Heliyon 7 (2021) e08482

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification of potent inhibitors of ATP synthase subunit c (AtpE) from *Mycobacterium tuberculosis* using *in silico* approach

Mustafa Alhaji Isa [a,*], Mustapha B. Abubakar [b], Mohammed Mustapha Mohammed [a], Muhammad Musa Ibrahim [a], Falmata Audu Gubio [a]

[a] Department of Microbiology, Faculty of Sciences, University of Maiduguri, P.M.B. 1069, Maiduguri, Nigeria
[b] Department of Veterinary Microbiology, Faculty of Veterinary Medicine, University of Maiduguri, Nigeria

ARTICLE INFO

Keywords:
MTB
AtpE
ADME
Homology modeling
MD Simulation

ABSTRACT

ATP synthase subunit c (AtpE) is an enzyme that catalyzes the production of ATP from ADP in the presence of sodium or proton gradient from *Mycobacterium tuberculosis* (MTB). This enzyme considered an essential target for drug design and shares the same pathway with the target of Isoniazid. Thus, this enzyme would serve as an alternative target of the Isoniazid. The three dimensional (3D) model structure of the AtpE was constructed based on the principle of homology modeling using the Modeller9.16. The developed model was subjected to energy minimization and refinement using molecular dynamic (MD) simulation. The minimized model structure was searched against Zinc and PubChem database to determine ligands that bind to the enzyme with minimum binding energy using RASPD and PyRx tool. A total of 4776 compounds capable of bindings to AtpE with minimum binding energy were selected. These compounds further screened for physicochemical properties (Lipinski rule of five). All the compounds that possessed the desirable property selected and used for molecular docking analysis. Five (5) compounds with minimum binding energies ranged between −8.69, and −8.44 kcal/mol, less than the free binding energy of ATP were selected. These compounds further screened for the absorption, distribution, metabolism, excretion, and toxicity (ADME and toxicity) properties. Of the five compounds, three (ZINC14732869, ZINC14742188, and ZINC12205447) fitted all the ADME and toxicity properties and subjected to MD simulation and Molecular Mechanics Generalized Born and Surface Area (MM-GBSA) analyses. The results indicated that the ligands formed relatively stable complexes and had free binding energies, less than the binding energy of the ATP. Therefore, these ligands considered as prospective inhibitors of MTB after successful experimental validation.

1. Introduction

Tuberculosis (TB) is among the major infectious diseases that responsible for mortality and morbidity worldwide. The disease has numerous available drugs for its treatment yet claims the lives of countless individuals [1, 2]. This is due to the endemic effects of multi-drug resistant TB (MDR-TB), extensively drug-resistant TB (XDR-TB), and totally drug-resistant TB (TDR-TB), which pose a serious menace to the tuberculosis control program [3]. India represents more than one-fourth of the world's TB cases and death [3]. Therefore, it is essential to develop new antituberculosis drugs which can inhibit both actively multiplying bacilli and a non-growing persistent population of *Mycobacterium tuberculosis* (MTB) to prevent reactivation of the infection. Isoniazid (INH) is the first-line antibiotic for the treatment of all types of TB caused by MTB. It was discovered in 1912 and successfully used for the treatment of TB in 1951. The drug acts by inhibiting the cell wall synthesis of the organisms. Initially, the compound is a prodrug, but it would activate by the enzymes called Catalase-peroxidase encoded by the *katG* gene. The activated drug form complex with NADH, thus inhibit the fatty acid synthesis of the bacterium, which is one of the significant components of the bacterium cell wall. Therefore, the drug has substantial early bactericidal action against quickly multiplying cells [4, 5]. Isoniazid, combined with rifampicin, has long been used for the treatment of TB. It is active against metabolically active multiplying or replicating bacilli. INH resistance is the most frequently occurring in MTB at a rate of 10^{5-6} frequency in vitro as compared to other tuberculosis drugs. The two critical molecular mechanisms for the drug resistance in INH commonly attributed to a mutation in the *katG* and *inhA* gene or its

* Corresponding author.
E-mail address: mustafaisa@unimaid.edu.ng (M.A. Isa).

https://doi.org/10.1016/j.heliyon.2021.e08482
Received 16 February 2021; Received in revised form 11 April 2021; Accepted 22 November 2021

promoter region. Indeed, a lot of researchers have conducted which demonstrated that these two genes mainly associated with INH resistance. Among these genes, S315T in *katG* is the most commonly found and accounts for 50–90% of all the INH-resistant found in clinical isolates [6, 7]. This mutation is associated with higher levels of resistance at MIC $>$ 1 μg/mL to INH [8] and appeared consistently in MDR-TB strains [7]. This process followed by a mutation in the promoter region of *inhA* which causes overexpression of *inhA* or a decrease in the binding affinity of the INH-NAD adducts [9]. However, ATP synthase subunit c (AtpE) has the same pathway as the target of Isoniazid [10, 11, 12]. Thus, this enzyme would serve as an alternative target of the Isoniazid since it plays a vital role by providing ATP during the dormancy state of the MTB [13]. ATP synthase subunit c is an enzyme that catalyzes the production of ATP from ADP in the presence of sodium or proton gradient. ATP synthase plays an important role in human wellbeing. Failure of this enzyme has been associated with a wide range of illnesses includes tuberculosis, neuropathy, Alzheimer's, and Parkinson's [14]. ATP synthase is also considered a drug target in the treatment of infection caused by *Streptococcus mutans* (The bacteria that causes dental caries via biofilm formation and acid production. Targeting ATP synthase from *S. mutans* inhibits biofilm formation and acid production [15]. Similarly, in *Mycobacterium* tuberculosis, mutations in the c-subunit (D32V and A63P) of the enzyme cause resistance to the tuberculosis drug diarylquinoline [16]. Therefore, this enzyme is considered a drug target in the curing diseases such as tuberculosis, heart disease, immune deficiency, cystic fibrosis, diabetes, ulcers, Parkinson's, and Alzheimer's [14, 17]. The enzyme (F-type ATPases) has two structural domains F1 and FO domains. F1 domain had the extramembranous catalytic core, while the FO domain had membrane proton channel joined together by the peripheral stalk and central stalk. The catalytic domain of F1 joins through a rotary mechanism during the catalytic process to the central stalk sections of proton translocation. Residues between the 10–14 subunits serve as homomeric c-ring and form the central rotor element of the F1. FO consists of the residues between 5–25 and 57–77 in its domain. However, ATP synthase subunit c is conserved in human beings, because the subtle difference between human and bacterial make it an attractive target for drug design and development [12].

Several compounds (natural and synthetic) have been identified to inhibit *E. coli* ATP synthase (Zheng and Ramirez, 2000). These include: 7-chloro-4-nitrobenzo-2-oxa-1, 3-diazole (NBD-Cl), sodium azide (NaN_3), aluminum fluoride (AlFx), scandium fluoride (ScFx), beryllium fluoride (BeFx), and dicyclohexylcarbodiimide (DCCD). Also, many drugs from the natural origin such as oligomycin, efrapeptins, aurovertins, leucinostatins, with some polyphenols like resveratrol, piceatannol, quercetin, morin, and epicatechin were known to inhibit *E. coli* ATP synthase [18]. However, no compounds were identified to inhibit the *M. tuberculosis* ATP synthase. These justify the reason for the selection of the enzyme mentioned above in this study.

2. Methods

2.1. Retrieval of sequence and search of templates

The sequence of AtpE from MTB retrieved from the National Centre for Biotechnological Information (NCBI). The retrieved AtpE sequence was subjected to Basic Local Alignment Search Tool (BLASTP) [19] against the Protein Data Bank (PDB) [20]. This step was carried out to identify an appropriate template for the homology modeling. After identifying the suitable template, the AtpE sequence, and the template were aligned using ClustalW [21] to obtained sequence identity and similarity using the default parameter.

2.2. Homology modeling of AtpE

Homology modeling was used to build the three-dimensional structure (3D) of the AtpE based on the principle of spatial strain using Modeller9.16 [22]. The process starts with the alignment of the target and the template sequence, then all the information of the template obtained during alignment was transferred to the target sequence. This information includes hydrogen bonds, main chain, side chain, and dihedral angle to build the 3D structure of AtpE. Ten (10) 3D model structure of the AtpE obtained and the one with least Discrete Optimized Protein Energy (DOPE) value was selected for energy minimization and evaluation.

2.3. Energy minimization and model evaluation

The 3D model of AtpE with the least Discrete Optimized Protein Energy (DOPE) value was selected and subjected to energy minimization and refinement using molecular dynamic simulation analysis for 10 ns via AMBERTOOLS10 [23]. The energy minimization and refinement was carried out to stabilize the model structure before the molecular docking studies. The minimized model structure was superposed to the template to determine their deviation based on the root mean square deviation (RMSD) value. Finally, the minimized model structure was assessed to determine the stereochemistry quality using the Ramachandran plot, ERRAT [24], and Verify_3D [25].

2.4. Selection of ligand for docking analysis

The model structure of AtpE complexed with ligand was submitted to the RASPD tool [26] to obtain ligands capable of binding to the protein. RASPD tool can identify ligands from Zinc database capable of binding to the protein complex with minimum binding energy. A total of four thousand seven hundred and seventy-six (4776) compounds were obtained and further validated their binding energies using PyRx program. The compounds with better binding energies were selected and screened for the Lipinski rule of five (Molecular weight ($\leq$500 Da), Log P ($\leq$5), hydrogen bond donor ($\leq$5), and hydrogen bond acceptor ($\leq$10)) to determine compounds with desirable physicochemical properties using DataWarrior program.

2.5. Molecular docking studies

Molecular docking studies were carried out to determine the interaction between the AtpE and the selected ligands using an AutoDock4.2 tool [27]. The process begins with the conversion of both AtpE and the ligands to the PDBQT type file, which contained all the information of the individual atoms such as atom type, partial charges, and the torsional degree of freedom. The gasteiger charges [28] calculated and the free binding energy was determined using Lamarckian genetic algorithms. The x, y, and z coordinate of the AtpE set, and the grid map was fixed at 60 $\times$ 60 $\times$ 60 with a spacing of 0.375 Å. Finally, the RMSD calculated, and the protein-ligand complex was analyzed using Pymol [29] and Ligplot + program [30, 31].

2.6. ADME and toxicity analysis

Absorption Distribution, Metabolism, Excretion, and Toxicity (ADME and Toxicity) analysis of all the selected ligands with good binding energies were carried out to determine compounds with desirable pharmacokinetic properties using ADME/TOX tool [32, 33], AdmetSAR tool [34], and DataWarrior [35]. The pharmacokinetic properties predicted in this study include Human Intestinal Absorption (HIA), Blood-Brain Barrier (BBB) penetration, Cytochrome P450 (CYP450 2D6) Inhibitor, Plasma Protein Binding (PPB), Mutagenicity, Tumorigenicity, Irritation, and Reproduction. All the ligands with suitable pharmacokinetic properties selected for the MD simulation analysis.

2.7. Molecular dynamic (MD) simulation analysis

The MD simulation analysis was carried out to determine the stability of the ligands at the binding pocket of the AtpE. In this study, all the ligands with the desirable ADME and toxicity properties were selected and subjected to MD simulation analysis using AMBERTOOLS10 package [23]. During the process, the protonate 3D and the antechamber were used for the addition of explicit hydrogen and missing parameter to the ligands respectively. The force field of the protein assigned with GAFF, while the ligand force field assigned with ff12SB. The coordinate and topology file of the protein-ligand complex was constructed using the tleap component of the Amber tool. The complex system was neutralized using a buffer solution of 10 Å within the octahedral box of TIP3P water. The system was minimized to remove structural artifact occurred during the model building. Also, further minimization of 2500 steps of steepest descent and 2500 steps of the conjugate gradient was carried out without any restrained. The initial and final temperature of 0 and 300 k respectively were used for heating the system using Langevin dynamics temperature regulator. Finally, the production of the simulation performed at constant temperature and pressure of 300 k and 1 atm using the time step of 2 fs. The analysis of the root mean square deviation (RMSD), root mean square fluctuation (RMSF), the radius of gyration was carried out using the PTRAJ component of the AMBERTOOL10. Also, Molecular Mechanics Generalized Born and Surface Area (MM-GBSA) analysis were carried out to determine the free binding energy of the protein-ligand complex using the MD simulation trajectory of the last 5 ns.

2.8. Molecular Mechanics Generalized Born and Surface Area (MM-GBSA) analysis

The MM-GBSA is a vital tool to use in quantitative measurement of the protein-ligand interactions [36]. The method uses the principle of MD simulation to determine the free binding energy of the protein-ligand complex, and it achieves a tremendous success over the years [37]. In this study, Amber14 was used for the MM-GBSA analysis to calculate the free binding energy of the DHQase-ligands complexes, based on the average of 500 snapshots obtained at every 10 ps from the molecular dynamics simulation trajectory of the last 5 ns. The process summarized in the below equation.

$$\Delta G_{binding} = G_{complex} - (G_{receptor} + G_{ligand}) \quad (1)$$

$$G_x = E_{MM} + G_{solv} - T\Delta S \quad (2)$$

$$E_{MM} = E_{vdW} + E_{ele} \quad (3)$$

$$G_{solv} = G_{polar} + G_{nonpolar} \quad (4)$$

See Ref. [38],

$$\Delta G_{\text{MM-GBSA}} = \Delta G_{\text{vdw}} + \Delta G_{\text{elec}} + \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \quad (5)$$

The total binding energy was determined from each snapshot based on the difference between the free binding energy of the complex ($G_{complex}$), receptor ($G_{receptor}$) and the ligand (Gl_{igand}), as shown in Eq. (1). The free binding of each component (G_x) was calculated using the sum of configurational entropy $T\Delta S$, the sum of solvation binding energy (G_{sol}), and the sum of molecular mechanical gas-phase free binding energy (E_{MM}) (Eq. (2)). The sum of the molecular mechanical gas-phase free binding energy (E_{MM}) further gave rise to the van der Waals energy (E_{vdW}) and Gas-phase electrostatic energy (E_{ele}) (Eq. (3)), while polar (Gpolar) and nonpolar (Gnonpolar) were derived from the solvation free energy (Gsolv) (Eq. (4)). Lastly, the total free binding energy was calculated based on the values of the gas-phase electrostatic energy (E_{ele}), van der Waals (E_{vdW}) polar (G_{polar}) and nonpolar ($G_{nonpolar}$) component (Eq. (5)).

3. Results and discussion

3.1. Homology modeling

The sequence of AtpE was retrieved from NCBI with accession number NP_215821.1. The sequence was subjected to BLASTP against proteins with known 3D structures in the PDB. Six protein templates (4V1F, 4MJN, 3ZK1, 3V3C, 2WIE, and 2W5J) were selected based on the sequence identity, sequence similarity and high statistical significance (Less e-value). But, 4V1F further chosen due to high-resolution and used as a template for homology modeling (predicting the 3D structure of protein using similar template). The result of pair sequence alignment between the template (4V1F) and the AtpE revealed that their sequences share 90.1% (73/81aa) sequence identity, 93.8% (76/81aa) sequence similarity, and 2.5% (2/81aa) gap (Figure 1). The presence of the high similarity between the AtpE and the template is a clear indication that their structures are highly conserved. It is therefore used as a template to build the modeled structure of the AtpE. During the model building, all the coordinates of the template structure such as structurally variable regions (SVRs), structurally conserved regions (SCRs), and N, and *C termini* were all transferred to the AtpE via satisfaction of spatial restraints. All the main chain and side chain were set using rotamers in the modeled structure (Figure 2a). The process generated ten modeled structures of AtpE, and the ones with minimum Discrete Optimized Protein Energy (DOPE) were selected. The selected modeled structure was subjected to energy minimization (Van der Waals repulsion energy and steric clashes) using MD simulation with AMBERTOOL10. The final structure after minimization used for structural superimposition with the Cα trace of the template (4V1F). The RMSD between the Cα trace of the template and the modeled structure is 0.578 Å. This RMSD value showed that the modeled structure was highly reliable resembling the template structure (Figure 2b). Analysis of the structural quality of the modeled structure, final minimized structure, and the template was carried out using PROCHECK. The program analyzes their stereochemistry quality based on phi and psi (Φ-ψ) plot, residue-by-residue geometry, overall structure geometry and G-factor. Also, Verify_3D and ERRAT Quality Factor for the template, modeled structure, and the final minimized structure were determined. ERRAT [24] was used to analyze and assess the statistics between different atom types based on non-bonded interaction. Verify_3D [25] was used to ascertain the compatibility of the 3D model with its amino acid sequence (ID) by giving a structural class to alpha, beta, loop, polar, nonpolar, etc. and comparing the output with proper structures. The result of the Ramachandran map of the modeled structure revealed that 91.8% of the entire residues were in the most favorable region [A, B, L]. Then followed by 6.6% in the additional allowed region [a, b, l, p], 1.6% and 0% in the generously allowed region [~a, ~b, ~l, ~p] and the disallowed region [XX] respectively. On the other hand, the minimized modeled structure and the template had 98.4% and 93.2% of all their residues in the most favorable regions [A, B, L]. However, 0% and 6.8% in additional allowed region [a, b, l, p] and 1.6% and 0% in the generously allowed region [~a, ~b, ~l, ~p] respectively (Table 4). Both the modeled and minimized modeled structures had excellent and acceptable quality with high structural integrity since a good quality model must have at least 90% in the most favorable region based on the analysis of 118 structures with 2 Å resolution and ≤20 R factor. G-factor was used to measure how unusual the 3D structure was if the overall G-factor threshold was < -0.5, the structure was considered as unusual, and if it was < -1.0 regarded as highly unusual. Errat results indicated the percentage of non-bonded interactions between various atoms, in which the calculated error values fell below the rejection limit of 95%. The Errat score of the modeled structure (before minimization) was 97.260%, while after minimization the score increased to 98.630%. It showed that the

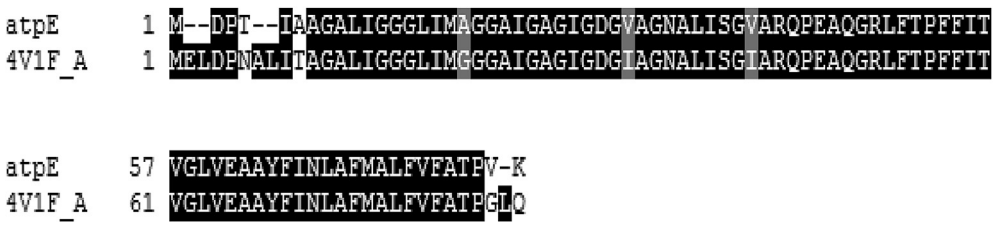

Figure 1. Pair sequence alignment between AtpE and the Mycobacterial ATP synthase rotor ring showing the conserved regions in the black color.

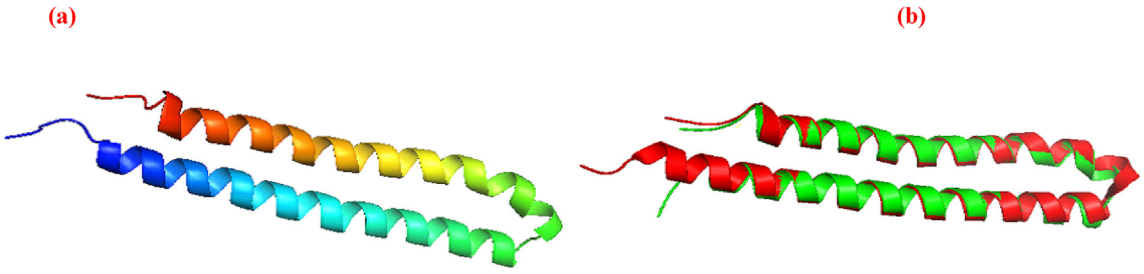

Figure 2. The 3D modeled structure of AtpE and the structural superimposition of the modeled structure and the template. (a) The Modeled structure of AtpE (b) Structural superimposition of the Cα traces of template (4V1F) (red) and the modeled structure (green) with the RMSD of 0.121 Å.

Table 1. Structural evaluation of the template (4V1F), modeled structure of AtpE and the minimized modeled structure of AtpE.

Proteins	Procheck				ERRAT Quality Factor (%)	Verify_3D (%)	Total Number of steric clashes	VDW repulsion energy (kcal/mol)
	Most favored	Additional allowed	Generously allowed	Disallowed				
AtpE	91.80	6.60	1.60	0.0	97.26	33.33	62	63.90
Minimized AtpE	98.40	0.00	1.60	0.0	98.63	34.57	49	30.49
4V1F	93.20	6.80	0.00	0.0	99.14	23.26	49	30.49

modeled structure became more stable after minimization (Table 1). Also, the number of steric clashes decreased from 62 to 49 after minimization. Similarly, the van der Waals repulsion energy also decreased from 63.8962 kcal/mol to 30.4879 kcal/mol. The decrease in the value of van der Waals repulsion energy in the minimized modeled structure accompanied by an increase in the stability (Table 1).

3.2. Molecular docking analysis

Molecular docking is a technique used in biological testing of lead molecules. It helps in identifying compounds that have the potential to serve as drugs in the future. This process depends on the capability of compounds to bind to the macromolecule (mostly protein) with minimum binding energy and form a stable complex. RASPD and PyRx were used to screened compounds from the large compound database (Zinc and PubChem database). A total of 4776 compounds capable of binding to AtpE with minimum binding energies (negative delta G values) were selected. These compounds further screened for physicochemical properties: Molecular weight (≤500), number of hydrogen bond acceptor (≤10), number of hydrogen bond donor (≤5) and LogP (≤5)) and drug-likeness (Table 2). All the compounds that possessed the desirable properties were used for molecular docking analysis, to determine their binding energies with AtpE. Five (5) compounds have more negative

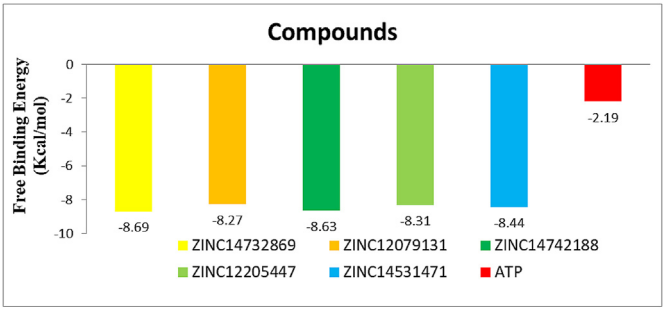

Figure 3. Distribution different free binding energies of the selected ligands Interacted with AtpE.

Table 2. Molecular properties and drug-likeness of the selected ligands interacted with AtpE.

S/No.	Zinc Code	Molecular Weight	cLogP	H-bond Acceptors	H-bond Donors	Drug-likeness
1	ZINC14732869	499.68	0.85	7	1	7.40
2	ZINC12079131	490.63	2.11	8	1	–3.45
3	ZINC14742188	486.64	0.26	7	1	6.76
4	ZINC12205447	488.61	0.53	8	1	6.51
5	ZINC14531471	494.70	0.84	7	2	6.20

binding energies than ATP. This indicates stronger interactions between the ligands and AtpE at the binding pocket (Figure 3). Based on the docking result, ZINC14732869 had the minimum binding energy of −8.69 kcal/mol and observed to interacted with AtpE via a hydrogen bond by accepting electrons from the hydroxyl group of Thr56 (distance = 2.72 Å). Also, it exhibited hydrophobic interactions with Ala34, Ile55, Leu35, Pro52, Phe53, Gly27, Ala31, Val30 and Leu49 (Table 3) (Figure 4a). All these residues mentioned above involved in both hydrogen and hydrophobic interactions formed the significant part of the catalytic domain (FO). Therefore, inhibition of these residues by the ligand would block the catalytic activity of the AtpE. This action would in turns affect the growth of the organism. Similarly, ZINC12079131 had the minimum binding energy of −8.27 kcal/mol and interacted and formed two hydrogen bonds with the hydroxyl group of Thr56 (distance$_1$ = 2.65 Å and distance$_2$ = 2.90 Å). It also presented hydrophobic interactions with Gly27, Val30, Ala31, Ala34, Leu35, Leu49, Pro52 and Phe53 (Figure 4b). ZINC14742188 exhibited minimum binding energy of −8.63 kcal/mol and interacted with AtpE and formed hydrogen bond by accepting electrons from the hydroxyl group of Thr56 (distance = 3.09 Å). Also, it exhibited hydrophobic interactions with Gly27, Val30, Ala31, Ala34, Leu35, Leu49, Pro52, Ile55 and Phe53 (Figure 4c). Lastly, ZINC14531471 had a binding affinity of −8.44 kcal/mol and interacted and formed two hydrogen bonds with Asn33 (distance$_1$ = 3.19 Å and distance$_2$ = 3.34 Å). Besides, it presented hydrophobic interactions with Val30, Ala31, Ala34, Leu35, Leu49, Phe53, Ile55 and Thr56 (Figure 4e). All these residues mentioned above formed the catalytic domain (F1) of AtpE. Therefore, the interactions of the ligands with these residues interfere with the growth of the MTB and subsequently led to its death (Table 3).

3.3. ADME and toxicity analyses

The result of the molecular docking analysis revealed that five compounds (ZINC14732869, ZINC12079131, ZINC14742188, ZINC12205447, and ZINC14531471) had good binding affinities. These compounds analyzed for ADMET properties such as Human Intestinal Absorption (HIA), Blood-Brain Barrier (BBB), Cytochrome P450 (CYP450 2D6) inhibitor, Aqueous Solubility, Plasma Protein Binding (PPB), Ames test, Carcinogens, Mutagenicity, Tumorigenic, Reproducibility, and Irritability (Table 4). Human Intestinal absorption and Blood-Brain Barrier of all the compounds were positive except ZINC14531471, which was contrary because it was unable to cross the Blood-Brain Barrier. Similarly, all the compounds were found to be non-inhibitors of CYP450 2D6 except ZINC14531471 which had an AC50 greater than 57 μM (>57 μM) based on the model calculation of Cheng *et al.* [29]. Concerning toxicity parameters (Mutagenicity, tumorigenic, reproducibility and irritability), the entire compound was predicted to be non-toxic except

Table 3. Docking score of ligands interacted with AtpE.

S/No.	Zinc Code	Minimum Free Energy of Binding (kcal/mol)	Interacting Residues	Distance (Å)	Residues involved in hydrophobic interaction
1.	ZINC14732869	−8.69	Thr56	2.72	Ala34, Ile55, Pro52, Phe53, Leu59, Val60, Gly23, Gly27, Ala31, Val130, Leu49, Leu35
2.	ZINC12079131	−8.27	Thr56 Thr56	2.65 2.90	Pro52, Ala34, Leu35, Leu49, Phe53, Leu59, Val60, Gly27, Val27, Val130, Ala31
3.	ZINC14742188	−8.63	Thr56	3.09	Leu35, Leu49, Phe53, Leu59, Val30, Gly23, Val60, Leu59, Ile55, Pro52, Ala31, Gly27
4.	ZINC12205447	−8.31	Thr56	2.73	Leu49, Ile55, Pro52, Ala31, Gly27, Leu59, Val160, Phe53, Leu35, Ala34
5.	ZINC14531471	−8.44	Asn33 Asn33	3.19 3.34	Leu59, Thr56, Val30, Phe53, Ala31, Ala34, Leu49, Leu35, Pro52, Ile55

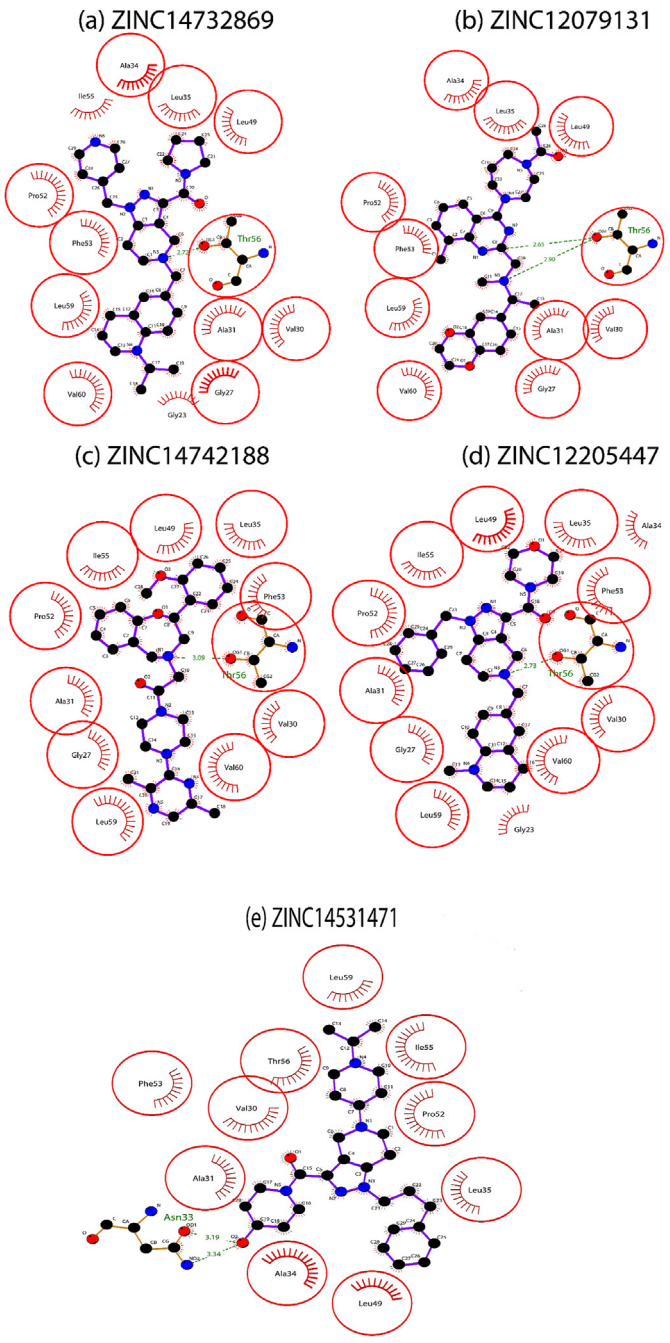

Figure 4. Interactions of AtpE residues with the selected ligands (a) ZINC14732869 (b) ZINC12079131 (c) ZINC14742188 (d) ZINC12205447 (e) ZINC14531471.

Table 4. ADME and toxicity analyses of selected ligands interacted with AtpE.

S/No	Compounds	HIA	BBB	CYP450 2D6 Inhibitor	PPB (%)	Aqueous Solubility	AMES Test	Carcinogens	Mutagenic	Tumorigenic	Reproducibility	Irritant
1	ZINC14732869	+	+	Non-inhibitor	71.07	–3.274	Non AMES toxic	Non-carcinogens	none	none	none	none
2	ZINC12079131	+	+	Non-inhibitor	44.78	–3.513	Non AMES toxic	Non-carcinogens	none	none	none	high
3	ZINC14742188	+	+	Non-inhibitor	65.45	–2.772	Non AMES toxic	Non-carcinogens	none	none	none	none
4	ZINC12205447	+	+	Non-inhibitor	44.60	–2.896	Non AMES toxic	Non-carcinogens	none	none	none	none
5	ZINC14531471	+	-	Inhibitor	11.87	–2.769	Non AMES toxic	Non-carcinogens	none	none	none	none

BBB = Blood-Brain Barrier, HIA = Human Intestinal Absorption, PPB = Plasma Protein Binding, Aqueous Solubility = Insoluble < –10 < Poorly soluble < –6 < Moderately soluble < –4 < Soluble < –2 < Very soluble <0 < Highly soluble.

ZINC12079131 which was found to be highly irritating (Table 4). However, one compound (ZINC14531471) was found to be non-toxic, but was unable to cross the Blood-Brain Barrier and was also found to be a CYP450 2D6 inhibitor. Of the five compounds, only three (ZINC14732869, ZINC14742188, and ZINC12205447) fitted all the ADME and toxicity properties. Therefore, these compounds selected for the MD simulation analysis (Table 4).

3.4. Molecular dynamic simulation analysis

From the results of the docking studies, five compounds with minimum binding energies were selected. These compounds further screened for ADME and toxicity properties. Three compounds (ZINC14732869, ZINC14742188, and ZINC12205447) were selected based on their pharmacokinetic properties. These compounds were subjected to the analysis of MD simulation to determine the structural stability of their complexes. Also, MD simulation of the AtpE complexed with ATP was carried out to compare its stability with the selected ligand. The stability of the four complexes (AtpE—ZINC14732869, AtpE—ZINC14742188, AtpE—ZINC12205447, and AtpE—ATP) was ascertained by carefully examining the RMSD during the 50 ns MD simulation. The deviations of the compound concerning their binding affinity as well as the motion of every residue within the complexes were determined based on the RMSF. The radius of gyration was also checked to determine the compactness of each complex based on the extent of how folded or unfolded the complex was. The AtpE—ZINC14732869 complex equilibrated at 5 ns and remained steady until fluctuated between 30 to 40 ns, but later stabilized throughout the 50 ns, with a mean value of 5.48 ± 0.06414 Å, maximum value of 9.77 Å and minimum value of 1.28 Å, higher than the mean RMSD value of AtpE—ATP complex (4.98 ± 0.04894 Å). The mean RMSD value of 5.48 ± 0.06414 Å suggested less flexibility probably due to interactions of the ligand with the flexible loop region of the AtpE, leading to high stability of the complex, although, it is less stable when compared to AtpE—ATP complex. Similarly, AtpE—ZINC14742188 and AtpE—ZINC12205447 complexes equilibrated and oscillated at 20 ns with the average mean values of 7.6037 ± 0.04346 Å and 6.5452 ± 0.06623, maximum values of 10.37 Å and 10.22 Å, and a minimum value of 1.72 Å and 1.13 Å respectively, higher than the mean RMSD value of AtpE—ATP complex (4.98 ± 0.04894 Å). The two (AtpE—ZINC14742188 and AtpE—ZINC12205447) complexes are less stable when compared with either AtpeE-ZINC14732869 or AtpE—ATP complex, although, it is not statistically significant ($p < 0.05$). However, both complexes had high RMSD values in most of their regions (Figure 5). The RMSF values of all the complexes (AtpeE—ZINC14732869, AtpE—ZINC14742188, and AtpE—ZINC12205447) shown in Figure 6. These represent the extent of movement of the initial position of each residue and atoms in the AtpE and the ligands. In the AtpeE—ZINC14732869 complex, all the residues of N-terminal, ω-loop, and C-terminal had RMSF values less than or equal

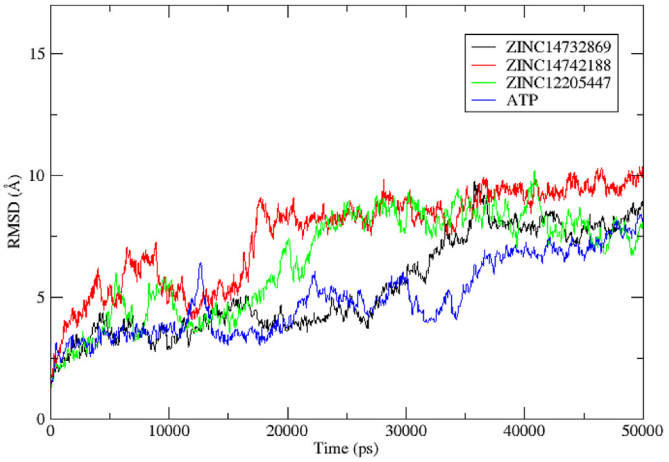

Figure 5. The MD simulation (RMSD analysis) of AtpeE—ZINC14732869, AtpE—ZINC14742188, AtpE—ZINC12205447, and AtpE—ATP complexes for 50 ns.

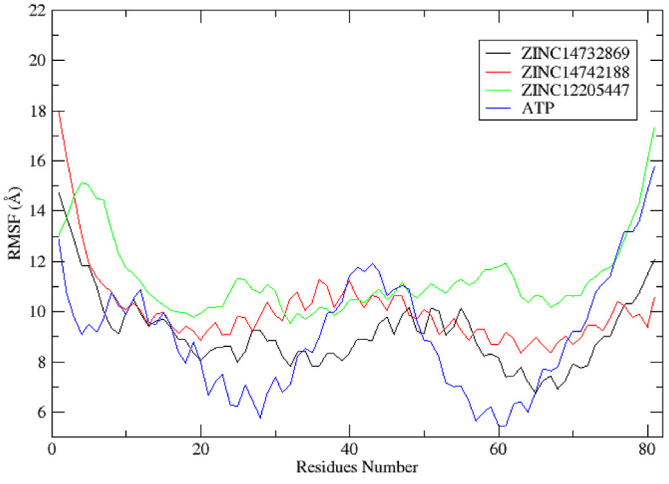

Figure 6. The MD simulation (RMSF analysis) of AtpeE—ZINC14732869, AtpE—ZINC14742188, AtpE—ZINC12205447, and AtpE—ATP complexes for 50 ns.

to 10 Å (≤10 Å), except residues between Met1—Ala6 and Val80—Lys81. The low RMSF values in all the regions of the complex occurred probably due to the interaction of the ZINC14732869 to the flexible loop region of AtpE. This result further strengthened the RMSD result, where the

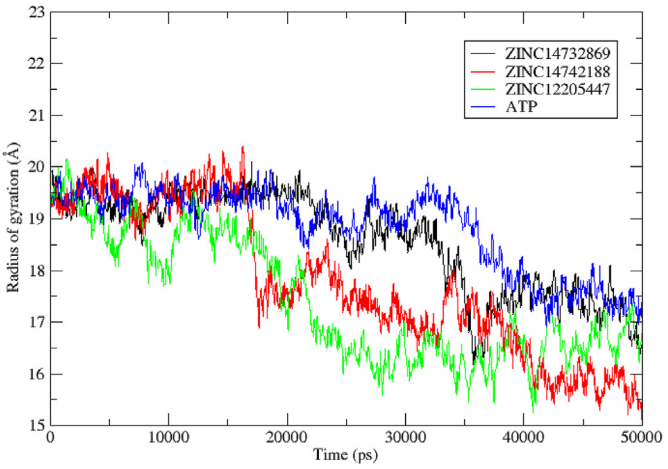

Figure 7. The MD simulation (Radius of gyration analysis) of AtpeE—ZINC14732869, AtpE—ZINC14742188, AtpE—ZINC12205447, and AtpE—ATP complexes for 50 ns.

4. Conclusion

A total of four thousand seven hundred and seventy-six (4776) compounds were obtained and further validated their binding energies using PyRx program. The selected ligands were used for the molecular docking analysis to determine the binding energy between the AtpE and the ligands. Five (5) compounds with minimum binding energies ranged between −8.69, and −8.44 kcal/mol, less than the free binding energy of ATP (−2.19 kcal/mol) were obtained. The compounds were further filtered for the ADME and toxicity properties. Of the five compounds, only three (ZINC14732869, ZINC14742188, and ZINC12205447) fitted all the ADME and toxicity properties. These compounds were subjected to MD simulation and MM-GBSA analyses. The results of the analyses show that all the ligands formed relatively stable complexes and had free binding energies, less than the binding energy of the ATP. Therefore, these ligands considered as prospective inhibitors of MTB after successful experimental validation.

Table 5. MM-GBSA analysis of the selected ligands.

Compounds	ΔG_{vdw}	ΔG_{ele}	ΔG_{polar}	$\Delta G_{nonpolar}$	$\Delta G_{MM\text{-}GBSA}$
ZINC14732869	−31.59 ± 0.3970	−29.83 ± 0.7662	40.60 ± 0.8820	−3.65 ± 0.0209	−24.47 ± 0.3135
ZINC14742188	−31.47 ± 0.3343	−3.37 ± 1.3207	11.72 ± 1.2036	−3.01 ± 0.0240	−26.13 ± 0.3176
ZINC12205447	−27.46 ± 0.3254	−12.03 ± 1.4011	19.55 ± 1.3813	−2.82 ± 0.0220	−22.76 ± 0.3194
ATP	−26.79 ± 0.5005	−9.79 ± 0.4811	32.73 ± 0.4812	−5.24 ± 0.0294	−9.09 ± 0.4646

complex had low RMSD values some of its regions. Similarly, the AtpE—ZINC14742188 complex had RMSF values of less than or equal to 10 Å (≤10 Å) in the N-terminal, C-terminal, and ω-loop regions except for residues between Met1—Ala7 and Ile36—Ser37, which had the RMSF value higher than 10 Å (>10 Å). These low RMSF values in the residues above were probably as a result of hydrophobic interactions between the residues and the ligand, which led to low RMSF values in the residues. Lastly, AtpE—ZINC12205447 complex residues located in N-terminal, C-terminal, and ω-loop regions had RMSF values of >10 Å, except Met17, Ala18, Gly19, Gly20, Gly32, Asn33, Ala34, Leu35 and Gly38 which had value < 10 Å (Figure 6). The radius of gyration of the complexes was analyzed to determine whether they were stably folded after the 50 ns MD simulation. All the complexes (AtpeE—ZINC14732869, AtpE—ZINC14742188, AtpE-ZINC12205447, and AtpE—ATP) fluctuated throughout the 50 ns with a mean value of 18.57954 ± 0.03009 Å, 17.7808 ± 0.03160 Å, 17.4353 ± 0.03786 Å, and 18.8310 ± 0.02508 Å respectively. Therefore, all the complex structures had unfolded polypeptide structures (Figure 7).

3.5. MM-GBSA analysis

The free binding energy of the protein-ligand complex was determined using MM-GBSA method implemented in Amber14. The analysis was carried out the using the average of 500 snapshots and at the interval of 10 ns of last 5 ns of the MD simulation trajectory. The energy used to determine the free binding energy include polar energy (G_{polar}), nonpolar ($G_{nonpolar}$) energy, van der Waals (E_{vdW}), and gas-phase electrostatic energy (E_{ele}), of the complexes. The result of the analysis showed that ZINC14742188 had favorable binding energy of −26.13 ± 0.3176 kcal/mol, followed by ZINC14732869 (−24.47 ± 0.3135 kca/mol) and ZINC12205447 (−22.76 ± 0.3194 kcal/mol), lower than the binding free energy of the ATP (−9.09 ± 0.4646 kcal/mol) (Table 5).

Declarations

Author contribution statement

Mustafa Alhaji Isa: Conceived and designed the experiments.
Mustapha B. Abubakar: Performed the experiments.
Mohammed Mustapha Mohammed; Falmata Audu Gubio: Analyzed and interpreted the data.
Muhammad Musa Ibrahim: Contributed to analyzed and interpret the data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

No data was used for the research described in the article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] M.A. Isa, R.S. Majumdar, S. Haider, S. Kandasamy, Molecular modelling and dynamic simulation of UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA)

from Mycobacterium tuberculosis using in silico approach, Informat. Med. Unlocked 12 (2018) 56–66.
[2] D. Jothieswari, K. Bhaskar Reddy, Molecular Docking studies of potential chemical inhibitors on multi-drug resistance genes in Mycobacterium tuberculosis, Int. J. Innovat. Drug Discov. 5 (1) (2015) 40–45.
[3] World Health Organization, Digital Health for the End TB Strategy: an Agenda for Action (No. WHO/HTM/TB/2015.21), World Health Organization, 2015. http://www.who.int/tb/areas-of-work/digital-health/Digital_ health_EndTBstrategy.pdf. (Accessed 8 August 2016).
[4] A. Jindani, V.R. Aber, E.A. Edwards, D.A. Mitchison, The early bactericidal activity of drugs in patients with pulmonary tuberculosis, Am. Rev. Respir. Dis. 121 (6) (1980) 939–949.
[5] R. Hafner, J.A. Cohn, D.J. Wright, N.E. Dunlap, M.J. Egorin, M.E. Enama, DATRI 008 Study Group, The early bactericidal activity of isoniazid in pulmonary tuberculosis: optimization of methodology, Am. J. Respir. Crit. Care Med. 156 (3) (1997) 918–923.
[6] A. Telenti, W.J. Philipp, S. Sreevatsan, C. Bernasconi, K.E. Stockbauer, B. Wieles, W.R. Jacobs Jr., The emb operon, a gene cluster of MTB involved in resistance to ethambutol, Nat. Med. 3 (5) (1997) 567.
[7] M.H. Hazbón, M. Brimacombe, M.B. del Valle, M. Cavatore, M.I. Guerrero, M. Varma-Basil, C.I. León, Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant MTB, Antimicrob. Agents Chemother. 50 (8) (2006) 2640–2649.
[8] L. Fenner, M. Egger, T. Bodmer, E. Altpeter, M. Zwahlen, K. Jaton, H.H. Siegrist, Effect of mutation and genetic background on drug resistance in MTB, Antimicrob. Agents Chemother. 56 (6) (2012) 3047–3053.
[9] D.A. Rozwarski, G.A. Grant, D.H. Barton, W.R. Jacobs, J.C. Sacchettini, Modification of the NADH of the isoniazid target (InhA) from MTB, Science 279 (5347) (1998) 98–102.
[10] A.S. Lee, A.S. Teo, S.Y. Wong, Novel mutations in ndh in isoniazid-resistant MTB isolates, Antimicrob. Agents Chemother. 45 (7) (2001) 2157–2159.
[11] R. Johnson, E.M. Streicher, G.E. Louw, R.M. Warren, P.D. Van Helden, T.C. Victor, Drug resistance in MTB, Curr. Issues Mol. Biol. 8 (2) (2006) 97–112.
[12] R. Cloete, E. Oppon, E. Murungi, W.D. Schubert, A. Christoffels, Resistance-related metabolic pathways for drug target identification in MTB, BMC Bioinf. 17 (1) (2016) 75.
[13] D.J. Murphy, J.R. Brown, Identification of gene targets against dormant phase MTB infections, BMC Infect. Dis. 7 (1) (2007) 84.
[14] Z. Ahmad, T. F Laughlin, Medicinal chemistry of ATP synthase: a potential drug target of dietary polyphenols and amphibian antimicrobial peptides, Curr. Med. Chem. 17 (25) (2010) 2822–2836.
[15] S. Duarte, S. Gregoire, A.P. Singh, N. Vorsa, K. Schaich, W.H. Bowen, H. Koo, Inhibitory effects of cranberry polyphenols on formation and acidogenicity of Streptococcus mutans biofilms, FEMS Microbiol. Lett. 257 (1) (2006) 50–56.
[16] K. Andries, P. Verhasselt, J. Guillemont, H.W. Göhlmann, J.M. Neefs, H. Winkler, V. Jarlier, A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis, Science 307 (5707) (2005) 223–227.
[17] S. Hong, P.L. Pedersen, ATP synthase and the actions of inhibitors utilized to study its roles in human health, disease, and other scientific areas, Microbiol. Mol. Biol. Rev. 72 (4) (2008) 590–641.
[18] J. Zheng, V.D. Ramirez, Inhibition of mitochondrial proton F0F1-ATPase/ATP synthase by polyphenolic phytochemicals, Br. J. Pharmacol. 130 (5) (2000) 1115–1123.
[19] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (17) (1997) 3389–3402.
[20] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.F. Meyer Jr., M.D. Brice, J.R. Rodgers, M. Tasumi, The Protein Data Bank: a computer-based archival file for macromolecular structures, Eur. J. Biochem. 80 (2) (1977) 319–324.
[21] D.J. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (22) (1994) 4673–4680.
[22] U. Pieper, B.M. Webb, D.T. Barkan, D. Schneidman-Duhovny, A. Schlessinger, H. Braberg, R.S. Datta, ModBase, a database of annotated comparative protein structure models, and associated resources, Nucleic Acids Res. 39 (suppl_1) (2010) D465–D474.
[23] D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, et al., AMBER 2015, University of California, San Francisco, 2015.
[24] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (9) (1993) 1511–1519.
[25] R. Lüthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (6364) (1992) 83.
[26] G. Mukherjee, B. Jayaram, Rapid identification of hit molecules for target proteins via physicochemical descriptors, Phys. Chem. Chem. Phys. 15 (23) (2013) 9107–9116.
[27] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (14) (1998) 1639–1662.
[28] Johann Gasteiger, Mario Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, Tetrahedron 36 (22) (1980) 3219–3228.
[29] W.L. DeLano, The PyMOL User's Manual, DeLano Scientific, San Carlos, CA, 2002, p. 452.
[30] R.A. Laskowski, N.M. Luscombe, M.B. Swindells, J.M. Thornton, Protein clefts in molecular recognition and function, Protein Sci. – Publ. Protein Soc. 5 (12) (1996) 2438.
[31] A.C. Wallace, R.A. Laskowski, J.M. Thornton, Derivation of 3D coordinate templates for searching structural databases: application to Ser-His-Asp catalytic triads in the serine proteinases and lipases, Protein Sci. 5 (6) (1996) 1001–1013.
[32] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 64 (2012) 4–17.
[33] D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, J. Med. Chem. 45 (12) (2002) 2615–2623.
[34] F. Cheng, W. Li, Y. Zhou, J. Shen, Z. Wu, G. Liu, Y. Tang, admetSAR: a Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties, 2012.
[35] T. Sander, J. Freyss, M. von Korff, C. Rufener, DataWarrior: an open-source program for chemistry aware data visualization and analysis, J. Chem. Inf. Model. 55 (2) (2015) 460–473.
[36] J.Y. He, C. Li, G. Wu, Discovery of potential drugs for human-infecting H7N9 virus containing R294K mutation, Drug Des. Dev. Ther. 8 (2014) 2377.
[37] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expet Opin. Drug Discov. 10 (5) (2015) 449–461.
[38] M.A. Isa, Homology modeling and molecular dynamic simulation of UDP-Nacetylmuramoyl-L-alanine-D-glutamate ligase (MurD) from *Mycobacterium tuberculosis* H37Rv using *in silico* approach, Comput. Biol. Chem. 78 (2019) 116–126.