INH by TBNAT makes the drug therapeutically inactive [22, 23]. In fact, Payton and colleagues [24] had previously demonstrated that the increased expression of NAT genes from *M. smegmatis* and *M. tuberculosis* results in isoniazid resistance which, in turn, is consistent with the hypothesis of a mechanism of competition between the catalase-peroxidase and the NAT enzymes for INH [23].

Several amino acid mutations caused by single nucleotide polymorphisms (SNPs) in the *tbnat* gene have been described [25]. One of these is responsible for replacement of a glycine at position 68 by an arginine (TBNAT_G68R) [26]. Although this mutation is observed in INH-resistant strains that have mutations in other important genes, the location of mutated amino acid (near to Cys 70, the nucleophilic cysteine of the catalytic triad) and the type of physical-chemical changes involved (in terms of volume and charge) is suggestive of a possible influence on the enzyme activity. More recently, the presence of polymorphisms in the N-acetyltransferase gene of 41 clinical isolates of isoniazid resistant *M. tuberculosis* with no mutations in the hot spots of the genes previously described to be involved in INH resistance (*katG*, *inhA* and *ahpC*) was investigated [18]. Two of these showed non-synonymous polymorphisms, C373A (Leu → Met) and T503G (Met → Arg). Particularly, the analysis of homology models showed that the L125M substitution (located on β−4 strand) can lead to local disturbances that can affect the conformation

and/or dynamics of residues situated at the bottom of the cleft for substrate and acetyl-CoA binding (e.g. Val95, Trp97, His110) [18].

Unfortunately, low solubility of recombinant TBNAT has hampered detailed structural crystallographic studies of this enzyme [27]. With this respect, taking into account that the structural information is a powerful ally for the understanding of how SNPs affect the structure-function relationships in proteins, the establishment of computational models to evaluate the interaction of proteins with substrates and inhibitors can be considered very useful.

In order to contribute to the discussion about the possible role of the *tbnat* gene product in isoniazid resistance of the *Mycobacterium tuberculosis* complex, in this work computational techniques were employed to investigate the interaction of wild type TBNAT and its isoforms G68R and L125M with INH. The strategy used involved the use of molecular dynamics (MD) simulations to refine the structural models obtained by homology modeling in a previous work. Distinct conformations were selected during the production stage of MD simulations for molecular docking experiments with the drug. Essential dynamics (ED) analysis and linear interaction energy calculations (LIE), a free energy approach which allows a good estimate of the relative protein-ligand affinities, were used to evaluate the impact of amino acid substitutions on the structural and binding properties of the enzymes. The results suggest that the wild type and the G68R TBNATs have a similar pattern of affinity for INH. On the other hand, the calculated enzyme-INH dissociation constant (K_D) was estimated 33 times lower for L125M isoform in comparison with wild type enzyme. This last finding is consistent with the hypothesis that isolated mutations in the *tbnat* gene can produce *M. tuberculosis* strains resistant to isoniazid.

Methods

Isoniazid setup

INH structural formula was obtained from National Center for Biotechnology Information (NCBI) site under CID code 3767. The ligand InChI™ (The IUPAC International Chemical Identifier) file was converted to MDL Molfile [28] format with the ACD/ChemSketch program [29]. ArgusLab [30] was used as graphical interface to generate the INH three-dimensional model and for geometry pre-optimization with the semi-empirical method Austin model 1 (AM1) [31].

INH topology, necessary for molecular dynamics simulations, was generated from the pre-optimized atomic coordinates with the PRODRG program [32], beta version. INH atomic charges generated from PRODRG were substituted by atomic charges calculated using the quantum model of molecular electrostatic potential (MEP) [33], charges from electrostatic grid based scheme (CHELPG) [34]. The atomic charges were calculated by single point procedure based on Hartree Fock method with a 6-31G* basis set. The calculations were carried out using the GAUSSIAN 03 program [35].

TBNAT structural models

Structural models for wild type TBNAT (TBNAT_WT) and L125M isoform (TBNAT_L125M) were obtained as described in [18]. The homology model of G68R isoform of TBNAT (TBNAT_G68R) was built in an analogous way, using the refined homology model of TBNAT_WT as template.

Molecular dynamics simulations of uncomplexed TBNATs

Molecular dynamics simulations were carried out using GROMOS96 53a6 force field [36] implemented in GROMACS package [37], version 4.0. All systems were simulated in NPT ensemble and periodic boundary conditions (dodecahedron). The dimensions of the central box were chosen in such a way that the minimum distance of any protein atom to the closest box wall was 12 Å. The simulations were carried out using explicit solvent water molecules described by the simple point charge (SPC) model [38]. The total charge of the systems was -7 for TBNAT_WT and TBNAT_L125M and -6 for TBNAT_G68R. Sodium ions were added to neutralize each system.

Initially, the protein structure in each system was submitted to a maximum of 500 steps of steepest descent energy minimization. To relax strong solvent-solvent and solvent-protein non-bonded interactions, 1 ps of MD simulation was performed restraining the protein structure. Initial velocities were assigned according to Maxwell distribution. All simulations were performed for 20 ns using an integration time step of 2 fs. Each system was heated with gradual increments in the following temperatures: 100 K (10 ps), 150 K (5 ps), 200 K (5 ps), and 250 K (5 ps). After, the temperatures of the systems were adjusted to 310 K. The first 15 ns of each simulation was considered as part of the heating (0.025 ns) and the equilibration (14.975 ns) steps and had not been used in the data analysis. The temperatures of solvent and solutes (protein and sodium ions) were independently coupled to a thermal bath with a relaxation time of 0.1 ps using the Berendsen thermostat [39]. The pressure in the systems was weakly coupled to a pressure bath of 1 atm applying an isotropic scaling and 0.5 ps of relaxation time using the Parrinello-Rahman barostat [40, 41]. Bond lengths were constrained using the LINCS algorithm [42] with 4th order expansion. Electrostatic interactions

among non-ligand atoms were evaluated by the PME method [43] with a charge grid spacing of approximately 1.0 Å. The charge grid was interpolated on a cubic grid with the direct sum tolerance set. Lennard-Jones interactions were evaluated using a 14 Å atom-based cutoff. The pair list was updated every 10 steps.

Essential dynamics (ED) analysis

ED analysis was performed upon the last 5 ns of trajectories. Overall rotation and translation motions were removed by least-squares fitting to the first structure of simulation. Non-mass weighted covariance matrix was constructed from backbone atomic displacements with the g_covar module of GROMACS package [37]. After, g_anaeig and g_analyze GROMACS modules were used to perform the analysis of principal modes. The choice of dimension of essential subspace, the number of principal modes retained at model, was based on the distribution of eigenvectors. The principal modes retained at model were those whose distributions were different of Gaussian distribution.

Docking protocol

All docking procedures utilized the Autodock 4.2 package [44–46]. Protein and ligand were prepared for docking simulations with AutoDock Tools (ADT), version 1.5.4 [47]. For each system (TBNAT_WT, TBNAT_G68R and TBNAT_L125M) three distinct transient conformations obtained from the last 5 ns of the MD simulation were chosen for the docking experiments: the conformation with lowest potential energy (**emim**), the conformation with lowest root mean square deviation (RMSD) value with respect to the starting structure (**rmin**) and the conformation with largest RMSD value with respect to the starting structure (**rmax**). All water molecules were removed from the original TBNAT files. Protein was considered rigid whereas the ligand was considered flexible. Gasteiger [48] partial charges were calculated after addition of all hydrogens. Nonpolar hydrogens of protein and ligand were subsequently merged. A cubic box of 60×60×60 points, with a spacing of 0.35 Å between the grid points, was generated for the whole protein target. The grid box was centered on the Sγ of Cys70. The global search Lamarckian genetic algorithm (LGA) [49] and the local search (LS) pseudo-Solis and Wets [50] methods were applied in the docking search. Except for the parameters shown Table 1, the rest of the parameters were set as the default values. The resulting docked conformations were clustered into families according to the RMSD. The lowest docking-energy conformation of the cluster with lowest energy was chosen as initial structure for the molecular dynamics simulations of the TBNAT-isoniazid complexes (next section).

Table 1 Docking parameters

Parameter	Value
Number of individuals in population	150
Maximum number of energy evaluations	100,000,000
Number of iterations of pseudo Solis and Wets local search	300
Number of hybrid LGA-LS runs	100
Maximum number of generations	27,000
Elitism	1
Rate of crossover	0.8
Rate of gene mutations	0.02

Molecular dynamics simulations of complexed TBNATs

Nine MD simulations were carried out for the TBNAT-isoniazid complexes. The simulations are identified in the text as follows: WT_**emim**, WT_**rmim**, WT_**rmax**, G68R_**emim**, G68R_**rmim**, G68R_**rmax**, L125M_**emim**, L125M_**rmim** and L125M_**rmax** where the characters WT, G68R and L125M are used to identify the respective isoforms and **emin**, **rmin** and **rmax** are used to identify the conformation used in the docking experiments. Except for the simulation time (5 ns), all other parameters used in the simulations of the complexes were the same described for the uncomplexed TBNATs.

Linear interaction energy protocol

For the LIE calculations (see below), INH structure was simulated free in solution using the same protocol described for the complexes. In this case, the total charge of the system was zero. Therefore, counter-ions were not added.

The binding affinity was calculated using LIE method [51] which can be used to estimate the binding free energy from MD simulations of ligand in bound and free states. A generalized version of LIE equation has the following form:

$$\Delta G = \alpha(\langle E^{vdW}\rangle_b - \langle E^{vdW}\rangle_f) + \beta(\langle E^{el}\rangle_b - \langle E^{el}\rangle_f) + \gamma \quad (1)$$

where $\langle\ \rangle$ denotes averages of the van der Waals (vdW) and electrostatic (el) interaction energies over the molecular dynamics trajectory. The "b" subscript denotes the ligand energy interaction with all parts of the system (protein and solvent). The "f" subscript denotes the ligand energy interaction with solvent. In Eq. 1 α is an empirically determined coefficient used to scale the non-polar contributions while β is a theoretically derived parameter used to scale the electrostatic energies. The third and additional parameter (γ) in Eq. 1 is a constant free energy term that should be used when there is a non-zero difference $\gamma_{prot} - \gamma_{wat}$ between constant terms in corresponding linear expressions for the solvation energies [51, 52]. The values for the parameters used in LIE equation were established by Åqvist and

colleagues (α=0.16, β=0.50 and γ=0) [51]. Considering that the main objective was to determine the relative affinities, there is no need to calibrate the LIE equation.

Structural analysis and nonbonded interaction profiles were performed with GROMACS modules [37] and Ligplot [53]. All molecular graphics images were produced using the SwissPDBViewer software [54].

Results and discussion

The presence of polymorphisms that affect the properties of NAT enzymes is well established [19]. For instance, mutations in *M. tuberculosis* NAT which reduce enzymatic activity appear to be associated with a slowing of the growth of the bacillus [25]. In addition, the increase in the level of the *nat* gene expression can increase isoniazid resistance [24]. Also, studies showed that the deletion of the *nat* gene is capable of producing approximately 2–3 fold change on INH susceptibility of *M. bovis* BCG [5] and of *M. smegmatis* [55]. The combination of these findings with the recent work of Coelho et al. [18] is strongly indicative that TBNAT mutants may play a potential role in the resistance of *M. tuberculosis* to isoniazid. These observations justify the conducting of structural studies to investigate the interaction of isoniazid with TBNAT.

Global structural analysis of the simulations of the uncomplexed TBNATs

RMSD, radius of gyration (Rg), total number of intramolecular hydrogen bonds (NHb) and surface accessible solvent area (SASA) structural properties were monitored during the course of the trajectory (Fig. 2). In Fig. 2a, the analysis of the RMSD values for the three isoforms shows that G68R mutation leads to a greater divergence from the initial

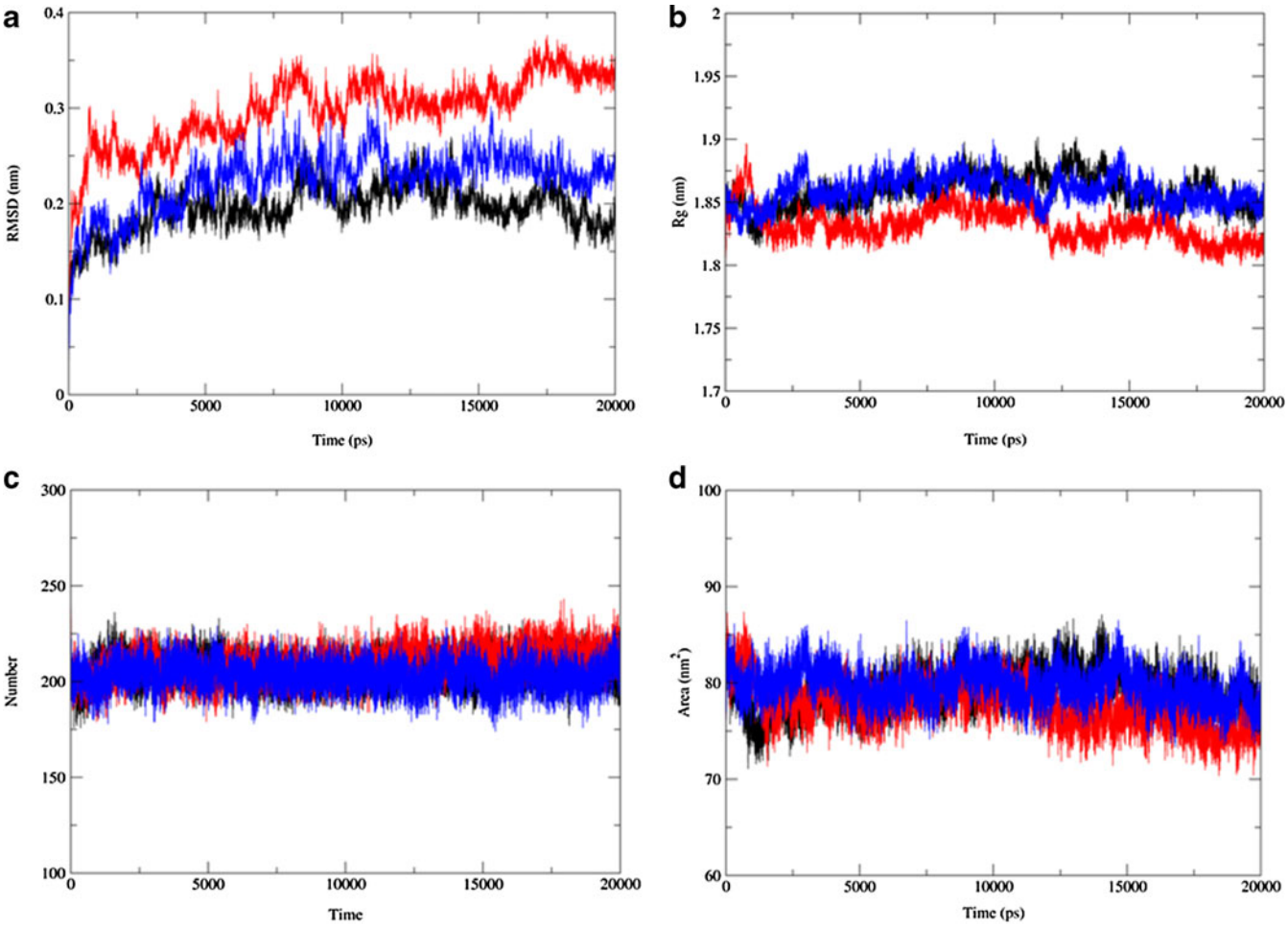

Fig. 2 Time dependence of structural parameters from MD simulations with respect to the initial structure. **a**: Atom-positional root mean-square deviation (RMSD) of Cα atoms; **b**: Radius of gyration (Rg); **c**: Total number of hydrogen bonds (NHb); **d**: Total solvent-accessible surface area (SASA). Color system: TBNAT_WT (*black*), TBNAT_G68R (*red*), TBNAT_L125M (*blue*)

structure. For this property, as presented in Table 2, the average value calculated from the last 5 ns of MD simulations was 0.19±0.01 nm (TBNAT_WT), 0.23±0.01 nm (TBNAT_L125M) and 0.33±0.01 nm (TBNAT_G68R). Among TBNATs, the greatest fluctuation in RMSD was around 5 % (0.1 Å), for the TBNAT_WT simulation. The Rg values of the three systems (Fig. 2b) fluctuated around an average value during most of the simulation time, indicating the stability with respect to this parameter. The fluctuations in NHb (Fig. 2c) and SASA (Fig. 2d) were approximately the same among the isoforms, ranging between 2-3 % for both properties. In a general way, the TBNAT_G68R presented the lowest Rg and SASA in the last 5 ns of the simulations (Table 2), indicating that the G68R substitution leads to a more compact structure.

From the data above, it is possible to conclude that the system achieved an adequate amount of sampling and that no significant structural deviations can occur after 15 ns of simulation. Structurally, the main differences observed between the TBNATs are located at domains I and III. In terms of secondary structure, the snapshots generated at 20 ns (Fig. 3) show the disappearance of the α-3 helix of the TBNAT_G68R isoform. In this case, the substitution of a glycine (neutral and short side-chain) by an arginine (charged and larger side-chain) induces local rearrangements in the segment located above the amino acid position 68 that directly affect the region between the helixes 3 and 4 (Fig. 4).

The differences found in domain III are more prominent and can be attributed to the fact that the C-terminal segment of these enzymes is inserted into the binding site [18]. The C-terminal region of the NAT proteins has been shown to have a role in the substrate specificity [56] and acetyl-CoA hydrolysis [57]. The TBNAT model built in [18] suggests that the protein domain III must be involved in interactions with acetyl-CoA through Thr209, Ser222, His229 and Lys236. Thus, local perturbations originated from the TBNAT binding site can be propagated from the C-terminal main-chain through the entire domain III.

The root mean square deviation (RMSF) in the essential subspace showed distinctive patterns of flexibility in this region for the three mutants. Indeed, as can be seen in Fig. 4, the mutations affect the flexibility of two loops belonging to the binding region (amino acid residues 97-107 and 128-134) which should be related to the conformational and dynamic differences found in domain III of the isoforms.

Table 2 Average of structural parameters of TBNAT from the last 5 ns of the simulations

Simulation	RMSD (nm)[a]	Rg (nm^3)[b]	NHb-intra[c]	SASA (nm^2)[d]
TBNAT_WT	0.19±0.01	1.85±0.00	207±6.8	79.0±1.5
TBNAT_G68R	0.33±0.01	1.82±0.00	214±7.1	75.7±1.6
TBNAT_L125M	0.23±0.01	1.85±0.00	201±7.5	79.1±1.9

[a] Root mean square deviation of TBNAT Cα atoms from starting coordinates (homology models); [b] Radius of gyration; [c] Number of hydrogen bonds (intramolecular); [d] Solvent-acessible surface area

Binding of INH in the TBNATs

The formation of a complex between isoniazid and NAT in the absence of either acetylation or acetyl CoA is consistent with earlier NMR experiments using the enzyme from *S. typhimurium* [58] and crystallographic data for the complex between isoniazid and the NAT from *M. smegmatis* [23]. As pointed out by Sandy and colleagues [23], the existence of these complexes may be related either to a "substrate inhibition" complex or a complex that exists before initiation of the ping-pong mechanism. In this case, it is unlikely that the key determinants of INH binding observed in these structures are substantially different from those that are exploited in the catalytic cycle.

Despite the interaction of wild type TBNAT with isoniazid is low affinity [59], amino acid substitutions able to affect the binding capacity of both acetyl-coenzyme and isoniazid can influence potentially the affinity and/or the catalytic ability of the mutant enzymes [18].

Nine MD simulations were performed, starting from the coordinates of complexes obtained from the docking of isoniazid onto three conformations of each TBNAT isoform: TBNAT_WT, TBNAT_G68R, TBNAT_L125M. The conformations were selected from the last 5 ns of the simulations of uncomplexed enzymes. In this case, the conformations with the lowest potential energy and those with the lowest and the highest RMSD with respect to the initial coordinates were chosen (see Methods section). For purposes of evaluating the binding mode, the complexes with the more favorable binding energy obtained by LIE method were chosen.

The LIE method has been used to calculate binding free energies for widely different compounds by averaging interaction energies obtained from MD or Monte Carlo (MC) simulations [60, 61]. The main consideration of the LIE method, deriving from linear response considerations, is that only (convergent) averages of the interaction energies between the ligand and its surroundings need to be evaluated to obtain an estimative of binding free energies. In order to circumvent any inadequacies of the model, the implicit limitations of the method, and the simplifications in its applications the relative free energies ($\Delta\Delta G_{bind}$) rather than absolute (ΔG_{bind}) energies will be considered in the discussion of the data [62].

According to Table 3, the more favorable the interaction of isoniazid was with the TBNAT_L125M isoform (L125M_**emin** simulation). The difference between the best binding energy calculated for the wild type and L125M

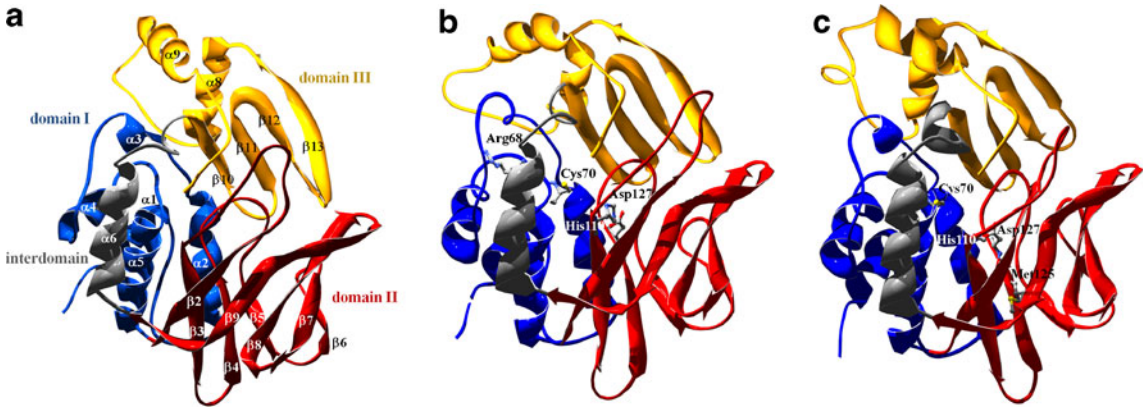

Fig. 3 Schematic representations of the refined TBNAT models. **a**: TBNAT_WT; **b**: TBNAT_G68R; **c**: TBNAT_L125M. Domain I in blue, domain II in red, interdomain in gray and domain III in yellow

isoform was 2.15 kcal mol^{-1}. This energy difference means a dissociation constant (K_D) thirty three times lower for L125M isoform in comparison with wild type enzyme. On the other hand, the difference of binding energy between the wild type and TBNAT_G68R isoform was non-significant. In fact, the TBNAT_G68R isoform was detected in *M. tuberculosis* isolates that have mutations in other *hot spots* associated with isoniazid resistance to the bacillus TB [26].

In Table 3 it is possible to see that the binding energy (ΔG_{bind}) in complexes is mainly favored by the hydrophobic interactions, which are indicated by the term that represents the difference of average interaction energy for the van der Walls component in the bound ($\langle E^{vdW}\rangle_b$) and free ($\langle E^{vdw}\rangle_f$) states. Comparatively speaking, the van der Waals component of the binding energy is 41.61 kcal mol^{-1} more negative for TBNAT_L125M isoform when compared with the wild type isoform. On the other hand, the electrostatic component (indicated by the difference of average interaction energy for the electrostatic component in the bound ($\langle E^{el}\rangle_b$) and free ($\langle E^{el}\rangle_f$) is 10.97 kcal mol^{-1} lower for TBNAT_L125M isoform when compared with TBNAT_WT isoform. These findings indicate that the increase of binding efficiency resulting from L125M substitution is mainly determined by the optimization of the hydrophobic contacts.

Figure 5 shows the frequency of isoniazid contacts with the three isoforms. The interactions were calculated for the last 2 ns of WT_**rmin,** G68R_**rmax** and L125M_**emin** MD simulations of using the LIGPLOT program [53]. Establishing as criterion efficiency for binding a minimum of 50% of contacts in the analyzed frames, it is observed that in WT_**rmin** isoniazid makes hydrogen bonds with Gly131 and Lys203 (Fig. 5_A1) whereas the hydrophobic interactions occur mainly with Phe38, Phe130, Cys70 and Phe204 (Fig. 5_A2). In G68R_**rmax** isoniazid interacts through hydrogen bonds with Cys70, Thr109 and Asn74 (Fig. 5_B1) whereas the hydrophobic interactions occur

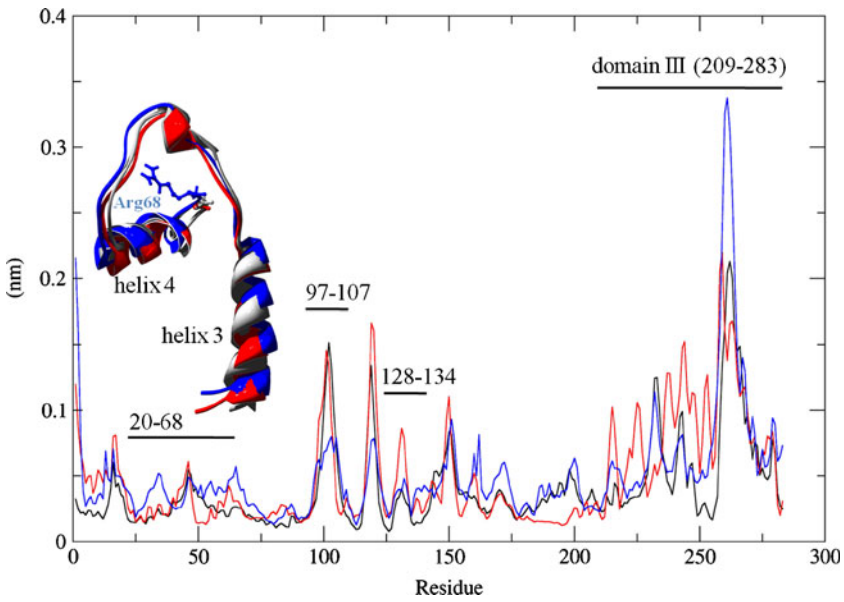

Fig. 4 RMS fluctuations of the trajectories calculated from projection of the MD trajectories onto the 5 principal components. Color system: TBNAT_WT (*black*), TBNAT_G68R (*red*), TBNAT_L125M (*blue*)

Table 3 Thermodynamic (kcal mol^{-1}) and kinetic parameters calculated for the selected TBNAT-INH complexes (see text)

Simulation	$\Delta G_{bind(docking)}$	$\langle E^{vdW}\rangle_b - \langle E^{vdW}\rangle_f$	$\langle E^{el}\rangle_b - \langle E^{el}\rangle_f$	$\Delta G_{bind(LIE)}$	$\Delta\Delta G_{bind(LIE)}$	K_D (mM)*
WT_rmin	-5.04	-40.85	3.60	-1.68	0.00	65.36
G68R_rmax	-5.17	-55.80	0.00	-2.40	-0.72	20.28
L125M_emin	-5.65	-85.46	-7.37	-3.83	-2.15	1.97

Data of energy averages determined for the last 2 ns of the MD simulations. *Dissociation constant estimated according to $K_D = e^{-\Delta G/RT}$

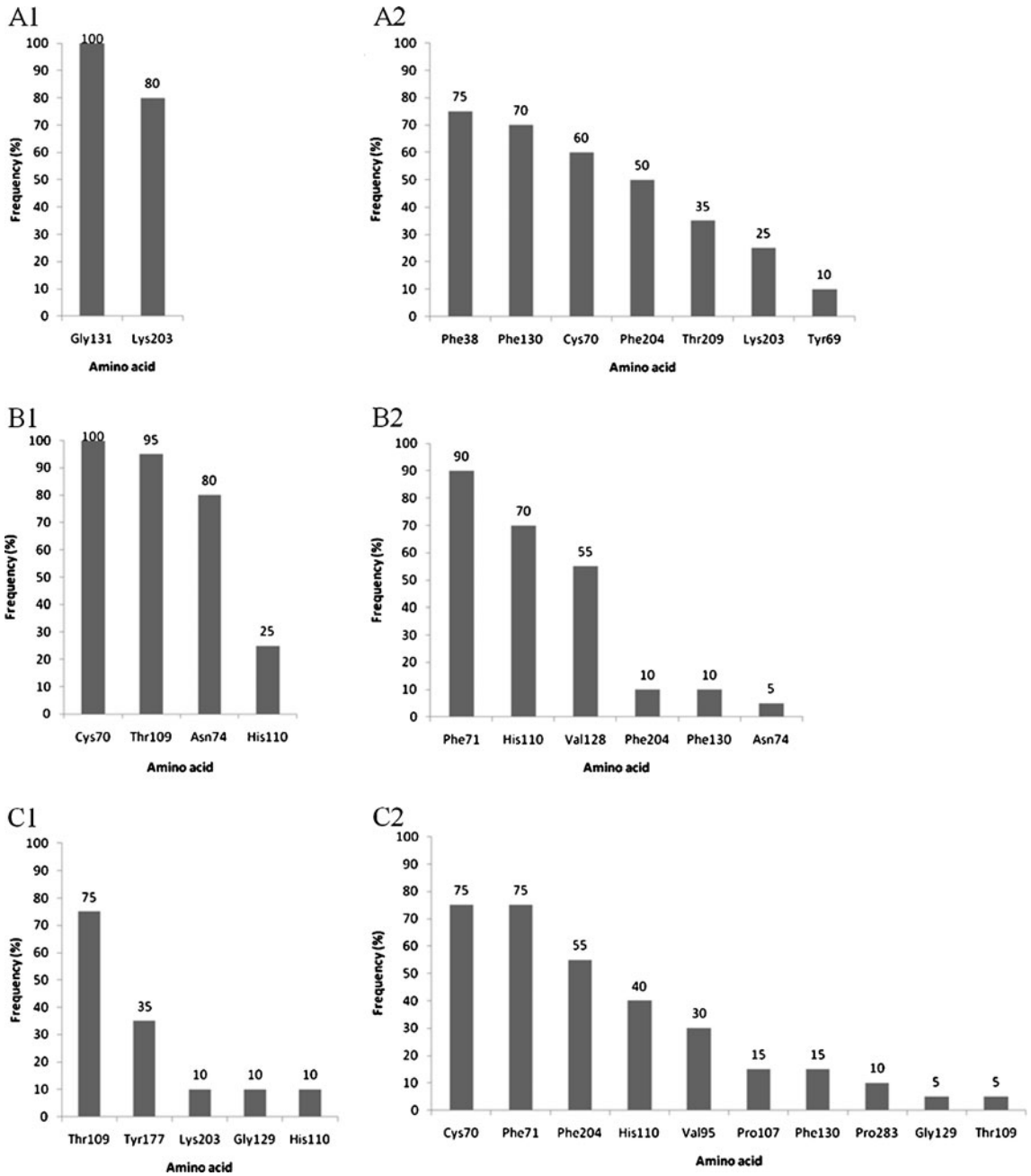

Fig. 5 Indentified contacts between INH and TBNAT isoforms calculated for the last 2.0 ns of MD simulations. *Left hand panels:* amino acids that interact through hydrogen bonds with INH in WT_**rmin** (A1), G68R_**rmax** (B1) and L125M_**emin** (C1) simulations; *Right-hand panels:* amino acids that make hydrophobic interactions with INH in WT_**rmin** (A2), G68R_**rmax** (B2) and L125M_**emi**n (C2) simulations. The numbers on the bars indicate the percentage of contacts for each amino acid

mainly with Phe71, His110 and Val128 (Fig. 5_B2). In turn, Fig. 5_C1 indicates that in L125M_**emin** isoniazid makes one hydrogen bond with Thr109 whereas the hydrophobic interactions occur with Cys70, Phe71 and Phe204 (Fig. 5_C2).

Figure 6 shows the graphic representation of TBNAT structures and INH interactions for the last frame (5 ns) of the MD simulations of the complexes. Despite representing a single snapshot, the interactions observed in the last frame

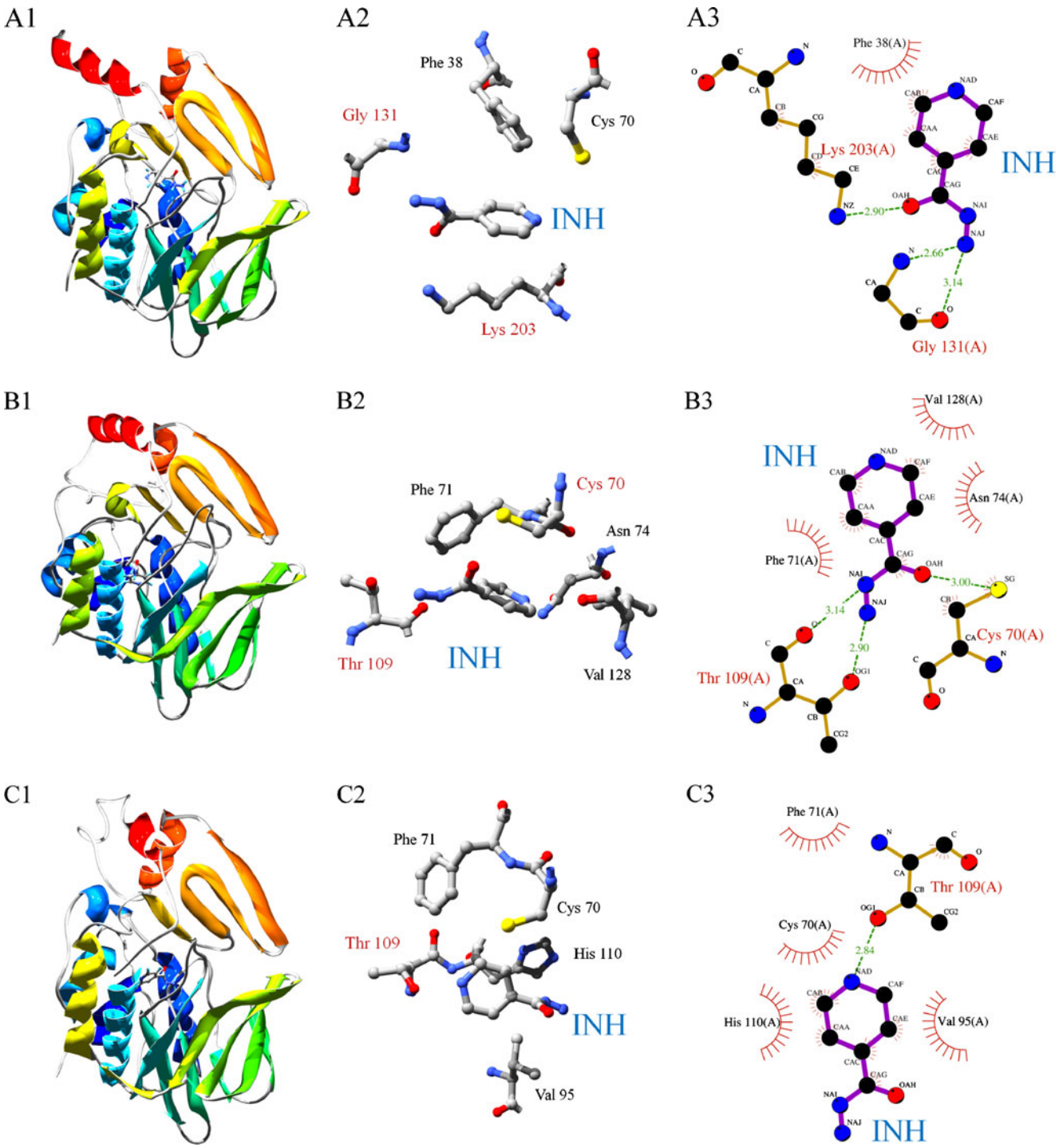

Fig. 6 Global structure of the complexes and binding mode of INH calculated for the last frame (5 ns) of MD simulations. *Left hand panels:* schematic representation of final structure of WT_**rmin** (A1), G68R_**rmax** (B1) and L125M_**emin** (C1) (TBNAT represented by ribbons and INH represented by sticks); *Central panels:* biding mode of INH in WT_**rmin** (A2), G68R_**rmax** (B2) and L125M_**emin** (C2); *Right-hand panels:* LIGPLOT diagrams for INH interaction in WT_**rmin** (A3), G68R_**rmax** (B3) and L125M_**emin** (C3)

of the MD simulations of the complexes are representative of the contacts identified during the last 2 ns of the simulations (Fig. 5). Isoniazid remained at the binding site in all simulations. Interestingly, when INH was manually docked in TBNAT taking into account the same position found for the drug in the crystal structure of the *M. smegmatis* complex [23] this resulted in the diffusion of the ligand to the bulk solution during the MD simulation (data not show). This probably means that the INH binding mode found in MSNAT crystallographic structure is not the same as that in TBNAT.

As expected, the last frame of WT_**rmin** shows hydrogen bonds between the carbonyl oxygen and the hydrazine NH2 group of INH with the ammonium group of Lys203 and the peptide NH group of Gly131, respectively (Fig. 6_A2 and A3). The only hydrophobic interaction observed at 5 ns is between the aromatic ring of INH and the side-chain of Phe38. As can be seem in Fig. 6_A2, the hydrazine moiety of INH is positioned away from the thiol group of Cys70. It is possible to argue that this binding mode represents a disadvantage to acetylation of the INH hydrazine moiety, which may be related with the low catalytic efficiency when this reaction is catalyzed by the wild type enzyme.

Like in SMNAT-isoniazid crystallographic complex, in G68R_**rmax** the hydrazine group of isoniazid makes two hydrogen bonds with Thr109 whereas the oxygen carbonyl makes one hydrogen bond with the thiol group of Cys70 (Fig. 6_B2 and B3). The hydrophobic contacts occur with the side-chains of Phe171, Ans74 and Val128. This binding mode seems to be more favorable for the acyl transfer from an acetylated form of the enzyme than that observed in WT_**rmin**.

The only hydrogen bond detected in the last frame of L125M_**emin** simulation is between the nitrogen of the pyridine ring of isoniazid and the carbonyl oxygen of Thr109 (Fig. 6_C2 and C3). This interaction is complemented by hydrophobic contacts with Cys70, Phe71, Val95 and His110. In this case, only some small arrangements would be needed to appropriately position the hydrazine group of INH to attack an eventually acetylated Cys70 amino acid residue. Considering that this binding mode is associated with the most favorable binding energy between the complexes, it is possible to argue that the TBNAT L125M isoform may play a role in the resistance of the *M. tuberculosis* to isoniazid.

A structural description for interaction of isoniazid with arylamine N-acetyltransferase from *M. tuberculosis* has not been reported yet. The only available information refers to the crystallographic complex of INH with the high homologous *M. smegmatis*. However, whereas isoniazid is an excellent substrate for MSNAT (with a reported Km value of 7.3 μM) [23] it is a very poor substrate for TBNAT (Km value of 106 mM) [59]. Thus, this study was designed to gain knowledge about the structural determinants of isoniazid interaction with TBNAT. From the data obtained in this work it was possible to obtain a description of the binding modes and binding energies related to the interaction of INH with three TBNAT isoforms.

Conclusions

In this work the interaction of three TBNAT isoforms with the antitubercular drug isoniazid was studied. The binding mode and binding energy calculated for the interaction of INH with the wild type isoform of TBNAT are consistent with experimental kinetic studies showing that this isoform is ineffective for the *N*-acetylation of isoniazid [59]. Even though the interaction energy of the best TBNAT_G68R complex is almost the same as that of the wild type complex, the spatial orientation of INH in the former is more consistent with the acetylation mechanism. Among the three isoforms, it was possible to identify that the L125M substitution leads to an enzyme with considerable more affinity to isoniazid when compared to the wild type isoform. The binding mode observed for the TBNAT_L125M isoform is consistent with the mechanism of acyl transfer from Cys70. It was also demonstrated that the estimated free energy of binding is dominated by hydrophobic contributions. The implications of these findings are potentially useful for drug design against new resistant strains of *M. tuberculosis*.

References

1. Zhang Y, Young DB (1993) Molecular mechanisms of isoniazid: a drug at the front line of tuberculosis control. Trends Microbiol 1:109–113
2. Timmins GS, Deretic V (2006) Mechanisms of action of isoniazid. Mol Microbiol 62:1220–1227
3. Balcells ME, Thomas SL, Godfrey-Faussett P, Grant AD (2006) Isoniazid preventive therapy and risk for resistant tuberculosis. Emerg Infect Dis 12:744–751
4. WHO (2011) World Health Organization. http://www.who.int/infectious-disease-report/pages/ch2text.html. Accessed 02 April 2011
5. Bhakta S, Besra GS, Upton AM, Parish T, Vernon CS, Gibson KJC, Knutton S, Gordon S, Silva RP, Anderton MC, Sim E (2004) Arylamine N-acetyltransferase is required for synthesis of mycolic acids and complex lipids in *Mycobacterium bovis* BCG and represents a novel drug target. J Exp Med 199:1191–1199
6. Sandy J, Mushtaq A, Holton SJ, Schartau P, Noble MEM, Sim E (2005) Investigation of the catalytic triad of arylamine N-acetyltransferases: essential residues required for acetyl transfer to arylamines. Biochem J 390:115–123
7. Mdluli K, Swanson J, Fischer E, Lee RE, Barry CE (1998) Mechanisms involved in the intrinsic isoniazid resistance of *Mycobacterium avium*. Mol Microbiol 27:1223–1233
8. Somoskovi A, Parsons LM, Salfinger M (2001) The molecular basis of resistance to isoniazid, rifampin, and pyrazinamide in *Mycobacterium tuberculosis*. Respir Res 2:164–168

9. Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, Collins D, Lisle G, Jacobs WR Jr (1994) inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. Science 263:227–230
10. Kelley CL, Rouse DA, Morris SL (1997) Analysis do ahpC gene mutations in isoniazid resistant clinical isolates of *Mycobacterium tuberculosis*. Antimicrob Agents Ch 41:2057–2058
11. Ramaswamy S, Musser JM (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis 79:3–29
12. Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, Crane DD, Musser JM, Barry CE (1998) Inhibition of a *Mycobacterium tuberculosis* beta-ketoacyl ACP synthase by isoniazid. Science 280:1607–1610
13. Lee ASG, Teo ASM, Wong SY (2001) Novel mutations in ndh in isoniazid-resistant *Mycobacterium tuberculosis* isolates. Antimicrob Agents Ch 45:2157–2159
14. Almeida Da Silva PE, Palomino JC (2011) Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs. J Antimicrob Chemother 66:1417–1430
15. van Rie A, Warren R, Mshanga I, Jordaan AM, van der Spuy GD, Richardson M, Simpson J, Gie RP, Enarson DA, Beyers N, van Helden PD, Victor TC (2001) Analysis for a limited number of gene codons can predict drug resistance of *Mycobacterium tuberculosis* in a high-incidence community. J Clin Microbiol 39:636–641
16. Silva MSN, Senna S, Ribeiro MO, Valim AM, Telles MA, Kritski AL, Morlock GP, Cooksey RC, Zaha A, Rossetti MLR (2003) Mutations in katG, inhA and ahpC genes of brazilian isoniazid-resistant isolates of *Mycobacterium tuberculosis*. J Clin Microbiol 41:4471–4474
17. Dalla Costa ER, Ribeiro MO, Silva MSN, Arnold LS, Rostirolla DC, Cafrune PI, Espinoza RC, Palaci M, Telles M, Ritacco V, Suffys PN, Lopes ML, Campelo CL, Miranda SS, Kremer K, Silva Almeida PE, Fonseca LS HOJL, Kritski AL, Rossetti MLR (2009) Correlations of mutations in katG, oxyR-ahpC and inhA genes and in vitro susceptibility in *Mycobacterium tuberculosis* clinical strains segregated by spoligotype families from tuberculosis prevalent countries in South America. BMC Microbiol 9:1–11
18. Coelho MB, Dalla Costa ER, Vasconcellos SEG, Linck N, Ramos RM, de Amorim HLN, Suffys PN, Santos AR, da Silva PEA, Ramos DF, Silva MSN, Rossetti MLR (2011) Sequence and structural characterization of tbnat gene in isoniazid-resistant *Mycobacterium tuberculosis*: identification of new mutations. Mutat Res 712:33–39
19. Sim E, Walters K, Boukouvala S (2008) Arylamine N-acetyltransferases: from structure to function. Drug Metab Rev 40:479–510
20. Weber WW, Hein DW (1985) N-acetylation pharmacogenetics. Pharmacol Rev 37:25–27
21. Sim E, Lack N, Wang CJ, Long H, Westwood I, Fullam E, Kawamura A (2008) Arylamine N-acetyltransferases: structural and functional implications of polymorphisms. Toxicology 254:170–183
22. Upton AM, Mushtaq A, Victor TC, Sampson SL, Sandy J, Smith DM, van Helden PD, Sim E (2001) Arylamine N-acetyltransferase of *Mycobacterium tuberculosis* is a polymorphic enzyme and a site of isoniazid metabolism. Mol Microbiol 42:309–317
23. Sandy J, Holton S, Fullam E, Sim E, Noble M (2005) Binding of the anti-tubercular drug isoniazid to the arylamine N-acetyltransferase protein from *Mycobacterium smegmatis*. Protein Sci 14:775–782
24. Payton M, Auty R, Delgoda R, Everett M, Sim E (1999) Cloning and characterization of arylamine N-acetyltransferase genes from *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: increased expression results in isoniazid resistance. J Bacteriol 181:1343–1347
25. Sholto-Douglas-Vernon C, Sandy J, Victor TC, Sim E, Helden PD (2005) Mutational and expression analysis of tbnat and its response to isoniazid. J Med Microbiol 54:1189–1197
26. Coelho MB (2008) Análise da presença de bombas de efluxo e mutações no gene nat em *Mycobacterium tuberculosis* resistentes a isoniazida. Universidade Luterana do Brasil, Dissertação de Mestrado
27. Fullam E, Westwood IM, Anderton MC, Lowe ED, Sim E, Noble ME (2008) Divergence of cofator recognition across evolution: coenzyme A binding in a prokaryotic arylamine N-acetyltransferase. J Mol Biol 375:178–191
28. Dalby A, Nourse JG, Hounshell WD, Gushurst AKI, Grier DL, Leland BA, Laufer J (1992) Description of several chemical structure file formats used by computer programs developed at Molecular Design Limited. J Chem Inf Comp Sci 32:244–255
29. Spessard GO (1998) ACD Labs/LogP dB 3.5 and ChemSketch 3.5. J Chem Inf Comput Sci 38:1250–1253
30. Thompson MA (2004) Planaria Software LLC, Seattle, WA, http://www.arguslab.com
31. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) AM1: a new general purpose quantum mechanical molecular model. J Am Chem Soc 107:3902–3909
32. Aalten DMF, Bywater R, Findlay JBC, Hendlich M, Hooft RWW, Vriend G (1996) PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J Comp Aid Molec Des 10:255–262
33. Chuang IL, Gershenfeld N, Kubinec M (1998) Experimental implementation of fast quantum searching. Phys Rev Lett 80:3408–3411
34. Breneman CM, Wiberg KB (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. J Comput Chem 11:361–373
35. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA (2004) Gaussian 03 Revision C.02. Gaussian Inc, Wallingford, CT http://www.gaussian.com/home.htm
36. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 25:1656–1676
37. van der Spoel D, Lindahl E, Hess B, van Buuren AR, Apol E, Meulenhoff PJ, Tieleman DP, Sijbers ALTM, Feenstra KA, van Drunen R, Berendsen HJC (2010) Gromacs User Manual version 4.5.4 http://www.gromacs.org
38. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Intermolecular forces. In: Pullman B (ed) Interaction models for water in relation to protein hydration. Reidel, Dordrecht, pp 331–342
39. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
40. Nosé S, Klein ML (1983) Constant pressure molecular dynamics for molecular systems. Mol Phys 50:1055–1076
41. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 52:7182–7190
42. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. J Comp Chem 18:1463–1472
43. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an Nlog-(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
44. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit 9:1–5
45. Goodsell DS (2005) Computational docking of biomolecular complexes with AutoDock. In: Golemis E, Adams P (eds) Protein-

protein interactions: a molecular cloning manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, p 885
46. Morris GM, Huey R, Olson AJ (2008) Using AutoDock for ligand-receptor docking. Curr Prot Bioinf 24:8.14.1–8.14.40
47. Sanner MF (1999) Python: a programming language for software integration and development. J Mol Graphics Mod 17:57–61
48. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. Tetrahedron 36:3219–3228
49. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function. J Comput Chem 19:1639–1662
50. Solis FJ, Wets RJB (1981) Minimization by random search techniques. Math Oper Res 6:19–30
51. Åqvist J, Medina C, Samuelsson JE (1994) A new method for predicting binding affinity in computer-aided drug design. Protein Eng 7:385–391
52. Hansson T, Marelius J, Aqvist J (1998) Ligand binding affinity prediction by linear interaction energy methods. J Comput Aided Mol Des 12:27–35
53. Wallace AC, Laskowski RA, Thornton JM (1996) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134
54. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18:2714–2723
55. Payton M, Gifford C, Schartau P, Hagemeier C, Mushtaq A, Lucas S, Pinter K, Sim E (2001) Evidence towards the role of arylamine N-acetyltransferase in *Mycobacterium smegmatis* and development of a specific antiserum against the homologous enzyme of *Mycobacterium tuberculosis*. Microbiology 147:3295–3302
56. Payton MA, Sim E (1998) Genotyping human arylamine N-acetyltransferase type 1 (NAT1): the identification of two novel allelic variants. Biochem Pharmacol 55:361–366
57. Mushtaq A, Payton M, Sim E (2002) The COOH terminus of arylamine N-acetyltransferase from *Salmonella typhimurium* controls enzymic activity. J Biol Chem 277:12175–12181
58. Delgoda R, Lian LY, Sandy J, Sim E (2003) NMR investigation of the catalytic mechanism of arylamine N-acetyltransferase from *Salmonella typhimurium*. Biochim Biophys Acta 1620:8–14
59. Sikora AL, Frankel BA, Blanchard JS (2008) Kinetic and chemical mechanism of arylamine N-acetyltransferase from *Mycobacterium tuberculosis*. Biochemistry 47:10781–9
60. Aqvist J, Medina C, Samuelsson JE (1994) A new method for predicting binding affinity in computer-aided drug design. Protein Eng 7:385–391
61. de Amorim HL, Caceres RA, Netz PA (2008) Linear interaction energy (LIE) method in lead discovery and optimization. Curr Drug Targets 9:1100–1105
62. Wang JP, Morin P, Wang W, Kollman PAJ (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. Am Chem Soc 123:5221–5230