Tuberculosis 114 (2019) 61–68

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: www.elsevier.com/locate/tube

Theoretical insights on the binding of isoniazid to the active site residues of *Mycobacterium tuberculosis* catalase-peroxidase

Yves Ira A. Reyes, Gerardo C. Janairo, Francisco C. Franco Jr.*

Chemistry Department, De La Salle University, 2401 Taft Avenue, 0922, Manila, Philippines

ARTICLE INFO

Keywords:
Isoniazid
Tuberculosis
Drug resistance
Computational chemistry
Drug design

ABSTRACT

Isoniazid (INH) is known to cause the exclusive lethal action to *Mycobacterium tuberculosis* (*M. tb.*) cells because of the pathogen's own catalase-peroxidase (katG) enzyme that converts INH to a very reactive radical. Thus, in order to gain insights on the interaction of INH with the individual active site residues (Res) of katG, this study presents a computational approach via molecular docking and density functional theory (DFT) using augmented models to study the individual INH-Res interactions. Seven amino acid residues directly interacts with INH: Arg104, Asp137, His108, Ile228, Trp107, Tyr229, and Val230. The residues with the highest interaction energies are Arg104 (−39.64 kcal/mol) and Asp137 (−32.85 kcal/mol) mainly due to strong ion-dipole and H-bonding interactions present in the complexes, while the weakest interaction was observed for Ile228 (−13.78 kcal/mol). Molecular electrostatic potential surface revealed complementary regions for dipole interactions and charge distribution analysis further shows that INH generally donates electrons to the residues. The results in this study agrees with the previous experimental findings and provides new insights into the catalase dependent activation of INH and the methods presented may be valuable in the study of biological metabolism of molecules.

1. Introduction

Tuberculosis (TB) is a global epidemic that is among the top ten diseases that pose a threat to human health [1,2]. It is caused by the infection of the bacterial pathogen, *Mycobacterium tuberculosis* (*M. tb*), in the human lungs [2,3]. Isonicotinylhydrazide (INH), more commonly known as Isoniazid, is one of the most efficient drugs currently being administered to treat TB patients [4]. INH is preferred due to its limited side effects to the patient because INH is usually harmless to living organisms. However, it becomes selectively lethal to *M. tb* cells because they possess an enzyme called catalase-peroxidase (katG), which binds and catalyzes the conversion of INH to a very reactive radical [5,6]. The activated INH, the radical, then reacts with nicotinamide adenine dinucleotide (NAD), a substrate of the enoyl acyl carrier protein reductase (InhA) of the pathogen to form an INH-NAD adduct which irreversibly binds to the InhA [7,8]. InhA is a crucial protein to produce an important cell wall component of the *M. tb* cells, thus its inhibition through the irreversible binding of the INH-NAD adduct proves fatal to the pathogen.

A subject of interest among many studies, especially using computational methods, is the mechanism of the katG dependent activation of INH since it provides information to explain the prevalence of INH resistant *M. tb* strains [5,7,9]. Since the crystal structure of the katG protein has already been resolved through X-ray crystallography [6], the use of molecular docking and molecular dynamics simulations have been prevalent in determining the interacting residues explaining how their mutations cause drug resistance [7,10] and the role of the co-factors in the reaction that converts INH to its radical form [7,11]. However, these studies are limited to simulations grounded on the molecular mechanics model of the INH-KatG system.

Since molecular mechanics model atoms as charged spheres, the redistribution of electrons in one molecule as it comes close to interact with another molecule is neglected in this kind of study. This leads to poor predictions of noncovalent interactions. Therefore, though the studies cited above give the big picture of the KatG-INH binding process, they are limited in terms of insights into the individual importance of each amino acid residue in binding the ligand. This is important in understanding the development of drug resistance through mutations and possibly, novel drug design. Most importantly, molecular mechanics does not provide an accurate description of the bonding interactions, especially weak bonding interactions. Thus, there is a need to further elucidate these bonding interactions using higher level calculations to have a better description of these systems.

The use of higher level of computations, i.e., quantum chemical methods, is necessary to fill in the said information gaps. Density functional theory (DFT) enables prediction of optimized geometries and

* Corresponding author.
E-mail address: francisco.franco@dlsu.edu.ph (F.C. Franco).

https://doi.org/10.1016/j.tube.2018.11.005
Received 14 July 2018; Received in revised form 29 September 2018; Accepted 14 November 2018

calculation of free electronic energies of molecules with considerations for electronic structures of involved atoms, at a reasonable computational cost [12]. Due to its balance between computational cost and accuracy, DFT is commonly used to understand the properties of materials and chemical systems [13–15] design new materials [16–18], and study reaction mechanisms [14,19]. However, use of quantum chemical methods for drug design has been limited so far. Isolated INH has been previously treated with DFT using B3LYP functional to successfully predict its geometry, molecular orbitals and vibrational modes [20]. But the significance of these data does not compare to a study of a drug-receptor complex. And though still impractical for whole protein systems, DFT has also been demonstrated to successfully characterize augmented models, including only core interacting residues and a pre-aligned ligand [21–23] using mainly B3LYP functional.

In this study, we applied DFT to understand the binding of INH into the active site of katG by forming a complex between INH and each individual residue in the active site. Molecular docking calculations were carried out to determine the active site residues (Res) interacting with INH, while DFT calculations were implemented to determine the structural and electronic properties of the INH-Res complexes. The accuracy of DFT computations rely on the choice of functionals. The successful application of B3LYP in biomolecular systems and with the INH structure, as described above, made it reasonable to use the same functional with the current study. With this study, we also aim to understand the importance of each individual active site residue in the binding of INH and to increase the use of quantum mechanical methods in studying drug-protein interactions for drug design applications.

2. Computational details

All DFT calculations were carried out using the General Atomic and Molecular Electronic Structure System (GAMESS) [24] quantum chemistry package. The gas-phase equilibrium structures were determined using 6-31G(d) as the basis set and B3LYP as the functional. Normal DFT methods do not consider long range electron correlations or simply, the interaction of electrons with electrons of a separate but nearby molecule. Because of this, effects of electron dispersion or commonly known as London Dispersion Forces are not properly simulated, resulting in an error in energy computations of complexes. Recently, Grimme et al. proposed a method for correcting this error, and the latest iteration of this method, Grimme's dispersion correction (D3) with Becke and Johnson damping (BJ) was shown to improve the accuracy of energy calculations of noncovalently bonded systems [25–27]. Thus, this correction method, D3-BJ was also implemented in this study. The specific DFT method characterized by the combination of functional and basis set described above is hereinafter referred to as DFT-D3(BJ)/B3LYP/6-31G(d). The equilibrium structures were then confirmed by frequency analysis to make sure that these structures correspond to true minima. Molecular docking simulations were done using AutoDock 4.0 using the interface of AutoDockTools [28].

The INH structure was manually drawn using Avogadro molecular viewer [29] and pretreated by geometry optimization using MMFF [30]. The resulting structure was then further optimized using DFT-D3/(BJ)/B3LYP/6-31G(d) as described above. The final DFT optimized structure was then docked to the crystal structure of the katG protein. Aside from not allowing bond rotations to maintain the DFT optimized geometry of INH, similar parameters with a study that docked the same system [10], were used for the docking: 200 genetic algorithm search, 150 population size, 2500000 maximum evaluations and 27000 maximum generations. The best docking pose, based on binding energy and cluster analysis, was analyzed using ligplot+ [31] to determine the residues most likely to interact directly with INH. The coordinates of the interacting residues along with the top INH binding pose was used for DFT analysis.

Each interacting residue (Res) was analyzed independently using DFT. Each residue was modeled as a “dipeptide” [32], illustrated in

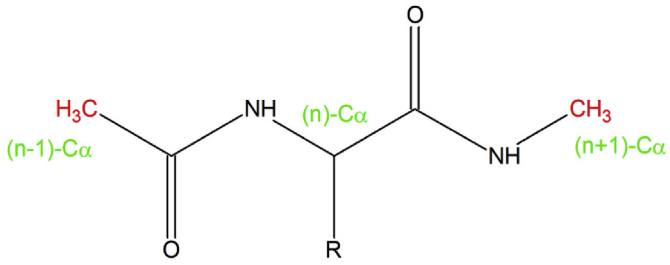

Fig. 1. Illustration of the dipeptide model of amino acid residues. Capping methyl groups, highlighted in red, were frozen during DFT optimization. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 1. In this model, the isolated amino acid residue being studied (n) is truncated at the alpha carbons of the flanking residues (n-1 and n+1) and capping these carbons with hydrogens. The models were then protonated to physiological pH. Thus, Arg104 had a positive protonated guanidino group while Asp137 had a negative deprotonated carboxyl group. Histidine is classified as a basic amino acid but the imidazole nitrogen (pKa ~6) would likely favor the neutral deprotonated state at physiological pH, thus His108 was modeled as a neutral residue. The capping hydrogens were optimized using MMFF but the resulting CH_3 caps were frozen in the subsequent DFT calculations. For each residue, the geometry of 1) the isolated residue and 2) the residue interacting with the docked INH (INH-Res) were optimized using DFT-D3(BJ)/B3LYP/6-31G(d) as described above. Single point energy calculations were then done on the optimized geometries of the samples using the same combinations of functional and basis set, DFT-D3(BJ)/B3LYP/6-31G(d). In applying DFT calculations to complexes, an error in energy calculations usually emerge due to the overlap of basis sets. This error is generally referred to as the basis set superposition error (BSSE). Thus, geometric Counter Poise (gCP) [33] correction was also implemented to remove the BSSE. The gCP-D3 corrections for the B3LYP/6-31G(d) energy calculations were calculated from Grimme's Webservice [34]. The gCP energy correction was then added to the calculated single point energies to get the corrected DFT-gCP-D3(BJ)B3LYP/6-31G(d) energies of the complexes. The gCP-D3 corrected interaction energy, $E_{\text{int}}^{\text{gCP-D3}}$, and uncorrected interaction energy, E_{int}, between each interacting amino acid residue and INH were calculated as:

$$E_{\text{int}}^{\text{gCP-D3}} = E_{\text{Complex}}^{\text{gCP-D3}} - E_{\text{INH}}^{\text{gCP-D3}} - E_{\text{Res}}^{\text{gCP-D3}} \quad (1)$$

$$E_{\text{int}} = E_{\text{Complex}} - E_{\text{INH}} - E_{\text{Res}} \quad (2)$$

where E_{complex} is the energy of the optimized complex of INH with the dipeptide model of a binding site amino acid residue, E_{INH} is the energy of the optimized isolated INH molecule and E_{Res} is the energy of the optimized isolated amino acid residue dipeptide model. Further analysis was done by examining the molecular electrostatic potential (MEPs) surface, frontier molecular orbitals (FMOs), and charge analysis of the optimized complexes. MEPs and FMOs were calculated and viewed using MacMolPlt [35], while NBO analysis was carried out using the NBO program [36].

3. Results and discussion

3.1. Geometric optimization of INH-Amino acid complexes

The docking of the DFT-optimized structure of INH ensured that only the most stable, and thus, the most probable conformation of INH would be considered by AutoDock. Out of 200 docking poses generated, the best pose belonged to a cluster with 73 conformations and had a binding energy of −6.8 kcal/mol. The binding site residues and their predicted interactions with the best INH pose are shown in Fig. 2. Fig. 2A is a 3D illustration of the enzyme-INH complex generated by

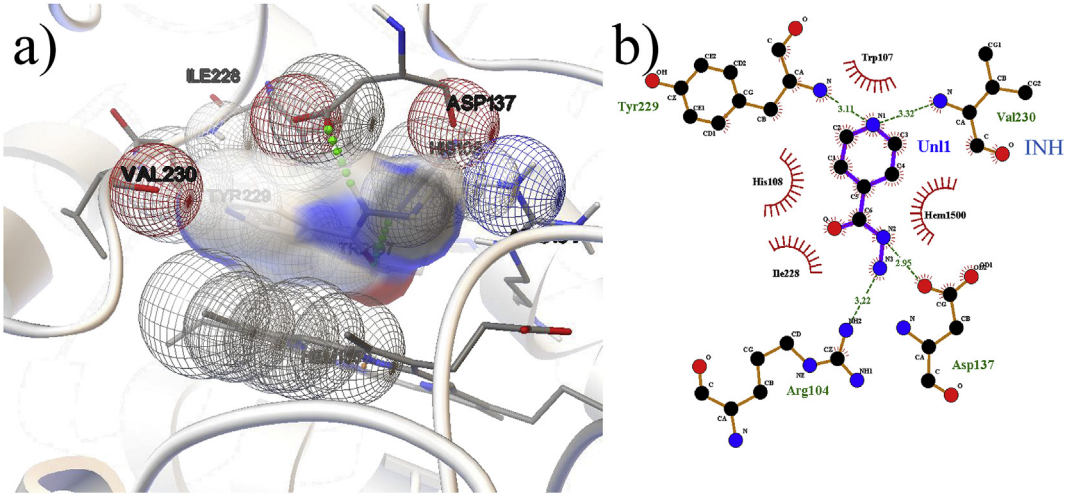

Fig. 2. Interaction diagrams of the best docking conformation of INH with amino acids in the binding site of katG shown in a) 3D (Autodock Tools) and b) 2D (ligplot +, with bond rotations). Hydrogen bonds are depicted as green dotted lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

AutoDock tools, where only polar hydrogens are shown explicitly, while Fig. 2B is a 2D representation and all hydrogens were omitted. Seven amino acid residues (Arg104, Trp107, His108, Asp137, Ile228, Tyr229, and Val230) were observed to directly interact with INH. It is important to note that the heme was also found to directly interact with INH in the binding site. However, the current study only focuses on the specific amino acids present in the active site and their importance in the binding of INH. In Fig. 2a, based on Autodock tools' default hydrogen parameters, only Asp137 and His108 are capable of H-bonding. On the other hand, in Fig. 2b, using ligplot +, it is shown that Asp137, Arg104, Tyr229, and Val230 will form H-bonds.

The primary difference between Autodock and ligplot + is the cutoff distance and the nature of the distance being measured. The Autodock tools protocol we employed measures the distance from the hydrogen to the H-bond acceptor and only considers it an H-bond if the distance falls below 2.25 Å. On the other hand, ligplot + in its default algorithm, removes all hydrogens before searching for H-bond donors (i.e., electronegative atoms that can be protonated) and H-bond acceptors. Then, the H-bonds are assessed based on the distance between the H-bond donor and acceptor, where an H-bond is predicted if the donor to acceptor distance falls below 3.35 Å. Therefore, ligplot + 's criteria for H-bond is more relaxed compared to Autodock resulting to more H-bonds detected. The only discrepancy is the residue His108, which is due to the limitation of the ligplot + in assigning H-bond acceptors for ligands [37]. In ligplot +, some expected H-bonds may not be detected by their algorithm due to mistakes in characterizing atoms in different ligands. The discrepancy in the H-bonding residues with INH using the docking analysis may be resolved by understanding the strength of H-bonding between INH and the individual residues using higher level computational methods, e.g., DFT.

The coordinates of the best docking pose were used as a reference for the subsequent DFT analysis of the binding complex, which is composed of the INH and the dipeptide model of the amino acid residues directly interacting with it. It is important to note that the isolated dipeptide model, illustrated in Fig. 1, of the amino acid residues, have limitations. Besides the anchoring effect to adjacent residues due to frozen n-1 and n + 1 methyl caps, the effect of surrounding katG protein structure is neglected. Furthermore, the subsequent model of the complex of each residue with INH is limited to the isolated effect of single residues and neglects the possible cooperative or competitive nature of simultaneous interactions of different residues to INH. Catalase-peroxidase class of enzymes are known to undergo post-translational modification that causes the side chains of a methionine, tyrosine, and tryptophan residues to be covalently bonded to form an adduct, commonly referred to as the M-Y-W adduct as shown in Fig. S1 of the Supplementary Materials. The sulfur atom of Met255 and the C7 of the indole group of Trp107 are covalently bonded to the carbons adjacent to the hydroxyl carbon of the phenol ring of Try229. From Fig. S1, it can be observed that Met255 do not directly interact with INH. On the other hand, Trp107 interacts with INH through the pyridine ring component of its indole group, while Tyr229 interacts with INH through its peptide backbone. Furthermore, the region where the covalent linkage between the adduct occurs do not interact directly with INH. Therefore, to simplify the modeled complexes, only the interactions between Tyr229-INH, and Trp107-INH were considered and not INH with the whole M-Y-W adduct. Lastly, as discussed further below, some atoms will be represented twice in adjacent residues. With these limitations in mind, the use of this model is still reasonable since it greatly reduces the computational cost to do a quantum mechanical description of the importance of each binding site amino acid residue and its interaction with INH.

The DFT-optimized complex of INH with each binding site amino acid residue is shown in Figs. 3–6. The interaction energy of individual amino acid residues with INH is calculated using equations (1) and (2) and is summarized in Table 1. Comparing the geometry of DFT optimized complexes (Figs. 3–6) to the result of molecular docking (Fig. 2), INH was observed to have moved closer to each residue since the DFT model isolates the interaction with each residue. Thus, new possible optimal interactions were observed. The B3LYP/6-31G(d) interaction energy for INH-Arg104 complex was observed to be −39.64 kcal/mol for gCP-D3 corrected and −37.91 kcal/mol without correction. It resulted to two strong H-bonds (Fig. 3a): N—H···O (d = 1.96 Å, 152.5°) and N—H···O (d = 2.00 Å, 148.5°). For INH-Asp137 complex, the interaction energy was observed to be −32.85 kcal/mol (−32.65 kcal/mol, uncorrected) and forms one strong H-bond (Fig. 3b): N—H···O (d = 1.80 Å, 172.5°). These two INH complexes show the strongest interaction energies compared to other INH-Res complexes mainly due to the ion-dipole interactions (shown as positive and negative regions in Fig. 7): positive Arg104 and negative Asp137 residues together with the presence of H-bonds. Due to the two H-bonds present in INH-Arg104, its interaction energy is stronger than INH-Asp137. For the neutral residues, the INH-His108 complex was observed to have an interaction energy of −14.17 kcal/mol (−13.40 kcal/mol, uncorrected) while forming one H-bond (Fig. 4a): N—H···O (d = 1.84 Å, 163.2°). For the INH-Ile228 complex, only a weak H-bonding was observed (Fig. 4b): N—H···N (d = 2.38 Å, 134.4°) resulting to the weakest interaction

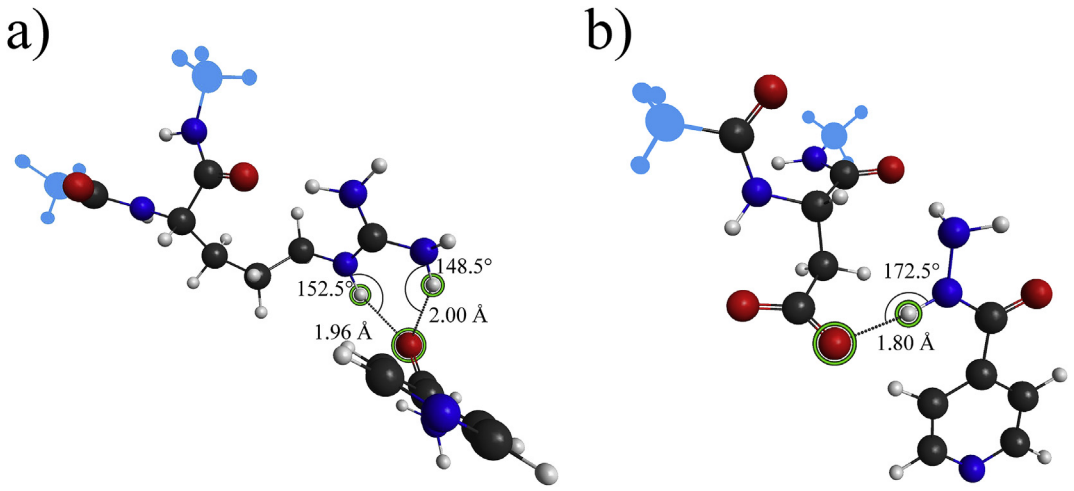

Fig. 3. DFT optimized complex of INH with a) Arg104 and b) Asp137. Methyl caps, highlighted in blue, were frozen during DFT optimizations. H-bonded atoms are encircled with a green line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

energy of −13.78 kcal/mol (−7.46 kcal/mol, uncorrected). It must be noted that without the D3 correction, weak van der Waals interaction are not considered which resulted to a much weaker interaction energy (−7.46 kcal/mol). Thus, it is very important to apply the dispersion (D3) corrections especially for weakly interacting systems. For INH-Trp107, a strong H-bond (Fig. 5a): N—H···O (d = 1.94 Å, 156.9°) was observed resulting to an interaction energy of −14.98 kcal/mol (−13.06 kcal/mol, uncorrected). In Tyr229, one strong H-bond (Fig. 5b): N—H···N (d = 2.01 Å, 171.7°) was observed which has an interaction energy of −18.66 kcal/mol (−14.00 kcal/mol, uncorrected). Lastly, for INH-Val230, an interaction energy of −14.56 kcal/mol (−12.30 kcal/mol) was observed mainly due to the H-bond (Fig. 6): N—H···N (d = 2.06 Å, 162.5°).

The dipeptide model extends the amino acid residue, n, up to the alpha carbons of the n+1 and n-1 residues. Because of this, the main chain atoms (alpha carbon, amino carbon, carbonyl carbon, amino hydrogen, amino nitrogen and carbonyl carbon), will be represented twice across adjacent residues and will contribute to the calculated interaction energies. For instance, since the dipeptide model of Tyr229 (Fig. 5b) contains the amino nitrogen and alpha carbon of the n+1 residue, which is Val230, the hydrogen bond observed between the amino hydrogen of Val230 (Fig. 6) and pyridine nitrogen of INH is also the same H-bond present in the dipeptide model of Tyr229 resulting to the presence of two hydrogen bonds in the INH-Tyr229 complex even though only one hydrogen bond can really be associated with Tyr229. The second hydrogen bond can be considered as a weak H-bonding (d = 2.46 Å, θ = 148.5°) but will nevertheless contribute to the lowering of the interaction energy. Also, the bond distance for the same H-bond in INH-Tyr229 and Val230 are similar (2.01 Å and 2.06 Å, respectively) but due to the presence of a different residue, and absence of the second H-bond in INH-Val230, the H-bond angle was slightly lower in INH-Val230 compared to INH-Tyr229. On a similar note, the H-bond in INH-Ile228 is due to the amino hydrogen of Tyr229. The same H-bond is duplicated in INH-Ile228 and INH-Tyr229 though the bond distance is shorter and the bond angle is lower in the INH-Ile228 model compared to INH-Tyr229 model.

Figs. 3a, 4a and 5a show that Arg104, His108, and Trp107 can all act as H-bond donors to the carbonyl carbon of INH but it is common knowledge that carbonyl oxygen can only accept at most two H-bonds. In the current model, it might be difficult to assure, which residue would form the interaction, if all the residues are present at the same time. It can be argued that Arg104 should take priority since it provides

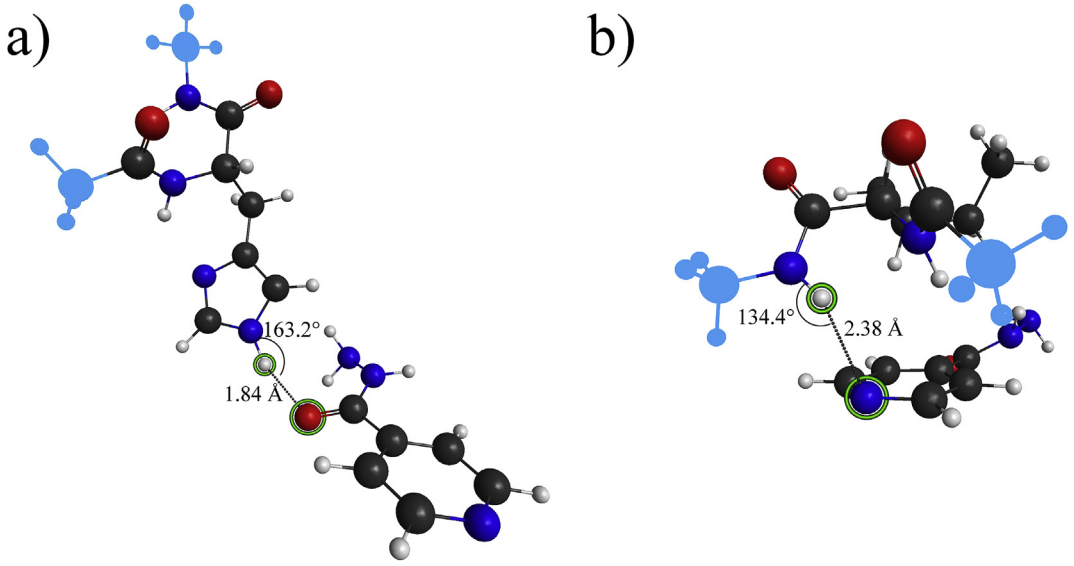

Fig. 4. DFT optimized complex of INH with a) His108 and b) Ile228. Methyl caps, highlighted in blue, were frozen during DFT optimizations. H-bonded atoms are encircled with a green line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

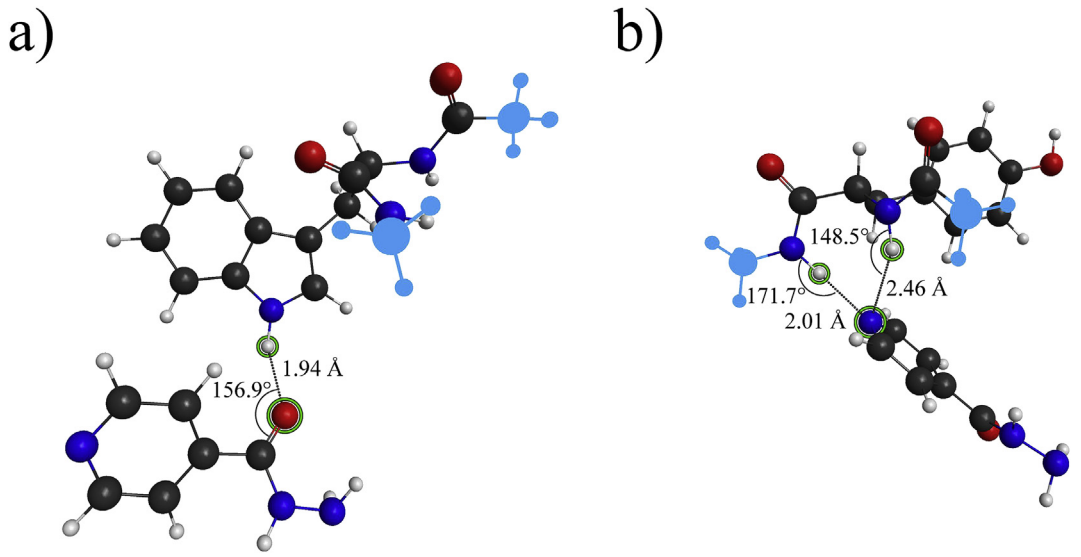

Fig. 5. DFT optimized complex of INH with a) Trp107 and b) Tyr229. Methyl caps, highlighted in blue, were frozen during DFT optimizations. H-bonded atoms are encircled with a green line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

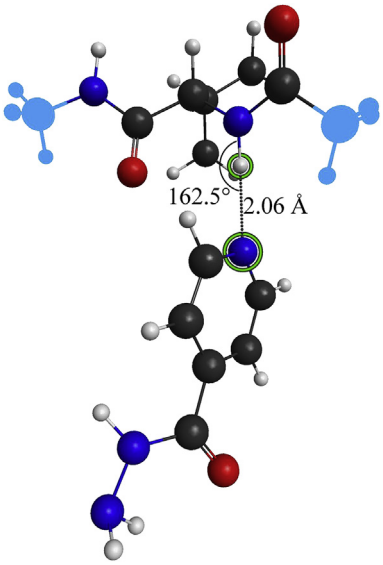

Fig. 6. DFT optimized complex of INH with Val230. Methyl caps, highlighted in blue, were frozen during DFT optimizations. H-bonded atoms are encircled with a green line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
H-bond distances and calculated interaction energies (gCP-D3 corrected and uncorrected using DFT/B3LYP/6-31G(d) for each amino acid-INH complex.

INH-Res	*d* (Å)	θ (°)	Interaction energies (kcal/mol)	
			Uncorrected	gCP-D3 corrected
Arg104	1.96	152.5	−37.91	−39.64
	2.00	148.5		
Asp137	1.80	172.5	−32.65	−32.85
His108	1.84	163.2	−13.40	−14.17
Ile228	2.38 (weak)	134.4	−7.46	−13.78
Trp107	1.94	156.9	−13.06	−14.98
Tyr229	2.01	171.7	−14.00	−18.66
	2.46 (weak)	148.5		
Val230	2.06	162.5	−12.30	−14.56

the strongest interaction. However, gathering all the possible mode interactions predicted in both the molecular docking and DFT optimizations, an optimal solution would be for Arg104 to interact with INH through the available, though less favorable, H-bond acceptor in the terminal β-nitrogen of the hydrazide group, leaving the carbonyl oxygen available to accept H-bonds from both His108 and Trp107.

3.2. Molecular electrostatic potential surface

Analysis of the molecular electrostatic potential (MEP) surfaces, shown in Fig. 7, of isoniazid and the interacting residues, provides further affirmation in the optimized geometry and interactions observed in Fig. 2 and Figs. 3–6. INH was found to have high electron densities in its pyridine nitrogen, carbonyl oxygen and β-nitrogen atoms while the hydrogens are relatively electron deficient. Arg104, being a basic amino acid residue, was modeled such that the guanidino group is protonated and thus, is positively charged, resulting in a molecule that is predominantly deficient in electrons. This affirms the previous prediction that this residue would interact with electron-rich regions of INH like the β-nitrogen, as initially predicted by the molecular docking result (Fig. 2), or the carbonyl oxygen, as predicted by the DFT optimization (Fig. 3a). In contrast, Asp137, an acidic residue, was modeled such that its γ carboxyl group is deprotonated resulting in a molecule that is predominantly electron rich. Thus, its interaction is expected to be directed to the hydrogens of INH. This agrees with the predictions in Figs. 2 and 3b where the carboxyl group of Asp137 was predicted to act as H-bond acceptor from the hydrogen attached to the α-Nitrogen of INH. Compared to other residues, the overwhelming density of electrons, in the acidic Asp137, and lack thereof, in the basic Arg104, can explain why they presented the strongest interaction energies (Table 1). Both INH-Arg104 and INH-Asp137 MEP surfaces showed that the INH and Arg104 and Asp137 residues interacted at suitable distances.

Unlike Arg104 and Asp107, the other residues are electrically neutral, but it is clear in Fig. 7, that the uneven distribution of electrons in these molecules result in electron rich and electron deficient regions which then results in possible electrostatic attraction to certain regions of INH. Analysis of Figs. 3–6 and Fig. 7 generally shows that the electron-rich regions of the residue are the regions oriented toward the electron deficient regions of INH or *vice versa*. The MEP surfaces for the complexes except INH-Ile228 showed that significant sharing of electron cloud densities are observed resulting to strong interactions.

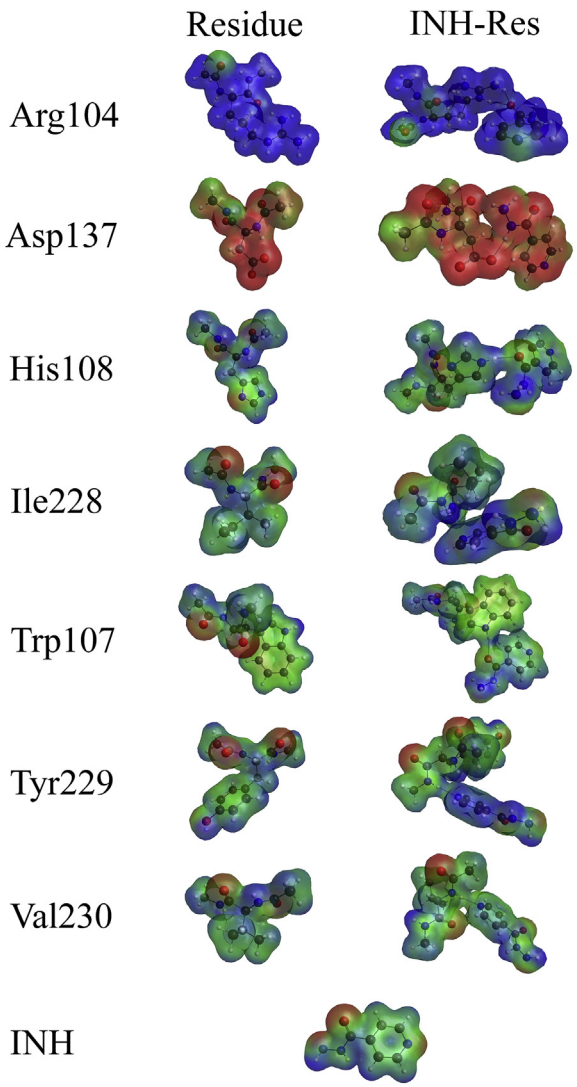

Fig. 7. Molecular electrostatic potential surfaces (isovalue = 0.01 a.u.) of the INH, residues, and INH-Res complexes. Red, green, and blue surfaces depict low, neutral, and high potential regions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Inspection of INH-Ile228 MEP surface reveals that the orientation of INH is unideal for favorable dipole interactions to form. Since Ile228 is mostly neutral or positive, except the carbonyl oxygen in the backbone, its only possible strong interaction with INH is through the electron-rich pyridine nitrogen of INH which is oriented in such a way that it is crowded by the neutral to positive rings. This further explains why it has the weakest interaction energy with INH.

3.3. Mulliken and NBO charge analysis

The charge transfer between INH and the residues were estimated using Mulliken ($Q_{Mulliken}$) and NBO (Q_{NBO}) charge analysis at the B3LYP/6-31G(d) level of theory. Although it is known that atomic charge analyses are dependent on the basis set used, using the same basis set for the systems studied will provide meaningful results [36,38]. Table 2 summarizes the charges for the INH molecules in the INH-Res complexes. Although the $Q_{Mulliken}$ and Q_{NBO} charges have different absolute values, the trends are the same for both analyses. It can be observed that in most complexes (except INH-Asp137), the INH molecule donates e^- charge to the residues, which results to a net positive charge for the INH molecule. The residue that accepts the most electron charge is the Arg104 residue (0.100 e^- and 0.052 e^-, for $Q_{Mulliken}$ and Q_{NBO}, respectively), mainly due to its positive charge in this model. For the INH-Asp137 complex, the INH molecule accepts electron charge (−0.191 e^- and −0.106 e^-, for $Q_{Mulliken}$ and Q_{NBO}, respectively) from the negatively charged residue, resulting to a net negative charge. The INH-Ile228 complex was observed to have the least charge transfer (0.036 e^- and 0.012 e^-, for $Q_{Mulliken}$ and Q_{NBO}, respectively). Since the main interaction between INH and the residues is H-bonding, where the O or N atom of INH forms an H-bond with the H atom of the residues, then it should be expected that the INH molecule will donate e^- charge. Also, the magnitude of the charge transfers may be correlated with the strength of interaction energies shown in Table 1, the higher the charge transfer, the stronger is the interaction energy. It is interesting to see that besides Arg104, the greatest charge transfer was calculated for Tyr229, followed by Trp107. According to Suarez et al., these two residues, along with Met255 form a covalently linked methionine-tyrosine-tryptophan (MYW) adduct near the binding site and the heme [39]. This MYW is supposed to be directly involved in the catalase cycle of the katG by donating a single electron to the porphyrin ring [39,40]. Thus, their tendency to accept electron from INH, as suggested in the charge transfer analysis, may hint at their direct involvement in the transformation of INH to a radical species.

Table 2
Mulliken and NBO charges for the INH molecule in the INH-Res complexes.

INH-Res	$Q_{Mulliken}$ (e^-)	Q_{NBO} (e^-)
-Arg104	0.100	0.052
-Asp137	−0.191	−0.106
-His108	0.052	0.034
-Ile228	0.036	0.012
-Trp107	0.052	0.027
-Tyr229	0.085	0.040
-Val230	0.041	0.025

3.4. Molecular orbitals

The frontier orbital energies and MO diagrams of the complexes formed between INH and each of the seven binding site amino acid residues are shown in Fig. 8. Generally, the residues contribute mostly to the HOMO orbital energy and surface (Fig. 8, thick black lines), while the INH contribute mostly to the LUMO orbital energy and surface (Fig. 8, thick red lines). In fact, the LUMO energy of INH is very similar to the LUMO energy of the INH-Res complexes, greatly lowering the LUMO energies of the residues (Fig. 8, thin red lines). The only exception to this is the HOMO of the Asp137-INH complex where a little density was present in the ligand, too. This means that for the said complex, the HOMO is formed from a significant overlap of molecular orbitals from the isolated INH and Asp137 molecules. This overlap might explain why Asp137 has very strong interaction energy with INH. These results may have great implication in the activation of INH to the reactive radical in the binding site of katG.

The overview of the analysis of the geometry, MEP, and MOs reveals that Asp137 and Arg104 are the two biggest contributors to the binding of INH in the active site of the katG protein while Ile228 provides the least interaction with INH. Although the mutated residues were not simulated in this study, our results agree with previous experimental finding as explained in the next sentences. Thus, the methods presented in this study predict that mutations of these residues would result to a decreased affinity with INH and potentially, the resistance of the *M. tb.* strains with such mutations to INH treatment. This is especially true for Arg104, Trp107, Asp137, and His108 where the side chains are directly involved in H-bonding. Consulting the Tuberculosis Drug Resistance Mutation Database (TBDReaMDB) [41], it was reported that mutations of Arg104, Trp107, and His108 cause drug resistance. The replacement

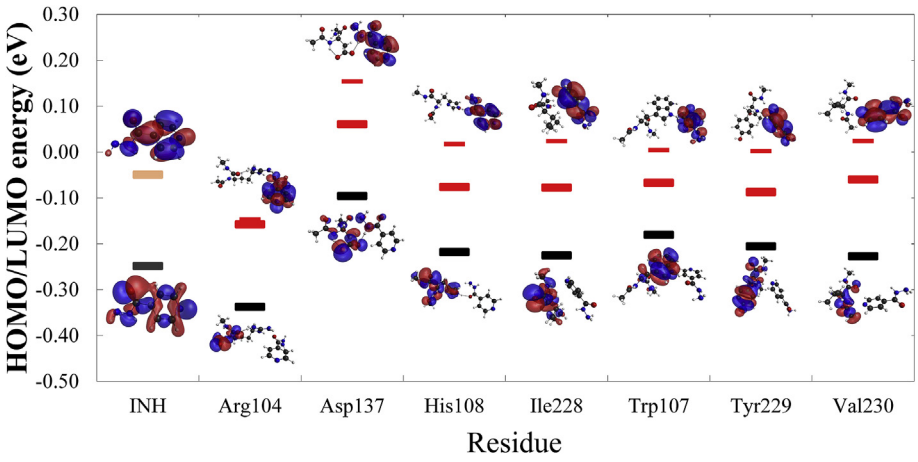

Fig. 8. Calculated HOMO and LUMO energy values and contour surfaces for INH and INH-Res complexes using DFT-D3(BJ)/B3LYP/6-31G(d). The thick black lines represent the HOMO energies of the complex, while the thick red lines represent the LUMO energies of the complexes. The HOMO and LUMO energies for INH is also shown, while the LUMO energies of the residues are shown as thin red lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of Arg104 by a Leu residue was reported to decrease the peroxidase activity by 100 fold [42]. Since Leu is a nonpolar alkyl amino acid, it cannot provide the electron deficient region and H-bond donating characteristics of an Arg residue observed in this study. Mutation of Ile228 has not been reported to cause drug resistance and with the findings in this study, it is highly probable that the mutation of this residue is indeed insignificant to the peroxidase activity of the protein. More interesting though is the fact that mutation of Asp137, having the second strongest interaction with INH according to this study, has never been correlated to isoniazid resistance. It may be that mutation at this position just does not happen frequently enough to be reported in experimental studies. The importance of Asp137 predicted by the presented model needs to be further confirmed. Experimentally, site-directed mutagenesis can force a mutation of this amino acid residue and determine its effect of peroxidase activity of the *M. tb* katG. With all that said, it is important to note that the current computational model, since it focuses on direct interactions, does not account for possible 3D protein structural changes that can be brought by the mutations being considered above. However, since INH and the residues will not be frozen for real systems, this study still shows the possible low energy interactions between INH and each of the residues and their importance in the binding of INH to the active site.

4. Conclusions

The application of quantum mechanical computations on a drug-protein system through augmented models was successfully demonstrated in this study. Preliminary molecular docking of INH to katG reveals seven amino acid residues involved in binding INH: Arg104, Trp107, His108, Asp137, Ile228, Tyr229, Val230. DFT analysis of the interaction of each residue with INH revealed that Asp137 and Arg104 interact most strongly with INH due to H-bonding with the side chain group of the said residues. Complementary dipole interactions were also observed through the analysis of the MEP surfaces. Charge transfer analysis shows that for most residues, INH is likely to donate electrons. Analysis of frontier MOs shows that the HOMO is generally exclusively contributed by the residue while the LUMO is contributed by INH. Ile228 on the other hand, had the weakest interaction with INH because their orientations limited formation of significant interaction.

Although the augmented model was used to minimize computational cost, and employed assumptions that may not hold true in the real environment, this study may still provide insights on the importance of each residue present in the active site of katG and its binding interaction with INH. Furthermore, these results may aid in the future development of drugs based on INH and its target protein katG for TB treatment.

Funding

This work was supported by the University Research Coordination Office of De La Salle University - Manila, Philippines (DLSU-URCO).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tube.2018.11.005.

References

[1] WHO. Global tuberculosis report 2017. 2017www.who.int/tb/publications/global_report/en/.
[2] Smith I. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. Clin Microbiol Rev 2003;16:463–96. https://doi.org/10.1128/CMR.16.3.463-496.2003.
[3] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 1998;393:537–44. https://doi.org/10.1038/31159.
[4] Dias MVB, Vasconcelos IB, Prado AMX, Fadel V, Basso LA, de Azevedo WF, et al. Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis. J Struct Biol 2007;159:369–80. https://doi.org/10.1016/j.jsb.2007.04.009.
[5] Musser JM, Kapur V, Williams DL, Kreiswirth BN, van Soolingen D, van Embden JDA. Characterization of the catalase-peroxidase gene (katG) and inhA locus in isoniazid-resistant and -susceptible strains of Mycobacterium tuberculosis by automated DNA sequencing: restricted array of mutations associated with drug resistance. J Infect Dis 1996;173:196–202. https://doi.org/10.1093/infdis/173.1.196.
[6] Bertrand T, Eady NAJ, Jones JN, Jesmin, Nagy JM, Jamart-Grégoire B, et al. Crystal structure of Mycobacterium tuberculosis catalase-peroxidase. J Biol Chem 2004;279:38991–9. https://doi.org/10.1074/jbc.M402382200.
[7] Jena L, Waghmare P, Kashikar S, Kumar S, Harinath BC. Computational approach to understanding the mechanism of action of isoniazid, an anti-TB drug. Int J Mycobacteriology 2014;3:276–82. https://doi.org/10.1016/j.ijmyco.2014.08.003.
[8] Unissa AN, Subbian S, Hanna LE, Selvakumar N. Overview on mechanisms of isoniazid action and resistance in Mycobacterium tuberculosis. Infect Genet Evol 2016;45:474–92. https://doi.org/10.1016/j.meegid.2016.09.004.
[9] Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. ethA, inhA, and katG loci of ethionamide-resistant clinical Mycobacterium tuberculosis isolates. Antimicrob Agents Chemother 2003;47:3799–805. https://doi.org/10.1128/AAC.47.12.3799-3805.2003.
[10] Srivastava G, Tripathi S, Kumar A, Sharma A. Molecular investigation of active binding site of isoniazid (INH) and insight into resistance mechanism of S315T-MtKatG in Mycobacterium tuberculosis. Tuberculosis 2017;105:18–27. https://doi.org/10.1016/j.tube.2017.04.002.
[11] Pierattelli R, Banci L, Eady NAJ, Bodiguel J, Jones JN, Moody PCE, et al. Enzyme-catalyzed mechanism of isoniazid activation in class I and class III peroxidases. J Biol Chem 2004;279:39000–9. https://doi.org/10.1074/jbc.M402384200.
[12] Kohn W, Becke AD, Parr RG. Density functional theory of electronic structure. J Phys Chem 1996;100:12974–80. https://doi.org/10.1021/jp960669l.
[13] Giustino F. Materials modelling using density functional theory properties and predictions. UK: Oxford University Press; 2014.
[14] Jain A, Shin Y, Persson KA. Computational predictions of energy materials using density functional theory. Nat Rev Mater 2016;1:15004. https://doi.org/10.1038/natrevmats.2015.4.
[15] Yin W-J, Yang J-H, Kang J, Yan Y, Wei S-H. Halide perovskite materials for solar

cells: a theoretical review. J Mater Chem 2015;3:8926–42. https://doi.org/10.1039/C4TA05033A.
[16] Franco FC, Padama AAB. DFT and TD-DFT study on the structural and optoelectronic characteristics of chemically modified donor-acceptor conjugated oligomers for organic polymer solar cells. Polymer 2016;97:55–62. https://doi.org/10.1016/j.polymer.2016.05.025.
[17] Franco FC, Padama AAB. On the structural and optoelectronic properties of chemically modified oligothiophenes with electron-withdrawing substituents for organic solar cell applications: a DFT/TDDFT study. J Phys Soc Japan 2017;86:064802. https://doi.org/10.7566/JPSJ.86.064802.
[18] Tang Q, Zhou Z, Chen Z. Innovation and discovery of graphene-like materials via density-functional theory computations. Wiley Interdiscip Rev Comput Mol Sci 2015;5:360–79. https://doi.org/10.1002/wcms.1224.
[19] Escorihuela J, Das A, Looijen WJE, van Delft FL, Aquino AJA, Lischka H, et al. Kinetics of the strain-promoted oxidation-controlled cycloalkyne-1,2-quinone cycloaddition: experimental and theoretical studies. J Org Chem 2018;83:244–52. https://doi.org/10.1021/acs.joc.7b02614.
[20] kumar Pandey A, Bajpai A, Baboo V, Dwivedi A. Structural, electronic, and vibrational properties of isoniazid and its derivative N-Cyclopentylidenepyridine-4-carbohydrazide: a quantum chemical study. J Theor Chem 2014:1–15. https://doi.org/10.1155/2014/894175. 2014.
[21] Stawoska I, Dudzik A, Wasylewski M, Jemioła-Rzemińska M, Skoczowski A, Strzałka K, et al. DFT-based prediction of reactivity of short-chain alcohol dehydrogenase. J Comput Aided Mol Des 2017;31:587–602. https://doi.org/10.1007/s10822-017-0026-5.
[22] Roos K, Hogner A, Ogg D, Packer MJ, Hansson E, Granberg KL, et al. Predicting the relative binding affinity of mineralocorticoid receptor antagonists by density functional methods. J Comput Aided Mol Des 2015;29:1109–22. https://doi.org/10.1007/s10822-015-9880-1.
[23] Czaja K, Kujawski J, Jodłowska-Siewert E, Szulc P, Ratajczak T, Krygier D, et al. On the interactions of fused pyrazole derivative with selected amino acids: DFT calculations. J Chem 2017;2017:1–9. https://doi.org/10.1155/2017/8124323.
[24] Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, et al. General atomic and molecular electronic structure system. J Comput Chem 1993;14:1347–63. https://doi.org/10.1002/jcc.540141112.
[25] Grimme S. Density functional theory with London dispersion corrections. Wiley Interdiscip Rev Comput Mol Sci 2011;1:211–28. https://doi.org/10.1002/wcms.30.
[26] Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 2010;132:154104. https://doi.org/10.1063/1.3382344.
[27] Grimme S, Ehrlich S, Goerigk L. Effect of the damping function in dispersion corrected density functional theory. J Comput Chem 2011;32:1456–65. https://doi.org/10.1002/jcc.21759.
[28] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 2009;30:2785–91. https://doi.org/10.1002/jcc.21256.
[29] Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminf 2012;4:17. https://doi.org/10.1186/1758-2946-4-17.
[30] Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 1996;17:490–519. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.
[31] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng Des Sel 1995;8:127–34. https://doi.org/10.1093/protein/8.2.127.
[32] Maxwell PI, Popelier PLA. Unfavorable regions in the ramachandran plot: is it really steric hindrance? The interacting quantum atoms perspective. J Comput Chem 2017;38:2459–74. https://doi.org/10.1002/jcc.24904.
[33] Kruse H, Grimme S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. J Chem Phys 2012;136:154101. https://doi.org/10.1063/1.3700154.
[34] Grimme S, Kruse H. gCP-D3 Webservice. 2016http://wwwtc.thch.uni-bonn.de/.
[35] Bode BM, Gordon MS. Macmolplt: a graphical user interface for GAMESS. J Mol Graph Model 1998;16:133–8. https://doi.org/10.1016/S1093-3263(99)00002-9.
[36] Glendening ED, Landis CR, Weinhold F. NBO 6.0 : natural bond orbital analysis program. J Comput Chem 2013;34:1429–37. https://doi.org/10.1002/jcc.23266.
[37] Wallace AC, Laskowski RA. LIGPLOT v.4.5.3. - operating manual. 2003. Retrieved April 28, 2018, from https://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/manual/index.html.
[38] Philips JJ, Hudspeth MA, Browne PM, Peralta JE. Basis set dependence of atomic spin populations. Chem Phys Lett 2010;495:146–50. https://doi.org/10.1016/j.cplett.2010.06.046.
[39] Suarez J, Ranguelova K, Jarzecki AA, Manzerova J, Krymov V, Zhao X, et al. An oxyferrous heme/protein-based radical intermediate is catalytically competent in the catalase reaction of Mycobacterium tuberculosis catalase-peroxidase (KatG). J Biol Chem 2009;284:7017–29. https://doi.org/10.1074/jbc.M808106200.
[40] Gasselhuber B, Graf MMH, Jakopitsch C, Zamocky M, Nicolussi A, Furtmüller PG, et al. Interaction with the redox cofactor MYW and functional role of a mobile arginine in eukaryotic catalase-peroxidase. Biochemistry 2016;55:3528–41. https://doi.org/10.1021/acs.biochem.6b00436.
[41] Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation Database. PLoS Med 2009;6:e1000002. https://doi.org/10.1371/journal.pmed.1000002.
[42] Rouse DA, DeVito JA, Li Z, Byer H, Morris SL. Site-directed mutagenesis of the katG gene of Mycobacterium tuberculosis : effects on catalase-peroxidase activities and isoniazid resistance. Mol Microbiol 1996;22:583–92. https://doi.org/10.1046/j.1365-2958.1996.00133.x.