

Article

The Structural Basis of *Mycobacterium tuberculosis* RpoB Drug-Resistant Clinical Mutations on Rifampicin Drug Binding

 Arnold Amusengeri ¹ , Asifullah Khan ^{2,*}  and Özlem Tastan Bishop ^{1,*} 
¹ Research Unit in Bioinformatics (RUBi), Department of Biochemistry and Microbiology, Rhodes University, Grahamstown 6140, South Africa; ahnoamu@gmail.com

² Department of Biochemistry, Abdul Wali Khan University Mardan (AWKUM), Mardan 23200, Pakistan

* Correspondence: asif@awkum.edu.pk (A.K.); O.TastanBishop@ru.ac.za (Ö.T.B.); Tel.: +92-333-9278471 (A.K.); +27-46-603-8072 (Ö.T.B.)

Abstract: Tuberculosis (TB), caused by the *Mycobacterium tuberculosis* infection, continues to be a leading cause of morbidity and mortality in developing countries. Resistance to the first-line anti-TB drugs, isoniazid (INH) and rifampicin (RIF), is a major drawback to effective TB treatment. Genetic mutations in the β -subunit of the DNA-directed RNA polymerase (*rpoB*) are reported to be a major reason of RIF resistance. However, the structural basis and mechanisms of these resistant mutations are insufficiently understood. In the present study, thirty drug-resistant mutants of *rpoB* were initially modeled and screened against RIF via a comparative molecular docking analysis with the wild-type (WT) model. These analyses prioritized six mutants (Asp441Val, Ser456Trp, Ser456Gln, Arg454Gln, His451Gly, and His451Pro) that showed adverse binding affinities, molecular interactions, and RIF binding hinderance properties, with respect to the WT. These mutant models were subsequently analyzed by molecular dynamics (MD) simulations. One-hundred nanosecond all-atom MD simulations, binding free energy calculations, and a dynamic residue network analysis (DRN) were employed to exhaustively assess the impact of mutations on RIF binding dynamics. Considering the global structural motions and protein–ligand binding affinities, the Asp441Val, Ser456Gln, and His454Pro mutations generally yielded detrimental effects on RIF binding. Locally, we found that the electrostatic contributions to binding, particularly by Arg454 and Glu487, might be adjusted to counteract resistance. The DRN analysis revealed that all mutations mostly distorted the communication values of the critical hubs and may, therefore, confer conformational changes in *rpoB* to perturb RIF binding. In principle, the approach combined fundamental molecular modeling tools for robust “global” and “local” level analyses of structural dynamics, making it well suited for investigating other similar drug resistance cases.

Keywords: drug resistance; mutations; rifampicin; *rpoB*; molecular dynamics simulations; dynamic residue network analysis



Citation: Amusengeri, A.; Khan, A.; Tastan Bishop, Ö. The Structural Basis of *Mycobacterium tuberculosis* RpoB Drug-Resistant Clinical Mutations on Rifampicin Drug Binding. *Molecules* **2022**, *27*, 885. <https://doi.org/10.3390/molecules27030885>

Academic Editors: Angelo Facchiano and Anna Marabotti

Received: 9 December 2021

Accepted: 25 January 2022

Published: 28 January 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Tuberculosis (TB) remains a major public health problem globally, with resistance to rifampicin (RIF) and isoniazid (INH), the main anti-TB first-line drugs, varying according to geographical location [1,2]. The number of TB patients diagnosed and treated for multidrug-resistant tuberculosis (MDR-TB) is increasing worldwide, and treatment success rates in patients with drug-resistant TB remain unacceptably low [1]. According to the World Health Organization (WHO), 9.0–11.1 million new cases and 1.1–1.3 million TB-related deaths were estimated in 2018 [1]. The global health problem due to tuberculosis has worsened with the increasing emergence of *Mycobacterium tuberculosis* complex strains that are resistant to RIF and INH. As recommended by the WHO, the timely detection of drug resistance is essential for the appropriate treatment of patients with tuberculosis and for limiting the further spread of MDR-TB [1,3]. The drug-resistant form of TB originates from

the development of different mutations [4–6]. Drug resistance poses an immense threat and presents new challenges in diagnostics, surveillance, and therapeutic methods employed in treating TB. Reports of increasing drug-resistant TB cases have emphasized the need for research to understand the basic mechanism of drug resistance [7–9].

RIF, a member of the rifamycin family of antibiotics, is a critical component of the combination therapy regimen used to treat active TB [10]. Since its approval for clinical use in late 1960s, RIF has made its mark as a core therapy for infections sustained by various bacteria, including TB (*Mycobacteria tuberculosis*), osteomyelitis (*Staphylococcus aureus*), meningococcal disease (*Neisseria meningitidis*), leprosy (*Mycobacterium leprae*), and gonorrhoea (*Neisseria gonorrhoeae*) [11]. Regarding *M. tuberculosis*, RIF exerts its effect by inhibiting DNA-dependent RNA polymerase (RNAP) [12]. RNAP is responsible for transcriptions and gene expressions in all living organisms. Bacterial RNAP is a multi-subunit complex consisting of five subunits ($\alpha_2\beta\beta'\omega$) (Figure 1A). The two largest RNAP subunits, β and β' , in coordination with Mg^{2+} ions, form the active center for the catalysis of nucleotide polymerization [13,14]. RIF binds to an adjacent pocket in the β -subunit encoded by the *rpoB* gene (Figure S1) [15,16], sterically inhibiting the extension of nascent RNA, and eventually blocking the bacterial proliferation [12].

It is estimated that 484,000 TB cases reported worldwide in 2018 were RIF-resistant [1]. More than 95% of the *M. tuberculosis* clinical strains resistant to RIF harbor mutations in the 81 base pair region of the *rpoB*, known as the RIF resistance-determining region (RRDR) (Figure 1B). The region spanning codons 433 to 458 (corresponding to 507 to 533 in the *Escherichia coli* numbering system, respectively) [7,17–19], which is a hotspot region, mainly consists of single amino acid substitutions [20]. The substitutions at codons 441, 451, and 456 (corresponding to codons 516, 526, and 531 of the *E. coli* coordinates) occur frequently and have often been used as excellent markers for the detection of resistance in RIF-resistant *M. tuberculosis* isolates [7,21–23]. Mutations outside the RRDR region in *rpoB* have also been reported in association with RIF resistance, but rarely [23–25]. Since mutations are considered critical factors associated with the acquisition of drug resistance, the elucidation of the mechanism(s) of action of the clinically important *rpoB* gene mutations that cause RIF resistance have become increasingly important in the development of novel anti-TB therapeutics. In this study, we seek to understand the structural implications of resistance-conferring mutations, occurring in the *rpoB* subunit, on RIF binding.

Previous computational studies have illustrated the structural consequences of amino acid variations in the RNA polymerase protein on drug binding affinities and drug-target stabilities [19,26–28]. The deterioration in electrostatic interactions and the rearrangement of the binding pocket to alter its shape complementarity, leading to poor binding affinities for mutations associated with positions 456 and 441 (Ser456Leu and Asp441Val), have been cited [15,19]. Likewise, substitutions at position 451 (His451Ser/Met/Glut/Asp/Tyr/Arg) induced detrimental effects on protein–ligand stability, as well as affinity [27,28]. While these studies have extensively explored thermodynamic aspects of RIF binding and have captured relevant biophysical interactions, less attention has been paid to conformation selection and intraprotein communication regulation. For instance, it is not clear whether residues within the RRDR region are key intermediaries in the intraprotein information flow, nor is it clear how mutations within this hotspot region influence the protein conformation freedom of the free energy landscape, ultimately dictating the drug binding kinetics. The systematic approach of combining molecular modeling and network analysis techniques has previously provided a versatile strategy of elucidating drug resistance mechanisms, and to a greater scope, the mechanistic ramifications of the mutations to drug binding [29–31]. Here, we screen non-synonymous single nucleotide polymorphisms (SNPs) reported in the mutation bioinformatics identification (MUBII-TB-DB) database [32], in order to unveil their influence on RIF binding in mycobacterial RNAP. We employ molecular docking, molecular dynamics (MD) simulations, and dynamic residue network (DRN) analysis techniques to prioritize the database mutations based on their impact toward RIF binding affinity. Here, the mutant–ligand stability, conformation selection, binding affinity, and

intraprotein communications are inferred with respect to the bound wild-type (WT) model to understand the structural insight and underlying mechanisms of RIF resistance.

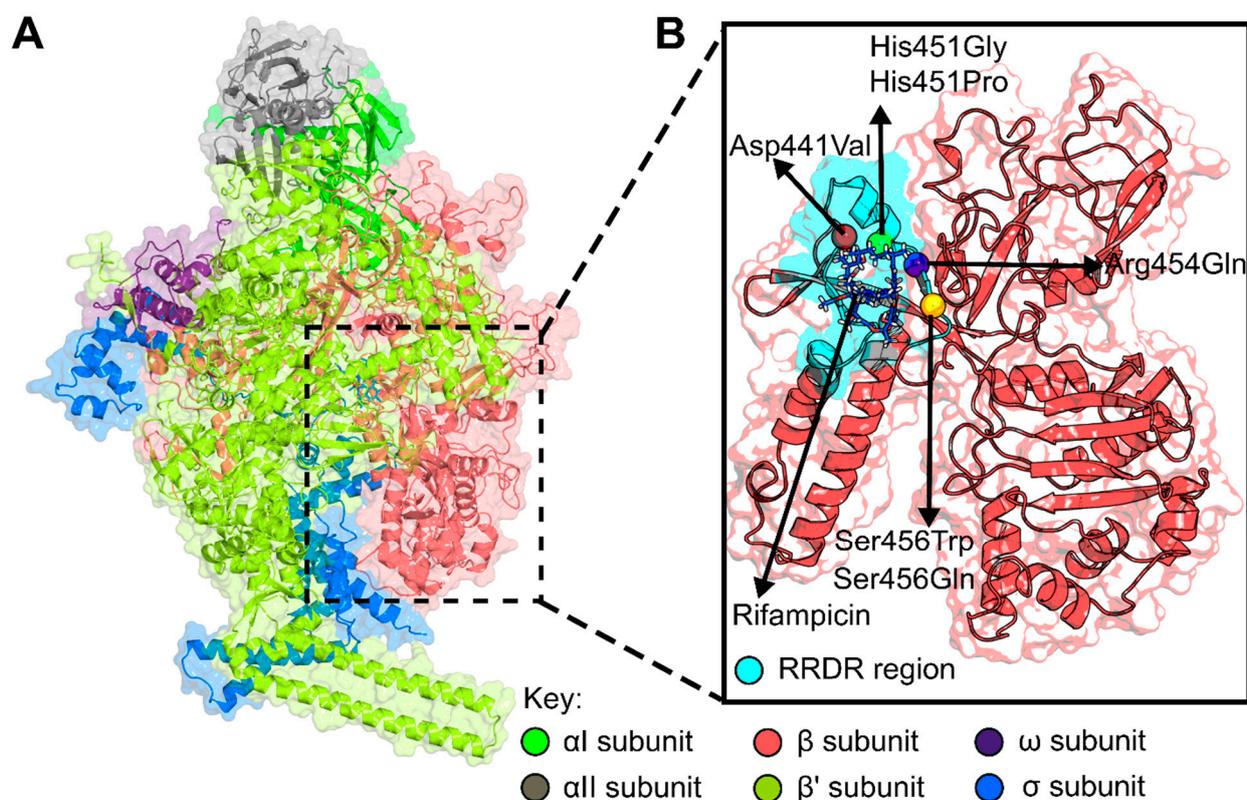


Figure 1. (A) Structural representation of the multiunit *M. Tuberculosis* RNAP catalytic core complex (consisting of α _I, α _{II}, β , β' , ω subunits) bound to an σ initiation factor to form a holoenzyme, PDB ID: 5UHC [33,34]. (B) Focused image of mutations under investigation (shown as spheres) within the *rpoB* model (residues 160 to 600). The RRDR region is mapped on the structure and colored cyan. RIF is shown as sticks and colored blue.

2. Results

2.1. Rifampicin Resistant-Associated Mutations Were Prioritized

The MUBII-TB-DB database held 133 RIF-*rpoB* drug-resistant mutations cataloged from the published literature (Supplementary_file S1, Supplementary_file S2) during the planning of this study (as of 6th September 2018). The most common RIF-resistant mutations in *M. tuberculosis* are reported in the RRDR region and are listed in the MUBII-TB-DB according to the *E. coli* genome coordinate position. Some *M. tuberculosis* strains were found to carry multiple point mutations. Therefore, the target region sequences were initially aligned to the H37Rv reference strain, as well as *E. coli*, using ClustalW in MEGA7 [35], in order to ascertain the position of target mutations [36,37]. Notably, the molecular interaction patterns governing RIF resistance and sensitivity for most mutations have previously been published in different in silico, and experimentally resolved, RIF-bound *rpoB* crystallographic structural assessments [38,39]. For this reason, we prioritized 30 mutations deposited to MUBII-TB-DB 5 years before the planning of the study, whose molecular insights, with respect to RIF resistance, have barely/not been assessed so far. The majority of these mutations were in the RRDR region of *rpoB*. Mutations occurring outside the hotspot region have been reported to cause an abrupt change in the *rpoB* structure and have been reported to influence the RIF binding efficiency [40]. About 30% of MDR-TB presented RIF-resistant mutations located outside the RRDR region [40].

The post-docking analysis revealed varied binding patterns of RIF with the WT as well as the mutated models in terms of docking scores, the formation of hydrogen bonds, and their associated distances.

Generally, the mutated complexes yielded higher binding free energy values relative to the WT rpoB-RIF complex (-13.8 kcal/mol), indicating that the mutations, regardless of their positions within the RIF binding pocket, reasonably abolish protein–drug binding (Table S1). Few mutations were found to completely abolish RIF binding, with substantially higher ligand root-mean-square deviation (RMSD) scores (ligands in the mutant model compared to ligands in the WT model) and poor docking scores. The intermolecular interactions were also investigated. The results evidently indicated that the prime difference in the binding affinity could be due to differences in binding pocket residues. The RIF-mutant interacting residues were found to be largely different from those of the native complex and, thus, could be responsible for the comparatively poor docking scores (Figure 2 and Figure S2).

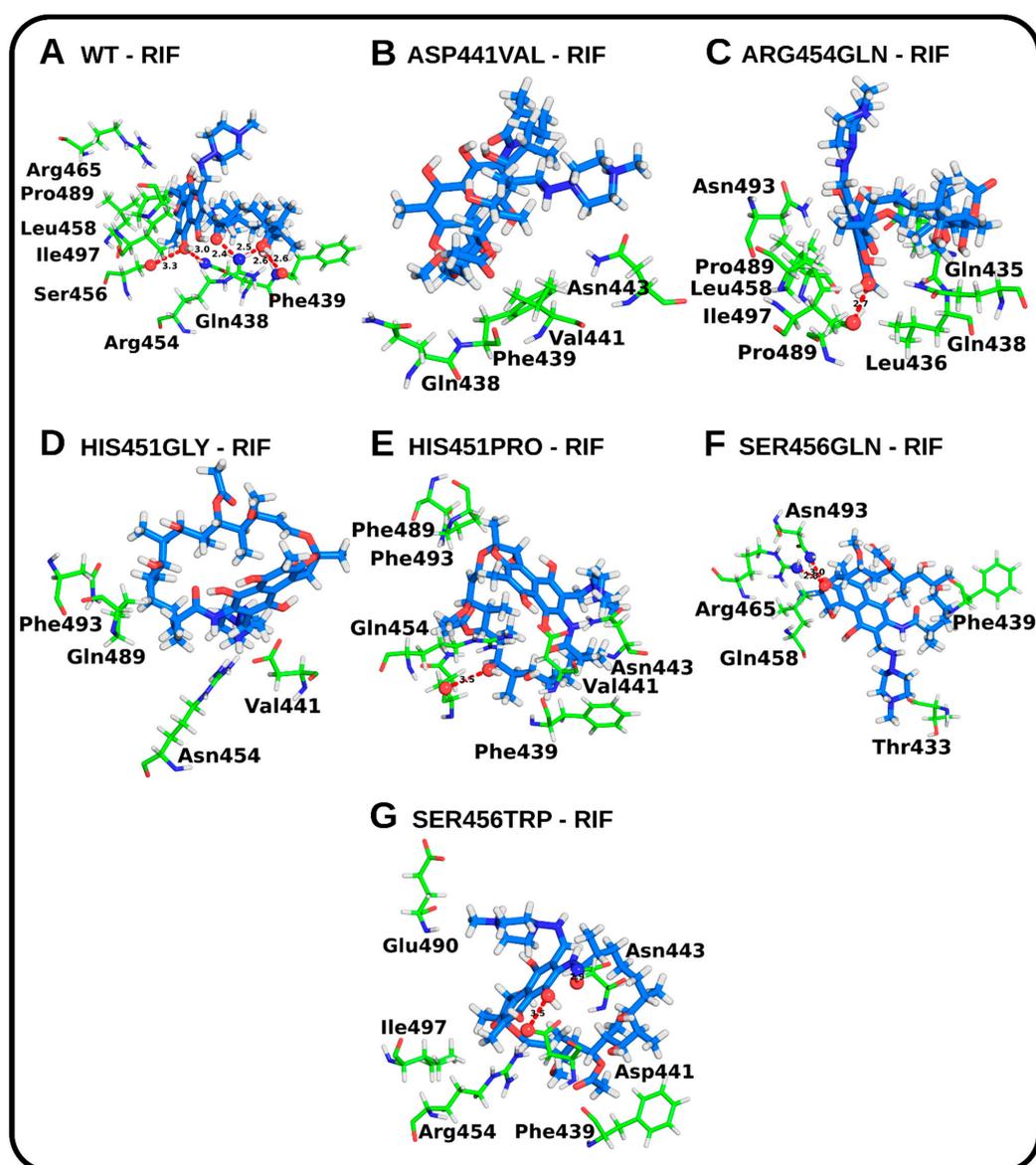


Figure 2. Post-docking analysis: The three-dimensional illustration of molecular interactions of RIF with the WT and mutated forms of *M. tuberculosis* rpoB. RIF is shown as blue sticks colored by heteroatom. Interacting atoms and associated residues are shown as spheres and sticks (colored green), respectively. Hydrogen bonds are shown as dashed red lines.

In the case of the WT rpoB protein, the RIF binding site residues, i.e., Arg454 and Phe439, formed hydrogen bonds with RIF. An amino group of Arg454 formed two hydrogen bonds with a carbonyl oxygen and hydroxyl (-OH) side chain, whereas the amide nitrogen of Phe439 formed one conventional hydrogen bond with an oxygen acceptor atom. The Asp441Val, His451Gly, His451Pro, Arg454Gln, Ser456Gln, and Ser456Trp models were found as the top mutants with respect to RIF binding hindrance and their poor ligand docking scores of -6.1 kcal/mol, -5.0 kcal/mol, -6.5 kcal/mol, -6.1 kcal/mol, -6.0 kcal/mol, and -5.4 kcal/mol, respectively, compared to the WT complex (-13.7955 kcal/mol). The visual inspection of the docked poses of the mutated rpoB complexes revealed that RIF either developed few hydrogen bond interactions with the binding sites residues or abolished them altogether (Figure 2). Likewise, high ligand RMSD scores were yielded during the superimposition of the mutant models against the WT complex (Table S1).

2.2. Mutations Variably Influenced the Protein and Ligand Stability

Six out of the thirty mutated rpoB models were identified as the top mutants possessing unfavorable protein–ligand interactions. To elucidate the impact of mutations on the dynamic behavior of ligand binding, 100ns MD simulation experiments were performed on the WT and the six mutated models in the complex with RIF. Duplicate runs were carried out for the WT (see Figures S3 and S4). MD simulations have been applied in various studies to exhaustively highlight structural mechanisms for differences in ligand bindings between the native protein and mutated targets [19,26–28,30,31]. Despite experimental evidence showing that RIF effectively binds to the WT rpoB protein at a pH of 7.4 based on physiological conditions [41,42], it remains unclear what the binding state of the drug and its associated conditions are there with the RIF-resistant mutants. Mutations are known to stimulate resistance by associating with the drugs at distinct pH levels [43–46]. It should be noted that the results discussed here illustrate the binding behavior of RIF in its neutral state, compared to the WT and the mutated proteins. To examine the general conformational divergence from the baseline structure (at zero nanoseconds), the protein and ligand RMSDs were computed. Calculations of the protein RMSD were based on the backbone atom positions. Initially, based on the RMSDs, it was found that both WT runs were in equilibrium. Moreover, the extent of divergence was small; therefore, dependable conclusions could be drawn from both runs: the models acquired comparatively similar conformational spaces considering the fair agreement in kernel density estimation plots, and the dimension projections of conformational shifts on the free energy landscape were lower (Figure S3). Protein RMSDs of both WT and the mutated models plateaued early to values in the range of 0.4 to 0.6 nm, on average; this suggests that the overall structural folds were fairly maintained (Figure S4). The average RMSD values acquired were below 0.53 nm, while the distribution median (interquartile range) recorded, in general, was 0.50 (0.09) nm (Table S2, Figure 3A,C(i)). The WT, Asp441Val, and Ser456Gln models displayed multimodal RMSD distribution patterns and the largest standard deviation (*sd*) values of 0.08 (on average), 0.1, and 0.09, respectively. This indicated that the systems were in equilibrium between multiple conformers. Conversely, the Arg454Gln, His451Gly, His451Pro, and Ser456Trp models showed normal RMSD distribution density curves and recorded the lowest *sd* values: 0.05, 0.06, 0.04, and 0.07, respectively. This suggests that a single, more stable equilibrium state was sampled.

The conformation adaptation of RIF during the simulation was assessed by measuring its RMSD. The ligands displayed system-specific conformation variations; RIF acquired normal (Arg456Gln and His451Pro), bimodal (WT (run 1 and run 2)), Ser456Gln, Ser456Trp), and trimodal (Asp441Val, His451Gly) RMSD frequency distribution density curves (Figure 3B,C(ii) and Figure S3A).

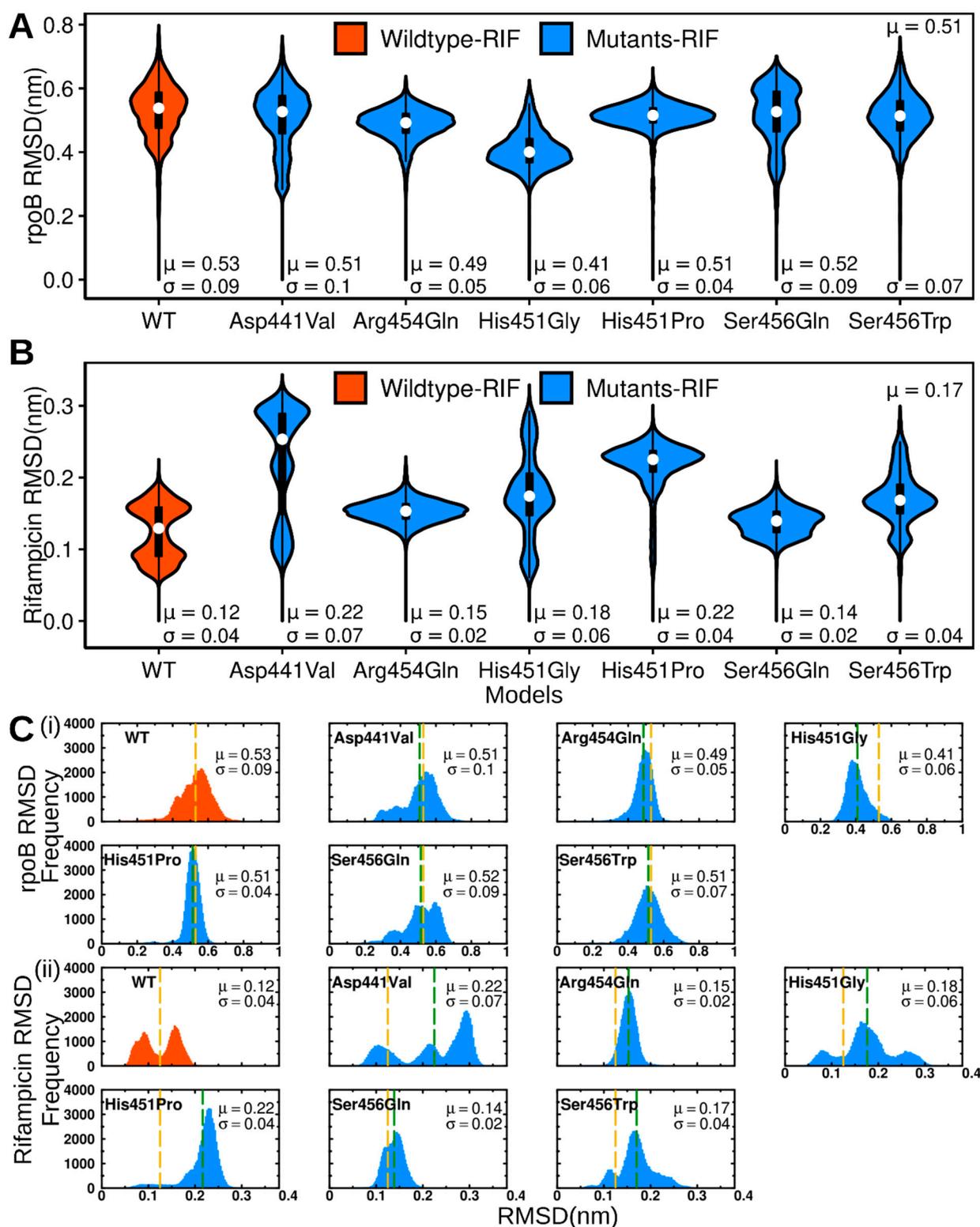


Figure 3. Violin plots illustrating the kernel density distribution frequency of (A): Protein RMSD, (B): Ligand RMSD. Color key: Red: bound WT, Blue: bound mutants. Embedded boxplots highlight the data distribution (25th (bottom) and 75th (top) percentiles). The white dot indicates the median value. (C) (i) and (ii). Histogram representation of RMSD distribution for (i) the protein, and (ii) RIF. Color key: red: bound WT, blue: bound mutants. The mean RMSD value is highlighted as dashed lines: yellow: representing the WT, and green: representing mutants.

The latter possessed the largest *sd* values (Asp441Val: 0.07 and His451Gly: 0.06, respectively) and exhibited a large RMSD variability. This suggests that the associated mutations largely destabilized the RIF's binding pose (Figure 3B and Figure S5). Taken together, (I) the WT protein accessed alternative conformers, possibly due to dual binding orientations acquired by RIF during simulation; and (II) mutated models displayed mixed (smaller and larger) conformation proportions with respect to the WT, indicating that the changes in binding site architecture can variably influence both protein and ligand stabilities.

Structural fluctuation at the residue level was monitored using root-mean-square fluctuation (RMSF) calculations, which highlight the per residue mobility in relation to its average position. Looking at the general layout of RMSF profiles, minimal differences were observed between the WT and mutated models: overall, residues 284–286, 397–418, and 595–600 were recorded with the largest fluctuation values (>0.34 nm, on average) (Figures S6 and S7). All the mutated models exhibited lower RMSF values compared to the WT, suggesting that the mutations favor the systemic relaxation.

The radius of gyration is a measure of compactness, with respect to the proteins' center of mass. Large *R_g* values correspond to a less tight packing of atoms, and vice versa. All models acquired average *R_g* values in the range of 2.58 to 2.61 nm, suggesting that minimal changes in protein compaction occurred in mutated models, with respect to the WT (Figure 4A,C(i), Figure S3B and S8). Based on *R_g* variability, the His451Gly (0.04 nm), WT, Asp441Val, His451Pro, Ser456Gln (0.03 nm each), Arg454Gln, and Ser456Trp (0.02 nm each) were found with large-to-small *sd* values, respectively. To investigate the local effects of the mutations on the packing density of residues within the RIF binding site, *R_g* calculations were performed on residues 166–181, 374–389, 430–440, 449–469, 473–475, and 488–501 (Figure 4B,C(ii) and Figure S3B). His451Pro, Ser456Gln, and Ser456Trp exhibited larger *R_g* variations (*sd* values of 0.02 each) relative to WT, Asp441Val, Arg454Gln, and His451Gly. These results suggest that: (I) Asp441Val confers insignificant effects on the protein packing density, both globally and locally (considering the RIF binding region). (II) His451Pro and Ser456Gln remain mainly intact, except for looser packing in the rifampicin binding region. (III) The His451Gly and Arg454Gln mutations majorly influence (increase and decrease, respectively) the spatial packing of the entire structure. (IV) While structural compaction is enhanced globally in Ser456Trp, structural expansion is favored in the RIF binding region. The above results clearly show that mutations can variably influence conformational compaction/expansion and, likewise, alter the shape of the complementarity aspects with respect to the ligand binding site.

2.3. Mutations Favor Conformational Rigidity

The investigation of conformation redistribution, due to mutation/ligand binding, is important for understanding resistance mechanisms. Essential dynamics were carried out in order to examine how the single point mutations reorganize the conformation populations in a reduced dimensional subspace. Often, principal components are sorted in such a way that the top few retain the most variation; the first two (PC1 and PC2) typically describe dominant systemic motions [47]. In general, PC1 and PC2 (from a total of 3659 eigenvectors) captured more than 40% of the total variance (Table S3). Two-dimensional projections suggested that the intrinsic motions were highly restricted, and the conformational subspaces were relatively limited (Figure S9). Relative to the WT, most mutants covered a restricted spatial area, suggesting an overall decrease in the accessible degrees-of-freedom. The projections of the proteins' free energy landscape (FEL), as a function of significant (slow) modes, provided a robust representation of dominating conformation ensembles and allowed a clear visual interpretation of the folding pathways. The intermediate basins (labelled "t") were best described as transitional conformations equivalent to metastable states that lead to the bottom of the folding funnel (the native state, labelled "C") (Figure 5) [48,49]. The WT explores two global-wide energy minimas (C1 and C2) separated by low lying energy barriers. This suggests that rpoB natively adopts two conformations during the RIF binding, as previously highlighted in RMSD calculations.

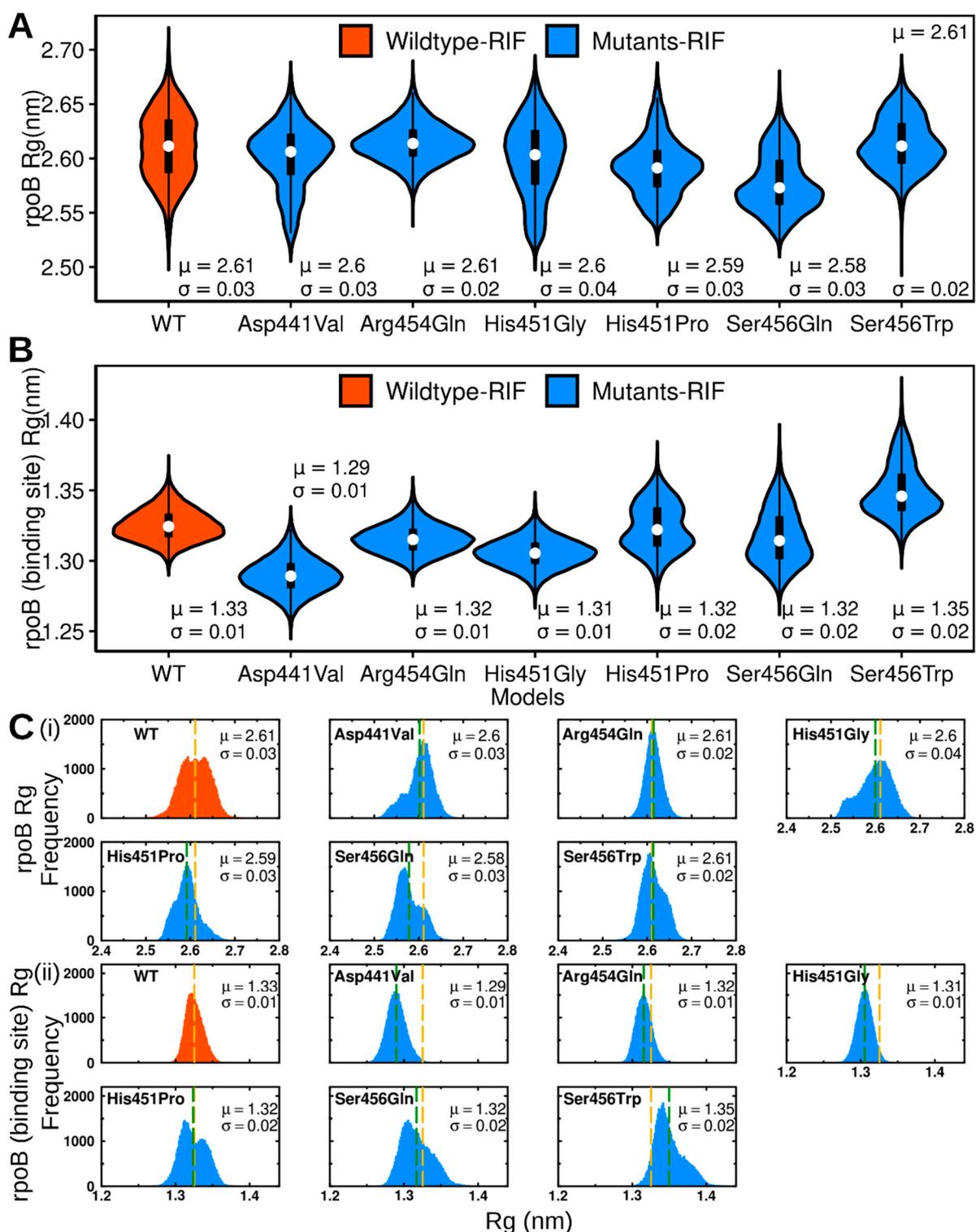


Figure 4. Violin plots illustrating the kernel density distribution frequency of the radius of gyration for (A): The entire protein structure, (B): residues within the RIF binding site. Color key: red: bound WT, blue: bound mutants. The boxplots within highlight the data distribution (25th (bottom) and 75th (top) percentiles). The white dot indicates the median value. (C) Histogram representation of R_g distribution for (i) the entire protein and (ii) RIF binding site. Color key: red: bound WT, blue: bound mutants. Dashed lines represent the mean R_g value: yellow: bound WT, green: bound mutants.

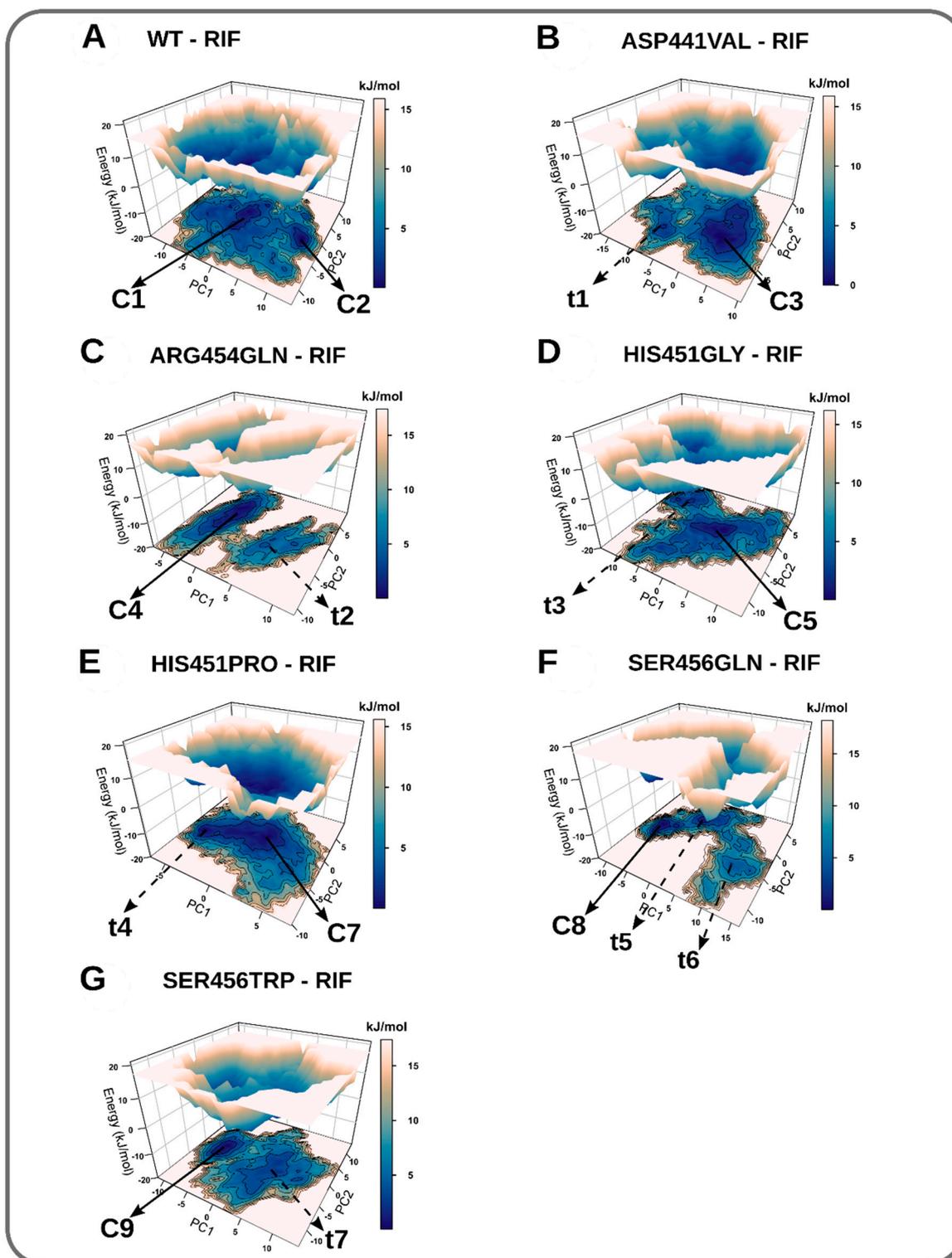


Figure 5. Protein free energy landscapes computed as a function of PC1 and PC2 for WT and mutated systems. Color ranges from white (energy maxima) to blue (energy minima). Labels: c: conformers, t: transition substates.

All mutants adopted single global free energy minima; however, they were trapped transiently in local energy minima wells towards the bottom of the funnel. Asp441Val, Arg454Gln, His451Gly, His451Pro, and Ser456Trp each sampled a single metastable substate (t1, t2, t3, t4, and t7, respectively), whereas Ser456Gln visited two metastable substates

(t5 and t6). These observations agree with the RMSD kernel density estimations plots (see Figure 3), which appropriately highlighted that Ser456Gln acquired large conformational variability. Lower trace values were observed for all the mutants (when compared to the WT), suggesting that the concerted atomic displacements yielded, due to mutations, were generally reduced (Table S3). Thus, the mutations confer structural rigidity. The projections of concerted motions using porcupine plots indicated that the major motions in the mutated models were dominated by the distant loop region (outside the RIF binding site) between residues 397–418. This region also exhibited large residual fluctuations, as described in the RMSF calculations (Figure S10). Globally, the WT model displayed larger atomic motions in general, whereas the mutated models showed reduced motions. The comparisons of the structures occupying dominant energy wells (conformers of the WT, versus the mutants) revealed large structural displacements across all mutants in the entire model following the superposition with the WT (Figure S11). Relative to the WT, mutated models recorded RMSD values in the range of 0.801 to 0.904 nm, emphasizing that the mutations induced considerable structural rearrangements (Table S4). In summary, mutations generally discouraged large atomic shifts relative to the WT, besides sizably restricting the conformational freedom of rpoB during the RIF binding. These observations could be considered a linchpin in the impairment of drug recognition in all mutants.

2.4. The Impact of Mutations on Protein-Ligand Binding Affinity

Intermolecular hydrogen bond numbers (H-bonds) play a key role in stabilizing protein–ligand complexes. Compared to the WT (three H-bonds), Arg454Gln (four H-bonds), and Ser456Trp (four H-bonds), which generally recorded higher intermolecular hydrogen numbers, Asp441Val (two H-bonds), His451Gly (two H-bonds), His451Pro (two H-bonds), and Ser456Gln (one H-bond) yielded lower numbers, on average (Figure S12). These results suggest that the latter's mutations largely destabilized the RIF–rpoB interactions. Protein–ligand interactions were examined by considering representative structures occupying energy minimas (dominant conformers) (Figure S13). Comparable to the WT, Arg454Gln, Ser456Trp, and Arg454Gln generally formed higher (four or more) conventional hydrogen bonds with RIF. The WT residues Arg465, Ser456, Gln435, Gln438, and Phe439 formed conventional hydrogen bonds with RIF. Notably, interactions with residue Phe439 were present across all mutations, except His451Gly and His451Pro. Both models (His451Gly and His451Pro) displayed diminished molecular interactions in general, suggesting that the mutations at position 451 conferred large detrimental effects on binding contacts.

Single point mutations occurring at ligand binding sites have been shown, in the past, to influence the binding thermodynamics of ligands [30,50]. Binding free energy (BFE) calculations provide a quantitative overview of drug–receptor affinities. MM/PBSA functions have been employed in previous works to estimate the total relative binding free energy, as well as the contribution of individual residues to binding through the per residue decomposition method, with promising accuracy [30,51–53]. Table 1 displays a summary of the results obtained. In general, the van der Waals interaction (ΔE_{vdW}) is the driving force for the RIF–rpoB interaction, whereas polar energy (ΔG_{polar}) is detrimental to binding. Relative to the WT (approximately -128.6 kJ/mol on average), RIF bound, with an increased affinity, to Arg454Gln (approximately -213.4 kJ/mol), His451Gly (approximately -138.2 kJ/mol), and Ser456Trp (approximately -216.3 kJ/mol) mutants, while its binding to Asp441Val (approximately -25.0 kJ/mol), His451Pro (approximately -90.2 kJ/mol), and Ser456Gln (approximately -71.5 kJ/mol) was not favored. While the computed total binding free energy values were expected to be nearly similar, with respect to the duplicate runs of the same system, the WT run 1 and run 2 recorded large variations. The electrostatic energy term (ΔE_{elec}) was the predominant factor dictating the observed fluctuations. This was not only evident within the duplicate runs of the WT protein, but also across the mutants. It should be pointed out that resultant binding free energy values are as accurate as the scoring functions. Additionally, MMPBSA computed the relative binding free energy,

which does not equal the real binding free energy, since we did not estimate the entropy contribution to binding. Despite these shortcomings, both WT runs showed remarkable consistencies in their per residue energy decomposition profiles (see Figure S14), giving us confidence in the obtained results. The residue decomposition of the BFE revealed three key residues in the WT model: Leu458, Pro489, and Ile497, contributing substantially to favorable ligand binding (Figure S14, Table S5). Interestingly, mutations arbitrarily disrupted the per residue thermodynamic profiles across all mutated models (Figure S14). Moreover, it is likely that the mutations induced far-reaching allosteric effects that strongly influenced the thermodynamic balance of binding. As such, the models yielded mutant-specific patterns of residues that contributed substantially to the total BFE (Figure S14, Table S5). The interactions with residue 454 (Arg454) did not favor ligand binding across all models (including WT), except the Asp441Val and Arg454Gln mutants. Markedly, the interactions with Glu487 that were identified as detrimental to binding in the WT system, contrarily, yielded favorable binding energy values in all mutants except Asp441Val.

Table 1. Summary of binding free energy values acquired for each protein–ligand complex.

Systems	ΔE_{vdW}	ΔE_{elec}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{binding}$ (kJmol ⁻¹)
WT-RIF run1	−195.7 ± 0.4	−34.6 ± 0.3	140.1 ± 0.2	−20.8 ± 0.0	−111.1 ± 0.5
WT-RIF run2	−218.3 ± 0.3	−62.6 ± 0.3	155.6 ± 0.9	−20.9 ± 0.0	−146.2 ± 0.9
Asp441Val-RIF	−190.1 ± 0.5	77.7 ± 0.5	106.8 ± 1.4	−19.4 ± 0.0	−25.0 ± 1.4
Arg454Gln-RIF	−210.9 ± 0.5	−204.5 ± 1.3	223.5 ± 1.7	−21.5 ± 0.0	−213.4 ± 0.6
His451Gly-RIF	−144.8 ± 0.6	−79.7 ± 0.9	101.7 ± 1.2	−15.4 ± 0.0	−138.2 ± 1.1
His451Pro-RIF	−160.4 ± 0.4	−59.7 ± 0.9	148.3 ± 1.1	−18.5 ± 0.0	−90.2 ± 0.8
Ser456Gln-RIF	−155.0 ± 0.3	11.7 ± 0.6	89.5 ± 0.7	−17.6 ± 0.0	−71.5 ± 0.7
Ser456Trp-RIF	−182.1 ± 0.3	−251.1 ± 1.2	236.9 ± 1.4	−19.9 ± 0.0	−216.3 ± 0.9

Previous computational studies on systems with mutations at positions 456 and 441 (i.e., Ser456Leu and Asp441Val) yielded concordant results with those of Ser456Gln and Asp441Val, considering the total binding free energy. Relative to the WT, Ser456Leu and Asp441Val exhibited a decrease in their total binding affinity in that order, which could be attributed to weakened electrostatic interactions [19]. Particularly, the weakened electrostatic interactions with Arg454 were highlighted as key to the development of resistance. Zhang et al. reported weakened binding affinities for complexes with various residue substitutions at position 451 (His451Asp, His451Tyr, H451Arg) [28]. However, our models yielded mixed results for this position. His451Gly (−138.2 kJ/mol) and His451Pro (−90.2 kJ/mol) were recorded with increased and decreased binding affinities, respectively. Taken together, the Arg454Gln and Ser456Trp mutations leveraged the drug–receptor complex stability, as demonstrated by higher hydrogen bond numbers larger binding affinity values, while the Asp441Val, His451Pro, and Ser456Gln mutations generally did not favored the drug binding and interactions.

2.5. Mutations Conferred Common Disparaging Effects on Critical Communication Hubs

Graphs of the residue interaction network have been widely employed in the identification of key amino acids involved in the regulation of protein functions [30,52–54]. Analyses of network graphs were conducted based on the average shortest path (L) and betweenness centrality (BC) metrics. When computed over a trajectory, the average L highlights how easily accessible a residue is for communication. In principle, regions with low average L values are better disposed to transmit information. Generally, both the WT and mutated models exhibited minimal differences in regard to residues possessing low average L values. The $1Xsd$ from the mean value was used as a threshold to identify low average L regions (dips) (Figure S15). Considering the WT model (native representation), the residues 177–188,

203–208, 213–214, 309–320, 338, 342, 369–385, 437, 455–458, 463, 474–484, 499, 505–514, and 532–535 yielded low average L values in general, signifying highly accessible communication hubs. Relative to the WT, all mutated models were found with lower threshold values, indicating an overall decrease in average shortest path lengths (Figure S15).

BC provides a hierarchical arrangement of highly influential residues in a social network by considering their ratios of participation in all pairs of the shortest paths. Residues possessing high average BC indexes are associated with signal propagation and are deemed critical intermediaries for intraprotein communication. Generally, both WT and the mutated models displayed largely similar average BC profiles and were strongly correlated, as exhibited by high pairwise Pearson's correlation values (Table S6, Figure 6A,C and Figure S16). The $2Xsd$ value was used to identify residues that yielded large average BC values (peaks) with respect to the WT model (the native representation). Generally, most residues that yielded high BC values were located within the ligand binding region (for instance, residues 179–182, 374–380, 384–387, and 456), indicating their importance in achieving native protein–drug interactions (Table S7). Among these, the residue 456 was placed within the RRDR [23,55,56], suggesting that the residue plays a key role in achieving standard drug potency. Moreover, residue 376, located outside the RRDR region, has previously been shown to confer drug resistance upon its mutation [23]. Interestingly, changes in average BC values, due to mutations, mostly occurred in these signaling hubs across all mutated models in general, including residues 181–182, 204–206, 375–380, 384–386, and 512, as shown in Figure 6B and Table S8. This is indicative of the common derogatory effects of mutations on critical communication hubs and could serve as molecular signatures distinguishing drug-susceptible WTs from drug-resistant mutants [30]. Residues unique to each model, with respect to our dataset, were also identified, including: 179, 379 (Asp441Val); 370, 379, 477, and 478 (Ser456Trp); 365, 366, 510, 539, and 540 (Arg454Gln); and 387, 390, 187, 209, 371, 388, 391, 392, and 461 (His451Gly) (Table S8). These residues might be useful in explaining the differences in phenotypic characteristics, or, perhaps, the severity of resistance exhibited by each mutant. Despite the pairwise Pearson's correlation showing the strong correlated arrangement of the network layout among the models (r range: 0.80–0.94) with respect to the average BC (Figure 6C, Table S6), the comparisons suggested that the mutated models might be distant from the WT in the order of His451Pro > Ser456Trp > Asp441Val > Arg454Gln > Ser456Gln > His451Gly > WT.

In summary, the mutations investigated here altered the native information flow network, characterizing the drug-susceptible WT system by increasing and decreasing the overall reachability and connectivity, respectively, of most residues charged with adjudicating intraprotein communications in their bound state. While these residues may not be directly associated with the RRDR region, their proximity to this area, as well as residues within the RIF binding pocket, could explain the associated detrimental impact of mutations on ligand binding.

Residue contact maps have been used in the past to analyze differences in residue interactions and as a tool for general protein structure comparisons [57,58]. We employed contact maps to scan through the trajectories and to identify shared and unique contacts between the WT and mutated models. Pairwise comparisons of the WT and mutants were carried out with respect to the mutated residue positions. Generally, the WT and all mutants were glued by numerous mutual molecular contacts (Figure 6, Figure S17). We hypothesized that the differential contacts, due to mutations, could be responsible for modulating the local thermodynamic changes and, hence, RIF stability. Specifically, we focused our analysis on the lost/emerging contacts that contributed substantially to the total binding free energy. Compared to the contacts exhibited in the WT model, residues Thr450 (Asp441Val), Gln438 (Arg454Gln), Thr380, Val381, Leu455; Ser177, Val175 (Ser456Gln), Gln435 (Ser456Trp), Phe439, and Gln442 (His451Pro) either emerged or lost contacts due to their associated mutations. Among these, Gln438, Gln435, Leu436, and Phe439 directly contributed to the total binding free energy in the WT (See Table S5). Other residues,

including Thr450, Leu455, Val175, and Ser177, as well as the neighbor residues His451, Arg454, Arg173, and Val176, were also charged with maintaining protein–ligand affinities. Based on the per residue binding free energy contribution values, these interactions might be categorized either as stabilizing (associated with positive energy values) or destabilizing (associated negative energy values).

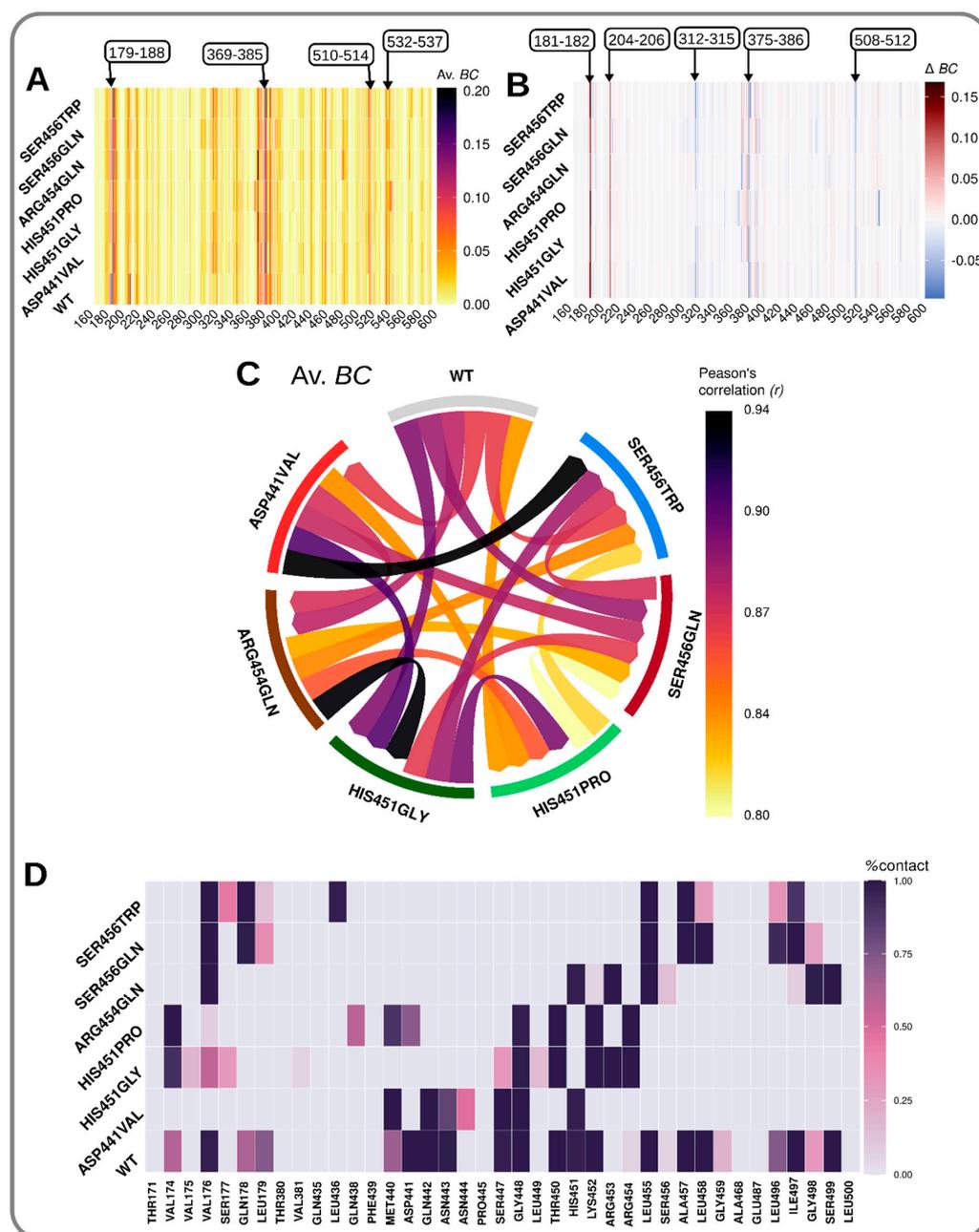


Figure 6. Analysis of average BC results: (A) Per residue average BC values illustrated using a heatmap. (B) The residue changes in average BC index (ΔBC) because of mutations (WT model and less mutated model). (C) Pairwise Pearson's correlation among models with respect to average BC profiles: comparison of linear relationships between the WT and the six mutated models in this dataset. The color ranges from yellow: the smallest correlation value, to black: the largest correlation value. (D) Per residue weighted contact heatmap. Color range is from white (0% frequency of occurrence) to violet (100% frequency of occurrence). Per residue contact network map representation is shown in Figure S17.

In summary, the above results provide evidence that the mutations confer resistance by losing/making novel contacts, particularly with residues participating in protein–ligand binding. While how the degree of association (weights) contributes to inefficiency in binding may not yet be established, it will certainly be exciting to investigate the consequences of fine-tuning these associations in order to counteract resistance.

3. Materials and Methods

3.1. Data Retrieval

3.1.1. Protein Preparation

The crystal structure of the DNA-directed rpoB chain was retrieved from the protein data bank (PDB) under the ID, 5UHC [33]. Before docking, the structure was checked for missing atoms, bonds, and contacts. The energy of the retrieved protein molecule was minimized after 3D-protonation using the default parameters of the MOE energy minimization algorithm (gradient: 0.05, force field: MMFF94X) [59].

3.1.2. Ligand Preparation

The ligand RIF used in this study against rpoB was retrieved from the crystal structure of the complex rpoB protein. The ligand was energy minimized using the default parameters of the MOE energy minimization algorithm (gradient: 0.05, force field: MMFF94X) [59]. The minimized molecule was saved in the mdb file format as a ligand input file.

3.1.3. The Retrieval of *Mycobacterium tuberculosis* rpoB Clinical Mutations

The MUBII-TB-DB repository holds *M. tuberculosis* mutations at seven loci associated with its resistance to first- and second-line antibiotics [32]. The database includes 133 RIF-rpoB drug-resistant mutations cataloged from the published studies (Supplementary_file S1, Supplementary_file S2). The repository holds mutations in the hotspot region (RRDR), as well as outside the hotspot region, within the rpoB [60,61]. The drug-resistant mutations, which have previously been evaluated with respect to RIF binding in an experimental layout, as well as via molecular docking and MD simulation approaches, were excluded.

3.2. Mutant Model Generation and RMSD Calculation

The 3D structures of rpoB mutant models were constructed using MOE software (MOE-2016). The established rpoB protein structure deposited in PDB under the ID, 5UHC, was used as template. All the mutants were 3D-protonated and energy-minimized using the default parameters of the MOE. The native (WT) structure was superimposed on the various mutant models and the root-mean-square deviation values (RMSD) of each pair (WT against mutant) were calculated using the MOE. The docked complexes were analyzed and the interactions between RIF and WT, as well as mutant proteins, were visualized using PyMOL [62].

3.3. Molecular Docking of Wild-Type and Mutant Protein Structures with RIF

Molecular docking simulations of RIF, into the binding pocket of the WT and the mutated rpoB models, were performed using the MOE [59]. The triangle matcher docking algorithm was implemented. All the ions were deleted, and the hydrogen atoms were added to the protein by 3D-protonation using the MOE software. The models were then energy-minimized using the default MOE parameters. The structure of the RIF compound was built in the MOE and was energy-minimized using the MMFF94x force field with a gradient value of 0.05. The inhibitor was re-docked into the RIF binding site of the target enzymes while applying the default parameters of the MOE (placement: triangle matcher, rescoring function: London dG) [59]. Ten conformations were generated for the ligand. The selection of the best docking solutions for further analyses were based on consensus-ranking by considering the following criteria: (I) How solid the affinity of RIF and rpoB was, while looking at the binding free energy values, and (II) how similar the

docked conformation of rifampicin in mutants, relative to that in the WT, was by looking at the RMSD deviation values.

3.4. All Atom Molecular Dynamics Simulations of the *rpoB*–RIF Complexes

The top effective mutants were prioritized for downstream MD simulation analyses based on the docking scores, RMSD calculations, and receptor–ligand chemical interactions perturbation. Molecular dynamics simulations were performed using GROMACS 2016.4 [63]. The AMBER03 force field was employed [64]. ACPYPE software was used to assign appropriate atomic partial charges and to generate RIF topologies. ACPYPE combines the general amber force field (GAFF) method and the semi-empirical quantum chemistry program, SQM [65]. Default options were applied. Mg^{2+} ions were excluded from all simulation systems [66]. A cubic simulation box with 2 Å space from the edge of the complex was defined and filled with TIP3 (SPC126) explicit water molecules. System charges were neutralized using a 0.15 M NaCl concentration. The final systems were composed of approximately 200,000 particles. While applying the steepest descent algorithm, the models were energy-minimized up to a gradient limit of ≤ 1000 kJ/mol/nm. One hundred thousand cycles of NVT equilibration, at 300 K as the initial temperature, and NPT equilibration at 1 bar of pressure, ensued. Temperature and pressure coupling at the 0.1 ps and 2.0 ps time constants were managed using the Berendsen thermostat and the Parrinello–Rahman barostat, respectively. All bonds were constrained using the LINCS solver. Non-bonded interactions were allowed a maximum distance of 1.4 nm. Long-range electrostatic interactions were treated using the particle mesh Ewald method, while the van der Waals interactions were evaluated using the Lennard Jones potential. Finally, 100 ns production runs (with a 2 fs timestep) were performed while applying periodic boundary conditions in all directions. Trajectory snapshots were written out every 2 ps.

3.5. Preliminary Trajectory Analysis

Systemic conditions including temperature, potential energy, kinetic energy, and pressure were initially checked to assess the quality of experiments (data not shown). Trajectories were analyzed using the GROMACS built-in utilities. *gmx rms*, *gmx rmsf*, *gmx gyration*, and *gmx hbond* were employed to calculate the RMSD, RMSF, Rg, and intermolecular hydrogen bond numbers, respectively.

3.6. Essential Dynamics

The principal component analysis (PCA) is a dimension reduction method that has been widely used to characterize important collective motions of proteins [30,52,53]. A detailed theoretical description of PCA is given by Amadei et al. [67]. The *gmx_covar* module of GROMACS was used to pre-align snapshots from the entire trajectory length and immediately after, generating the covariance matrix (C) associated with the C- α atomic positions. The C is a $(3 \times 3)m$ matrix where m is the number of residues considered. The Eigenvalue decomposition of the C-matrix and the pre-sorting from the highest to the lowest variance was carried out using the *gmx_anaeig* tool. Typically, biologically relevant motions can be captured by the combination of top orthogonal modes/principal components (PCs) [68]. Once the functionally relevant motions have been captured, the converged equilibrium probabilities ($\Delta G(X)$) equivalent to the free energy landscape can be expressed as a function of the Boltzmann constant (K_B) and the absolute temperature (T):

$$\Delta G(X) = -K_B T \ln P(X) \quad (1)$$

where $P(X)$ is the probability distribution of the protein conformations along the PCs.

3.7. Binding Free Energy Computations

The effects of mutations on the protein–ligand binding affinity were determined by calculations of binding free energy. The binding free energy computations were carried out on 7500 trajectory snapshots spanning the equilibrated trajectory phase of 85 ns to 100 ns

and were sampled at 10 ps time intervals by following the molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) method [69]. The *g_mmpbsa* tool of GROMACS was employed [64]. Briefly, the total binding free energy $\Delta G_{binding}$ of a protein–ligand complex can be computed as follows:

$$\Delta G_{binding} = \Delta G_{complex} - (\Delta G_{rpoB} - \Delta G_{rifampicin}) \quad (2)$$

$$\Delta G_x \approx Emm - (TS + G_{solv}) \quad (3)$$

where $\Delta G_{complex}$, ΔG_{rpoB} , and $\Delta G_{rifampicin}$ denote the Gibbs free energy values of the complex (bound protein), the receptor (free protein), and the ligand, respectively. The *g_mmpbsa* calculates the relative binding free energy; the Gibbs free energy of each component (ΔG_x) is equivalent to the sum of molecular mechanical energy (*Emm*), the solvent accessible area (G_{solv}), and the entropic potential (*TS*). The module utilizes the single trajectory approach, which follows the assumption that the receptor and the ligand endpoint states (bound and free) occupy identical conformations.

3.8. Dynamic Residue Network Analysis

The dynamic residue interaction network was performed using MD-TASK [70]. MD-TASK employs the NetworkX algorithm to construct residue interaction network (RIN) graphs. In RIN, amino acid residues represent nodes (in this case, the C β atom of amino acids (C α for Glycine)), and the non-covalent interactions defined by a threshold distance of ≤ 6.7 Å between residues represents the existing links/edges [71]. The protein RIN was assessed based on the average shortest path (*L*) and the betweenness centrality (*BC*) metrics. *L* highlights the accessibility of a residue for communication. The average shortest path can be constructed using the following equation:

$$L = \sum_{i,j \in M} \frac{d(i,j)}{n(n-1)} \quad (4)$$

where *M* is the set of nodes (residues) in the protein, and *i* and *j* are elements of set *M*. $d(i,j)$ represents the shortest path from residue *i* to *j*, whereas *n* is the total number of residues in the system. *BC* resolves the degree of connectivity of a residue by looking at its magnitude of participation along the shortest communication paths constructed by all residues, hence its importance as an intermediary for communication. The *BC* index of a residue *m* in a network can be computed as follows:

$$BC(m) = \sum_{i,j \in M} \frac{\sigma(i,j|m)}{\sigma(i,j)} \quad (5)$$

where $\sigma(i,j)$ is the total number of the different shortest paths between *i* and *j*, and $\sigma(i,j|m)$ is the count of these *i-j* paths that pass through residue *m*. Calculations were performed on 7500 snapshots sampled at 10 ps intervals spanning the equilibrated last 15 ns of the trajectories. The overall accessibility and connectivity indices were represented by the average value, namely, the average *L* and the average *BC*. In order to make comparisons, the values were initially normalized, as described before [52]. A weighted residue contact map analysis was performed using the MD-TASK webserver (unpublished). Likewise, covalent contacts between a target amino acid and a neighboring residue were considered to be formed if the residue was placed within a radius distance of 6.7 Å or less.

4. Conclusions

Tuberculosis is a major health problem that is further exacerbated by the emergence of MDR-TB. Mutations in the chemotherapeutic target, *rpoB*, play important roles in determining the cellular sensitivity and resistance to RIF, the associated first-line anti-TB drug. This work presents a comprehensive examination of molecular mechanisms underlying

RIF resistance in selected mutants. At first, six mutated models were prioritized following comparative docking simulations. Relative to the WT, the mutated models Asp441Val, Ser456Trp, Ser456Gln, Arg454Gln, His451Gly, and His451Pro yielded the least favorable protein–ligand binding energies, molecular interactions, and RIF binding pose similarity characteristics. Next, MD simulations, coupled with BFE calculations and DRN, were performed on the seven (six mutants and one WT) protein–ligand complexes to unravel the dynamic effects of mutations on RIF binding. The following testable assertions were demonstrated: (I) Asp441Val and His451Gly largely destabilized the binding orientation of RIF; (II) His451Gly favored structural expansion, whereas Arg454Gln and Ser456Trp favored structural compaction; (III) all mutations induced opposing concerted atomic motions with respect to the WT, while altogether restraining the conformational freedom, where the most significant conformational changes were represented by the motions of loop 397–418; (IV) Asp441Val, Ser456Gln, and His451Pro were detrimental to RIF binding affinity; and (V) all mutations conferred common derogatory effects on critical intraprotein communication hubs, whose impact on ligand binding could be linked to the immediate proximity to the RRDR region and the RIF binding pocket. Lastly, we should consider the contribution of electrostatic interactions for their role in stabilizing protein–ligand affinities and, generally, the molecular interactions with Arg454 and Glu487 might be modified in order to re-sensitize the resistant mutants.

Finally, the experimental verification of the effects of the mutants mentioned here is crucial, and we hope that this study will inspire wet-lab investigation.

Supplementary Materials: The data presented in this study are available in supplementary material. Table S1: Docking scores, RMSD, and post-docking interactions (hydrogen bonds) of rifampicin (RIF) in WT and mutated models of rpoB. Figure S1: A: Graphical representation of the crystal structure of *Mycobacterium tuberculosis* transcription initiation complex, PDB ID: 5UHC [1], containing 3nt RNA in a complex with rifampicin. The structure is superimposed to the wild-type and mutated models. Color code: green: multiunit RNA polymerase complex; brown: 3nt RNA strand; yellow: modeled wild-type structure of the β -subunit; grey: modeled mutated structures of the β -subunit. Rifampicin ligands docked to the structures are shown in sticks. B: 5UHC structure without the RNA strand: Depiction of the rifampicin binding pocket adjacent to the active center [2,3]. C: Zoomed-in image of the rifampicin binding pocket depicting all residues within 3.7 Å distance from the ligand atoms. Color code: red: co-crystallized rifampicin bound to 5UHC; blue: rifampicin docked to the wild-type protein. Figure S2: Post-docking analysis: 2D binding modes of RIF with the WT and mutated forms of *M. tuberculosis* rpoB. Molecular interactions were visualized using Discovery Studio Visualizer 4 (DS Visualizer) [72]. Figure S3: Post MD results: Preliminary assessment of the wild-type model runs 1 and 2. (A) The protein and ligand RMSDs represented as time evolution (i), kernel density estimation violin plots (ii), and frequency histograms. (B) The radius of gyration calculated for the entire protein, as well as residues within rifampicin binding pocket. Figure S4: 100ns protein RMSD evolution computed with respect to backbone atom positions. Color key: black: WT, red: mutants. Table S2: tabulated summary of average RMSD values acquired for each model. Figure S5: time-dependent RMSD evolution of rifampicin for each system. Color key: black: WT, red: mutant. Figure S6: Average per residue RMS fluctuation of rpoB computed based on C- α atom positions. Ligand binding regions are shaded grey. Blue bar indicates the position of associated mutation. Color code: black: WT, red: mutants. Figure S7: Structural mapping of the average per residue RMSF (calculated across all models). The color ranges from blue (low RMSF values) to red (high RMSF values). Figure S8: Time (ns) dependent Rg evolution of rpoB computed based on backbone atom positions. Color code: black: WT, red: mutants. Table S3: The proportion of variance captured by the top four principal components (PC1, PC2, PC3, and PC4). Generally, the first four PCs accounted for >60% of the variance. Trace values obtained following the diagonalization of the covariance are also tabulated. Figure S9: 2D projections of PC1 and PC2 as a function of time. Figure S10: Porcupine plots displaying concerted atomic motion acquired during simulation. Arrows represent the general direction of dominant motion, whereas the porcupine length represents the magnitude of motion. Figure S11: Structural comparison of dominant protein conformers extracted from low energy minimas on the conformational landscape (see Figure 5): WT (C1) versus mutants (C3, C4, C5, C6, C8, and C9). Color Key: WT: deep salmon, mutants: green. Table S4: Structural comparison of dominant

protein conformers: WT (conformer: C1) versus mutants. Figure S12: Time-dependent hydrogen bond numbers formed between rpoB and rifampicin during simulation. Figure S13: Protein–ligand interactions visualized using Discovery Studio Visualizer 4 (DS Visualizer) [72]. The representative structures were extracted from low energy minimas representing dominant protein conformations. Figure S14: Bar plots showing per residue contribution to the total binding free energy for each model. Shaded regions indicate areas located within the ligand binding pocket. Table S5: Tabulated summary of residues contributing substantially ($> \pm 3Xsd$) to the total binding free energy in each model. -ve indicates the negative binding free energy value, while +ve indicates that a positive energy value was recorded. Figure S15: Per residue average L plots. Table S6: Average BC pairwise Pearson's correlation values computed among models. The WT–RIF input is computed from per residue average values of run 1 and run 2. Figure S16: Plots of per residue average BC . Table S7: Residues possessing large average BC values (peaks) with respect to the WT systems (native representation). Table S8: Tabulated summary of residues that yielded large changes in average BC (ΔBC) due to mutations. Residues that yielded high average BC values are shown in bold. Residues unique to each model are underlined. Figure S17: Ensemble averaged residue contact map of the WT and mutated models. Supplementary_file S1: *rpoB* mutations retrieved from MUBII-TB-DB database. Supplementary_file S2: *rpoB* mutations prioritized in this study.

Author Contributions: Conceptualization, A.K. and Ö.T.B.; Data curation, A.K.; Formal analysis, A.A. and A.K.; Funding acquisition, Ö.T.B.; Methodology, A.A., A.K. and Ö.T.B.; Project administration, Ö.T.B.; Resources, Ö.T.B.; Supervision, Ö.T.B.; Validation, A.A. and Ö.T.B.; Visualization, A.A.; Writing—original draft, A.A.; Writing—review & editing, A.K. and Ö.T.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported through the Grand Challenges Africa Drug Discovery programme [GCA/DD/rnd3/023] awarded to Ö.T.B. Grand Challenges Africa is a programme of the African Academy of Sciences (AAS) implemented through the Alliance for Accelerating Excellence in Science in Africa (AESA) platform, an initiative of the AAS and the African Union Development Agency (AUDA-NEPAD). GC Africa is supported by the Bill & Melinda Gates Foundation (BMGF), Swedish International Development Cooperation Agency (SIDA), German Federal Ministry of Education and Research (BMBF), Medicines for Malaria Venture (MMV), and Drug Discovery and Development Centre of University of Cape Town (H3D). The views expressed herein are those of the author(s) and not necessarily those of the AAS and its partners.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets generated for this study are included in the article supplementary material.

Acknowledgments: The authors thank the Centre for High Performance Computing (CHPC), South Africa, for computing resources.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Sample Availability: Not available.

References

1. World Health Organization. *Global Tuberculosis Report (2019)*; WHO/CDS/TB; WHO: Geneva, Switzerland, 2019; ISBN 9789241565714.
2. Badie, F.; Arshadi, M.; Mohsenpoor, M.; Gharibvand, S.S. Drug Resistance Pattern of Mycobacterium tuberculosis Isolates From Patients Referred to TB Reference Laboratory in Ahvaz. *Osong Public Health Res. Perspect.* **2016**, *7*, 32–35. [[CrossRef](#)] [[PubMed](#)]
3. Boehme, C.C.; Nabeta, P.; Hillemann, D.; Nicol, M.P.; Shenai, S.; Krapp, F.; Allen, J.; Tahirli, R.; Blakemore, R.; Rustomjee, R.; et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N. Engl. J. Med.* **2010**, *363*, 1005–1015. [[CrossRef](#)] [[PubMed](#)]
4. Casali, N.; Nikolayevskyy, V.; Balabanova, Y.; Ignatyeva, O.; Kontsevaya, I.; Harris, S.R.; Bentley, S.D.; Parkhill, J.; Nejentsev, S.; Hoffner, S.E.; et al. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res.* **2012**, *22*, 735–745. [[CrossRef](#)] [[PubMed](#)]
5. Farhat, M.R.; Shapiro, B.J.; Kieser, K.J.; Sultana, R.; Jacobson, K.R.; Victor, T.C.; Warren, R.M.; Streicher, E.M.; Calver, A.; Sloutsky, A.; et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. *Nat. Genet.* **2013**, *45*, 1183–1189. [[CrossRef](#)] [[PubMed](#)]

6. Zhang, H.; Li, D.; Zhao, L.; Fleming, J.; Lin, N.; Wang, T.; Liu, Z.; Li, C.; Galwey, N.; Deng, J.; et al. Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* **2013**, *45*, 1255–1260. [CrossRef]
7. Singh, A.; Grover, S.; Sinha, S.; Das, M.; Somvanshi, P.; Grover, A. Mechanistic Principles Behind Molecular Mechanism of Rifampicin Resistance in Mutant RNA Polymerase Beta Subunit of Mycobacterium tuberculosis. *J. Cell. Biochem.* **2017**, *118*, 4594–4606. [CrossRef]
8. Kurz, S.G.; Furin, J.J.; Bark, C.M. Drug-Resistant Tuberculosis: Challenges and Progress. *Infect. Dis. Clin. N. Am.* **2016**, *30*, 509–522. [CrossRef]
9. Singh, R.; Dwivedi, S.P.; Gaharwar, U.S.; Meena, R.; Rajamani, P.; Prasad, T. Recent updates on drug resistance in Mycobacterium tuberculosis. *J. Appl. Microbiol.* **2020**, *128*, 1547–1567. [CrossRef]
10. Sharma, S.K.; Sharma, A.; Kadhiraivan, T.; Tharyan, P. Rifamycins (rifampicin, rifabutin and rifapentine) compared to isoniazid for preventing tuberculosis in HIV-negative people at risk of active TB. *Cochrane Database Syst. Rev.* **2013**, *7*. Available online: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD007545.pub2/full> (accessed on 8 December 2021). [CrossRef]
11. Alifano, P.; Palumbo, C.; Pasanisi, D.; Talà, A. Rifampicin-resistance, rpoB polymorphism and RNA polymerase genetic engineering. *J. Biotechnol.* **2015**, *202*, 60–77. [CrossRef]
12. Floss, H.G.; Yu, T.W. Rifamycin—Mode of action, resistance, and biosynthesis. *Chem. Rev.* **2005**, *105*, 621–632. [CrossRef] [PubMed]
13. Zenkin, N. Multiple personalities of the RNA polymerase active centre. *Microbiology* **2014**, *160*, 1316–1320. [CrossRef] [PubMed]
14. Mustaev, A.; Kozlov, M.; Markovtsov, V.; Zaychikov, E.; Denissova, L.; Goldfarb, A. Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6641–6645. [CrossRef] [PubMed]
15. Molodtsov, V.; Scharf, N.T.; Stefan, M.A.; Garcia, G.A.; Murakami, K.S. Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in Mycobacterium tuberculosis. *Mol. Microbiol.* **2017**, *103*, 1034–1045. [CrossRef] [PubMed]
16. Campbell, E.A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S.A. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **2001**, *104*, 901–912. [CrossRef]
17. Ramaswamy, S.; Musser, J.M. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber. Lung Dis.* **1998**, *79*, 3–29. [CrossRef]
18. Herrera, L.; Jiménez, S.; Valverde, A.; García-Aranda, M.A.; Sóez-Nieto, J.A. Molecular analysis of rifampicin-resistant Mycobacterium tuberculosis isolated in Spain (1996–2001). Description of new mutations in the rpoB gene and review of the literature. *Int. J. Antimicrob. Agents* **2003**, *403–408*. [CrossRef]
19. Zhang, Q.; Tan, S.; Xiao, T.; Liu, H.; Shah, S.J.A.; Liu, H. Probing the molecular mechanism of rifampin resistance caused by the point mutations S456L and D441V on mycobacterium tuberculosis RNA polymerase through gaussian accelerated molecular dynamics simulation. *Antimicrob. Agents Chemother.* **2020**, *64*, e02476-19. [CrossRef]
20. Gill, S.K.; Garcia, G.A. Rifamycin inhibition of WT and Rif-resistant Mycobacterium tuberculosis and Escherichia coli RNA polymerases in vitro. *Tuberculosis* **2011**, *91*, 361–369. [CrossRef]
21. Wang, H.Y.; Uh, Y.; Kim, S.; Cho, E.; Lee, J.S.; Lee, H. Detection of rifampicin- and isoniazid-resistant Mycobacterium tuberculosis using the quantamatrix multiplexed assay platform system. *Ann. Lab. Med.* **2018**, *38*, 569–577. [CrossRef]
22. Tessema, B.; Beer, J.; Emmrich, F.; Sack, U.; Rodloff, A.C. Analysis of gene mutations associated with isoniazid, rifampicin and ethambutol resistance among Mycobacterium tuberculosis isolates from Ethiopia. *BMC Infect. Dis.* **2012**, *12*, 37. [CrossRef] [PubMed]
23. Zaw, M.T.; Emran, N.A.; Lin, Z. Mutations inside rifampicin-resistance determining region of rpoB gene associated with rifampicin-resistance in Mycobacterium tuberculosis. *J. Infect. Public Health* **2018**, *11*, 605–610. [CrossRef]
24. Siu, G.K.H.; Zhang, Y.; Lau, T.C.K.; Lau, R.W.T.; Ho, P.L.; Yew, W.W.; Tsui, S.K.W.; Cheng, V.C.C.; Yuen, K.Y.; Yam, W.C. Mutations outside the rifampicin resistance-determining region associated with rifampicin resistance in Mycobacterium tuberculosis. *J. Antimicrob. Chemother.* **2011**, *66*, 730–733. [CrossRef] [PubMed]
25. Heep, M.; Rieger, U.; Beck, D.; Lehn, N. Mutations in the beginning of the rpoB gene can induce resistance to rifamycins in both Helicobacter priori and Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **2000**, *44*, 1075–1077. [CrossRef] [PubMed]
26. Nisha, J.; Shanthi, V. Computational simulation techniques to understand rifampicin resistance mutation (S425L) of rpoB in M. leprae. *J. Cell. Biochem.* **2015**, *116*, 1278–1285. [CrossRef] [PubMed]
27. Vedithi, S.C.; Rodrigues, C.H.M.; Portelli, S.; Skwark, M.J.; Das, M.; Ascher, D.B.; Blundell, T.L.; Malhotra, S. Computational saturation mutagenesis to predict structural consequences of systematic mutations in the beta subunit of RNA polymerase in Mycobacterium leprae. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 271–286. [CrossRef]
28. Zhang, Q.; An, X.; Liu, H.; Wang, S.; Xiao, T.; Liu, H. Uncovering the Resistance Mechanism of Mycobacterium tuberculosis to Rifampicin Due to RNA Polymerase H451D/Y/R Mutations From Computational Perspective. *Front. Chem.* **2019**, *7*, 819. [CrossRef]
29. Brown, D.K.; Sheik Amamuddy, O.; Tasthan Bishop, Ö. Structure-Based Analysis of Single Nucleotide Variants in the Renin-Angiotensinogen Complex. *Glob. Heart* **2017**, *12*, 121–132. [CrossRef]
30. Amusengeri, A.; Tata, R.B.; Tasthan Bishop, Ö. Understanding the Pyrimethamine Drug Resistance Mechanism via Combined Molecular Dynamics and Dynamic Residue Network Analysis. *Molecules* **2020**, *25*, 904. [CrossRef]

31. Sheik Amamuddy, O.; Musyoka, T.M.; Boateng, R.A.; Zabo, S.; Tasthan Bishop, Ö. Determining the unbinding events and conserved motions associated with the pyrazinamide release due to resistance mutations of Mycobacterium tuberculosis pyrazinamidase. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1103–1120. [[CrossRef](#)]
32. Flandrois, J.P.; Lina, G.; Dumitrescu, O. MUBII-TB-DB: A database of mutations associated with antibiotic resistance in Mycobacterium tuberculosis. *BMC Bioinform.* **2014**, *15*, 107. [[CrossRef](#)] [[PubMed](#)]
33. Lin, W.; Mandal, S.; Degen, D.; Liu, Y.; Ebricht, Y.W.; Li, S.; Feng, Y.; Zhang, Y.; Mandal, S.; Jiang, Y.; et al. Structural Basis of Mycobacterium tuberculosis Transcription and Transcription Inhibition. *Mol. Cell* **2017**, *66*, 169–179. [[CrossRef](#)] [[PubMed](#)]
34. Borukhov, S.; Nudler, E. RNA polymerase holoenzyme: Structure, function and biological implications. *Curr. Opin. Microbiol.* **2003**, *6*, 93–100. [[CrossRef](#)]
35. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)] [[PubMed](#)]
36. Steenken, W. Lysis of Tubercle Bacilli in Vitro. *Proc. Soc. Exp. Biol. Med.* **1935**, *33*, 253–255. [[CrossRef](#)]
37. Steenken, W.; Gardner, L.U. History of H37 strain of tubercle bacillus. *Am. Rev. Tuberc.* **1946**, *54*, 62–66. [[CrossRef](#)]
38. Portelli, S.; Phelan, J.E.; Ascher, D.B.; Clark, T.G.; Furnham, N. Understanding molecular consequences of putative drug resistant mutations in Mycobacterium tuberculosis. *Sci. Rep.* **2018**, *8*, 15356. [[CrossRef](#)]
39. Portelli, S.; Myung, Y.; Furnham, N.; Vedithi, S.C.; Pires, D.E.V.; Ascher, D.B. Prediction of rifampicin resistance beyond the RRDR using structure-based machine learning approaches. *Sci. Rep.* **2020**, *10*, 18120. [[CrossRef](#)]
40. Sanchez-Padilla, E.; Merker, M.; Beckert, P.; Jochims, F.; Dlamini, T.; Kahn, P.; Bonnet, M.; Niemann, S. Detection of drug-resistant tuberculosis by xpert MTB/RIF in Swaziland. *N. Engl. J. Med.* **2015**, *372*, 1181–1182. [[CrossRef](#)]
41. Woods, G.L.; Brown-Elliott, B.A.; Conville, P.S.; Desmond, E.P.; Hall, G.S.; Lin, G. *M24 Susceptibility Testing of Mycobacteria, Nocardiae, and Other Aerobic Actinomycetes*; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2011; ISBN 1562387464.
42. Santoveña-Estévez, A.; Suárez-González, J.; Cáceres-Pérez, A.R.; Ruiz-Noda, Z.; Machado-Rodríguez, S.; Echezarreta, M.; Soriano, M.; Fariña, J.B. Stability Study of Isoniazid and Rifampicin Oral Solutions Using Hydroxypropyl-B-Cyclodextrin to Treat Tuberculosis in Paediatrics. *Pharmaceutics* **2020**, *12*, 195. [[CrossRef](#)]
43. Ogwang, S.; Asiimwe, B.B.; Traore, H.; Mumbowa, F.; Okwera, A.; Eisenach, K.D.; Kayes, S.; Jones-López, E.C.; McNerney, R.; Worodria, W.; et al. Comparison of rapid tests for detection of rifampicin-resistant Mycobacterium tuberculosis in Kampala, Uganda. *BMC Infect. Dis.* **2009**, *9*, 139. [[CrossRef](#)] [[PubMed](#)]
44. Cocco, L.; Roth, B.; Temple, C.; Montgomery, J.A.; London, R.E.; Blakley, R.L. Protonated state of methotrexate, trimethoprim, and pyrimethamine bound to dihydrofolate reductase. *Arch. Biochem. Biophys.* **1983**, *226*, 567–577. [[CrossRef](#)]
45. Abdizadeh, H.; Tamer, Y.T.; Acar, O.; Toprak, E.; Atilgan, A.R.; Atilgan, C. Increased substrate affinity in the Escherichia coli L28R dihydrofolate reductase mutant causes trimethoprim resistance. *Phys. Chem. Chem. Phys.* **2017**, *19*, 11416–11428. [[CrossRef](#)] [[PubMed](#)]
46. Wang, H.Y.; Ahn, K.; Uh, Y.; Lee, H.; Kim, S.; Chang, Y.; Chang, C.L.; Shim, T.S. Direct detection of rifampin-resistant mycobacterium tuberculosis in respiratory specimens using Quantamatrix Multiplexed Assay Platform (QMAP) system: A multicenter study in Korea. *Front. Microbiol.* **2018**, *9*, 1804. [[CrossRef](#)] [[PubMed](#)]
47. Wolf, A.; Kirschner, K.N. Principal component and clustering analysis on molecular dynamics data of the ribosomal L11·23S subdomain. *J. Mol. Model.* **2013**, *19*, 539–549. [[CrossRef](#)]
48. Liu, S.-Q.; Ji, X.-L.; Tao, Y.; Tan, D.-Y.; Zhang, K.-Q.; Fu, Y.-X. Protein Folding, Binding and Energy Landscape: A Synthesis. In *Protein Engineering*; INTECH Open Access Publisher: London, UK, 2012.
49. Nussinov, R.; Tsai, C.J. Free energy diagrams for protein function. *Chem. Biol.* **2014**, *21*, 311–318. [[CrossRef](#)]
50. Ben-Levy, R.; Peles, E.; Goldman-Michael, R.; Yarden, Y. An oncogenic point mutation confers high affinity ligand binding to the neu receptor. Implications for the generation of site heterogeneity. *J. Biol. Chem.* **1992**, *267*, 17304–17313. [[CrossRef](#)]
51. Musyoka, T.M.; Njuguna, J.N.; Tasthan Bishop, Ö. Comparing sequence and structure of falcipains and human homologs at prodomain and catalytic active site for malarial peptide based inhibitor design. *Malar. J.* **2019**, *18*, 159. [[CrossRef](#)]
52. Amusengeri, A.; Tasthan Bishop, Ö. Discorhabdin N, a South African Natural Compound, for Hsp72 and Hsc70 Allosteric Modulation: Combined Study of Molecular Modeling and Dynamic Residue Network Analysis. *Molecules* **2019**, *24*, 188. [[CrossRef](#)]
53. Amusengeri, A.; Astl, L.; Lobb, K.; Verkhivker, G.M.; Tasthan Bishop, Ö. Establishing computational approaches towards identifying malarial allosteric modulators: A case study of plasmodium falciparum hsp70s. *Int. J. Mol. Sci.* **2019**, *20*, 5574. [[CrossRef](#)]
54. Penkler, D.; Atilgan, C.; Tasthan Bishop, Ö. Allosteric Modulation of Human Hsp90 α Conformational Dynamics. *J. Chem. Inf. Model.* **2018**, *58*, 383–404. [[CrossRef](#)] [[PubMed](#)]
55. Palomino, J.C.; Martin, A. Drug resistance mechanisms in Mycobacterium tuberculosis. *Antibiotics* **2014**, *3*, 317–340. [[CrossRef](#)] [[PubMed](#)]
56. Jamal, S.; Khubaib, M.; Gangwar, R.; Grover, S.; Grover, A.; Hasnain, S.E. Artificial Intelligence and Machine learning based prediction of resistant and susceptible mutations in Mycobacterium tuberculosis. *Sci. Rep.* **2020**, *10*, 5487. [[CrossRef](#)] [[PubMed](#)]
57. Nyamai, D.W.; Tasthan Bishop, Ö. Identification of selective novel hits against plasmodium falciparum prolyl tRNA synthetase active site and a predicted allosteric site using in silico approaches. *Int. J. Mol. Sci.* **2020**, *21*, 3803. [[CrossRef](#)] [[PubMed](#)]

58. Vehlow, C.; Stehr, H.; Winkelmann, M.; Duarte, J.M.; Petzold, L.; Dinse, J.; Lappe, M. CMView: Interactive contact map visualization and analysis. *Bioinformatics* **2011**, *27*, 1573–1574. [[CrossRef](#)] [[PubMed](#)]
59. Chemical Computing Group Inc. *Molecular Operating Environment (MOE)*, 2015.01; Chemical Computing Group Inc.: Montreal, QC, Canada, 2015.
60. Racheal, S.D.-M.; Zephaniah, G.; Reggie, M.; Kerina, D.; Babill, S.-P.; Peter, M. Diagnosis of Multi-drug Resistant Tuberculosis Mutations Using Hain Line Probe Assay and GeneXpert: A Study Done in Zimbabwe. *J. Adv. Med. Med. Res.* **2015**, *5*, 1044–1052. Available online: <http://www.sciencedomain.org/abstract/6677> (accessed on 4 July 2020). [[CrossRef](#)]
61. Mohajeri, P.; Sadri, H.; Farahani, A.; Norozi, B.; Atashi, S. Frequency of Mutations Associated with Rifampicin Resistance in Mycobacterium tuberculosis Strains Isolated from Patients in West of Iran. *Microb. Drug Resist.* **2015**, *21*, 315–319. [[CrossRef](#)]
62. DeLano WL The PyMOL Molecular Graphics System, Version 1.7 Schrödinger, LLC | www.pymol.org. Available online: <https://www.pymol.org/> (accessed on 16 May 2016).
63. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [[CrossRef](#)]
64. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M.C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012. [[CrossRef](#)]
65. da Silva, A.W.S.; Vranken, W.F. ACPYPE—AnteChamber PYthon Parser interface. *BMC Res. Notes* **2012**, *5*, 367. [[CrossRef](#)]
66. Feklistov, A.; Mekler, V.; Jiang, Q.; Westblade, L.F.; Irschik, H.; Jansen, R.; Mustae, A.; Darst, S.A.; Ebricht, R.H. Rifamycins do not function by allosteric modulation of binding of Mg²⁺ to the RNA polymerase active center. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14820. [[CrossRef](#)] [[PubMed](#)]
67. Amadei, A.; Linssen, A.B.M.; Berendsen, H.J.C. Essential dynamics of proteins. *Proteins Struct. Funct. Bioinform.* **1993**, *17*, 412–425. [[CrossRef](#)] [[PubMed](#)]
68. Brüschweiler, R. Collective protein dynamics and nuclear spin relaxation. *J. Chem. Phys.* **1995**, *102*, 3396. [[CrossRef](#)]
69. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897. [[CrossRef](#)]
70. Brown, D.; Penkler, D.; Sheik Amamuddy, O.; Ross, C.; Atilgan, A.R.; Atilgan, C.; Tastan Bishop, Ö. MD-TASK: A software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **2017**, *33*, 2768–2771. [[CrossRef](#)]
71. Atilgan, A.R.; Akan, P.; Baysal, C. Small-world communication of residues and significance for protein dynamics. *Biophys. J.* **2004**, *86*, 85–91. [[CrossRef](#)]
72. BIOVA BIOVIA Discovery Studio | Predictive Modeling & Science Simulation Software App. Available online: <http://accelrys.com/products/collaborative-science/biovia-discovery-studio/> (accessed on 27 May 2016).