Chemistry Africa (2020) 3:391–408
https://doi.org/10.1007/s42250-020-00132-9

ORIGINAL ARTICLE

Quantitative Structure–Activity Relationship Model, Molecular Docking Simulation and Computational Design of Some Novel Compounds Against DNA Gyrase Receptor

Shola Elijah Adeniji[1] · David Ebuka Arthur[2] · Mustapha Abdullahi[1] · Abdurrashid Haruna[1]

Received: 14 December 2019 / Accepted: 14 March 2020 / Published online: 31 March 2020

Abstract
Time consumed and expenses in discovering and synthesizing new hypothetical drugs with improved biological activity have been a major challenge toward the treatment of multi-drug resistance strain *Mycobacterium tuberculosis* (TB). To solve the above problem, Quantitative structure activity relationship (QSAR) is a recent approach developed to discover a novel drug with a better biological against *M. Tuberculosis*. A validated QSAR model developed in this study to predict the biological activities of some anti-tubercular compounds and to design new hypothetical drugs is influenced with the molecular descriptors; MATS2s, nHBint3, maxtsC, TDB9u, RDF90i and RDF110s. Molecular docking studies was as well carried for all the studied compounds in order to show the interactions and binding modes between the ligand and the receptor (DNA gyrase). The lead compound (compound 41) with higher anti-tubercular activity was observed with prominent binding affinity of − 21.9 kcal/mol compared to the recommended drugs; Isoniazid (− 14.6 kcal/mol). Therefore, compound 41 served as a template structure to designed compounds with more efficient activities. Among the compounds designed; compounds 41p was observed with better anti-tubercular activities with more prominent binding affinities of − 24.3 kcal/mol. The findings in the research will be valued to pharmacology, medicinal chemists and pharmacist to design and synthesis a novel drug candidate against the tuberculosis. Moreover, in vitro and in vivo test could be carried out to validate the computational results.

Keywords Model · Triazole QSAR · Tuberculosis

1 Introduction

Multi-drug resistance strain *Mycobacterium tuberculosis* (TB) has pose a challenge toward the treatment of tuberculosis in the global community. World Health Organization in (2017), has reported 9.0 million people infected with tuberculosis, 360,000 HIV patient whom were leaving with tuberculosis, death of 230,000 children and death of 1.6 million people worldwide [1]. Some of the notable commercial sold drugs administered to people infected with tuberculosis are isoniazide (INH), pyrazinamide (PZA), rifampicin (RMP) and para-amino salicylic acid (PAS) [2]. The emergence of multi-drug resistance strain of *M. tuberculosis* toward the aforementioned drugs has led to advances in searching for new and better approach that is precise and fast in developing a novel compound with improved biological activity against *M. tuberculosis* [2, 3].

For the time being, QSAR is a theoretical approach with widely used computational method in predicting and designing new hypothetical drug candidate [2]. Multi-variant QSAR model is expressed mathematically to relates the biological activity of each compound with its respective molecular structures. Meanwhile, some prominent researchers [4–7] have successful established QSAR models to show the relationship between some anti-*M. tuberculosis* inhibitor's such as; chalcone, quinolone, 7-methyijuglone, pyrrole and their respective biological activities using QSAR approach. However, QSAR alongside with molecular docking simulation study have not been fully established to relate the structures and activities of the inhibitory compounds as well as the interaction mode with the receptor (DNA gyrase). Hence, this research was aimed to build a robust QSAR model with high predictability, carry out a molecular

✉ Shola Elijah Adeniji
shola4343@gmail.com

[1] Chemistry Department, Ahmadu Bello University, Zaria, Nigeria

[2] Chemistry Department, Baze University, Abuja, Nigeria

docking simulation and to design new potent hypothetical compounds with better anti-tubercular activities against *M-tuberculosis*.

2 Materials and Methods

2.1 Data Collection

Fifty (50) molecules comprising the derivatives of 1,2,4-Triazole reported as anti-*mycobacterium tuberculosis* that were used in this study were obtained from the literature [3]. The biological activities of these compounds and the list of the compounds were presented in Table 1.

2.2 Molecular Optimization

Spartan 14 software version 1.1.4 [https://down.cd/10055/buy-WaveFunction-Spartan-14-1.1.4-download] was used to optimize all the inhibitory compounds in order for the compounds to attain stable conformation at a minimal energy. The strain energy from the molecules were removed by employing Molecular Mechanics Force Field (MMFF) and complete optimization was achieved with the aid of Density Functional Theory (DFT) by utilizing the (B3LYP/6-31G*) basic set [5].

2.3 Generation of Molecular Descriptor

A descriptor is a mathematical logic that defines the properties of a molecule in a numeral term based on the connection between the biological activity of each molecule and its molecular structure. Descriptors for all the inhibitory molecules was calculated with the aid of PaDEL descriptor software version 2.20 [http://www.yapcwsoft.com/dd/padeldescriptor/] and a total of 1879 molecular descriptors were generated.

2.4 Normalization and Pretreatment of Data

For each of the variable (descriptor) to have the same chance at the inception so as to influence the QSAR model, the descriptors values generated from PaDEL descriptor software version 2.20 were subjected to normalization using Eq. 1 [2, 8].

$$D = \frac{d_1 - d_{\min}}{d_{\max} - d_{\min}} \tag{1}$$

where d_{max} and d_{min} are the maximum and minimum value for each descriptors column of D. d_1 is the descriptor value for each of the molecule. Immediately after the data have been normalized, the normalized data were then subjected to pretreatment [http://teqip.jdvu.ac.in/QSAR_Tools/]. so as to remove redundant descriptors.

2.5 Generation Training and Test Set

The whole compounds that made up the data set was divided into training and test set in proportion of 70 to 30% using Kennard and Stone's algorithm which was incorporated in DTC lab software [http://teqip.jdvu.ac.in/QSAR_Tools/]. The development of the QSAR model and internal validation test were performed on the training set while the confirmation of the developed model was performed on test set.

2.6 Building of QSAR Models and Internal Validation Test

The QSAR models were built by adopting the Genetic Function Approximation (GFA) technique incorporated in the Material Studio software version 8.0 [https://www.3dsbiovia.com/products/collaborative-science/biovia-materials-studio/] to select the optimum descriptors for the training set. Meanwhile, Multi-linear regression Approach (MLR) was used as a modelling tool to develop the multi-variant equations by placing the activity data in the last column of Microsoft Excel 2013 spread sheet which was later imported into the Material Studio software version 8.0 to generate the QSAR model. The internal validation test to affirm the built model is robust and also have a high predictability was also performed in Material Studio software version 8.0 and reported.

2.7 Evaluation of Leverage Values (Applicability Domain)

Influential and outlier molecule present in the both the training and test set were determined by employing the applicability domain approach. The leverage h_i approach as defined in Eq. 2 was used define applicability domain space ± 3 for outlier molecule [9, 10].

$$h_i = M_i\left(M^T M\right)^{-1} M_i^T \tag{2}$$

where M_i represent the matrix of i for the training set. M represent the $n \times d$ descriptor matrix for the training set and M^T is the transpose of the training set (M). M_i^T represent the transpose matrix M_i. Meanwhile, the warning leverage h^* defined in Eq. 3 is the limit boundary to check for an influential molecule.

$$h^* = 3\frac{(d+1)}{N} \tag{3}$$

Table 1 Molecular structures of inhibitory compounds and their derivatives as anti-tubercular agents

S/N	Molecules	Experimental activity (pBA)	Predicted activity (pBA)	Residual	Leverage
1	1-allyl-3-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.9432	7.9218	0.0214	0.46159
2[a]	1-allyl-5-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.4535	7.6812	− 0.2277	0.1122
3	1-allyl-3-methyl-5-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.9759	7.8435	0.1324	0.0475
4	1-allyl-3-ethyl-5-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.9759	7.1969	− 0.221	0.3967
5[a]	3,5-dibromo-1-(prop-2-yn-1-yl)-1*H*-1,2,4-triazole	7.9294	7.9274	0.002	0.3679
6[a]	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-5-(tert-butyl)-1*H*-1,2,4-triazol-3-yl)thio)methyl)-1H-1,2,3-triazole	5.4543	5.6835	− 0.2292	0.3325
7	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-5-(4-methoxyphenyl)-1*H*-1,2,4-triazol-3-yl)thio)methyl)-1*H*-1,2,3-triazole	4.7441	4.7913	− 0.0472	0.0739
8	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-5-(4-chlorophenyl)-1*H*-1,2,4-triazol-3-yl)thio)methyl)-1*H*-1,2,3-triazole	6.1674	6.2865	− 0.1191	0.0855
9[a]	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	6.3456	6.4919	− 0.1463	0.0565
10	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-3-methyl-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	7.4134	7.1414	0.272	0.0903
11	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-3-(tert-butyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	5.7441	6.0728	− 0.3287	0.0799
12	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-3-(4-nitrophenyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	5.9258	5.6249	0.3009	0.0689
13[a]	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-3-(4-methoxyphenyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	5.6754	5.47	0.2054	0.0645
14	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-3-(4-chlorophenyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	6.3793	6.3309	0.0484	0.1437
15	3-(allylthio)-1*H*-1,2,4-triazole	6.1667	6.5298	− 0.3631	0.0313
16[a]	5-(allylthio)-3-methyl-1*H*-1,2,4-triazole	5.8765	6.8631	− 0.9866	0.0488
17	5-(allylthio)-3-(tert-butyl)-1*H*-1,2,4-triazole	6.4171	6.122	0.2951	0.3531
18	5-(allylthio)-3-(4-nitrophenyl)-1*H*-1,2,4-triazole	5.9413	6.01177	− 0.07047	0.2506
19	3-(allylthio)-5-(4-methoxyphenyl)-1*H*-1,2,4-triazole	7.6397	7.5921	0.0476	0.4036
20	3-(allylthio)-5-(4-chlorophenyl)-1*H*-1,2,4-triazole	7.6899	7.8302	− 0.1403	0.2527
21	1-allyl-3-(allylthio)-1*H*-1,2,4-triazole	6.3981	6.196	0.2021	0.2443
22	1-allyl-3-(allylthio)-5-methyl-1*H*-1,2,4-triazole	5.8131	6.4174	− 0.6043	0.399
23	1-allyl-3,5-dimethyl-1*H*-1,2,4-triazole	6.1213	5.9031	0.2182	0.2461
24	3,5-dibromo-1-(2-methylallyl)-1*H*-1,2,4-triazole	5.4406	5.2561	0.1845	0.0449
25	1-benzyl-4-((3,5-dibromo-1*H*-1,2,4-triazol-1-yl)methyl)-1*H*-1,2,3-triazole	4.9074	4.8361	0.0713	0.43062
26[a]	1-allyl-3-(allylthio)-5-(4-chlorophenyl)-1*H*-1,2,4-triazole	7.0123	6.8337	0.1786	0.2386
27	1-allyl-5-(allylthio)-1*H*-1,2,4-triazole	6.5267	6.5716	− 0.0449	0.2085
28	1-allyl-5-(allylthio)-3-methyl-1*H*-1,2,4-triazole	5.7405	6.0828	− 0.3423	0.0858
29[a]	1-allyl-5-(allylthio)-3-(tert-butyl)-1*H*-1,2,4-triazole	5.6533	5.4662	0.8171	0.0623
30[a]	1-allyl-5-(allylthio)-3-(4-nitrophenyl)-1*H*-1,2,4-triazole	6.1923	6.0292	0.1631	0.1091
31	1-allyl-5-(allylthio)-3-(4-methoxyphenyl)-1*H*-1,2,4-triazole	7.3233	7.4961	− 0.1728	0.176
32	1-allyl-5-(allylthio)-3-(4-chlorophenyl)-1*H*-1,2,4-triazole	6.0097	6.3017	− 0.292	0.0616
33	3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	6.0928	5.9367	0.1561	0.1445
34	1-allyl-5-(tert-butyl)-3-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.8656	7.7939	0.0717	0.2079
35	5-(tert-butyl)-3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	6.8568	6.8627	− 0.0059	0.2036
36[a]	5-(4-nitrophenyl)-3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	6.2234	7.6353	− 2.4119	0.1921
37	5-(4-methoxyphenyl)-3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	7.3079	7.2271	0.0808	0.3495
38	5-(4-chlorophenyl)-3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	7.314	7.4916	− 0.1776	0.0444
39[a]	1-allyl-3-(tert-butyl)-5-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.7412	5.6835	2.0577	0.0742
40	1-allyl-5-(tert-butyl)-3-(prop-2-yn-1-ylthio)-1*H*-1,2,4-triazole	7.6615	7.0875	0.574	0.079
41	5-methyl-3-(prop-2-yn-1-ylsulfonyl)-1*H*-1,2,4-triazole	8.0214	8.0856	− 0.0642	0.3174

Table 1 (continued)

S/N	Molecules	Experimental activity (pBA)	Predicted activity (pBA)	Residual	Leverage
42[a]	3-(allylthio)-5-(tert-butyl)-1-(prop-2-yn-1-yl)-1*H*-1,2,4-triazole	6.8494	7.6612	1.1882	0.4217
43	1-benzyl-4-(((3-methyl-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	4.925	4.8788	0.0462	0.0513
44[a]	4-(((1*H*-1,2,4-triazol-5-yl)thio)methyl)-1-benzyl-1*H*-1,2,3-triazole	5.0345	4.8582	0.1763	0.2688
45	1-benzyl-4-(((3-(tert-butyl)-1*H*-1,2,4-triazol-5-yl)thio)methyl)-1*H*-1,2,3-triazole	5.0064	5.0807	− 0.0743	0.3304
46	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-1*H*-1,2,4-triazol-3-yl)thio) methyl)-1*H*-1,2,3-triazole	5.7386	5.8174	− 0.0788	0.5099
47	1-benzyl-4-(((1-((1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl)-5-methyl-1*H*-1,2,4-triazol-3-yl)thio)methyl)-1*H*-1,2,3-triazole	5.5994	5.5669	0.0325	0.8062
48[a]	1-allyl-3-(allylthio)-5-(tert-butyl)-1*H*-1,2,4-triazole	6.2878	6.29117	− 0.00337	0.2564
49	1-allyl-3-(allylthio)-5-(4-nitrophenyl)-1*H*-1,2,4-triazole	5.7268	5.9799	− 0.2531	0.3462
50[a]	1-allyl-3-(allylthio)-5-(4-methoxyphenyl)-1*H*-1,2,4-triazole	7.366	7.531	− 0.165	0.1898

[a]Test set

where *d* is the total number of descriptors present in the built model and *N* is the total number of compounds that made up the training set.

2.8 Y-Randomization Validation Test

Y-Randomization test [http://teqip.jdvu.ac.in/QSAR_Tools/] is one of the external validation criteria which has to be considered in order to ascertain that the developed model is not built by chance [10, 11]. Random shuffling of the data was performed on the training set following the principle laid by [10, 12]. The activity data (dependent variable) were shuffled while the descriptors (independent variables) were kept unchanged in order to generate the Multi-linear regression (MLR) model. For the developed QSAR to pass the Y-Randomization test, the R^2 and Q^2 values for the model must be significantly low for numbers of trials while Y-randomization Coefficient (cR_p^2) shown in Eq. 4 must be ≥ 0.5 in order to establish the robustness of the model.

$$cR_p^2 = R \times \left[R^2 - \left(R_r\right)^2\right]^2 \quad (4)$$

where cR_p^2 is Y-randomization Coefficient, R is correlation coefficient and Rr is average 'R' of random models.

2.9 Affirmation of the Build Model

The internal and external validation criteria for both test and training set reported were compared with the generally accepted threshold value shown in Table 6 for any QSAR model [2, 10–13] in order to affirm the reliability, fitting, stability, robustness and predictability of the developed models.

2.10 Docking Studies

2.10.1 Preparation of receptor

The crystal structure of DNA gyrase shown in Fig. 1 was obtained from protein data bank with PDB code 31FZ [15]. Crystal structure of DNA gyrase was prepared by removing all bound substances (ligands and cofactors) and solvent molecules associated with the receptor. DNA gyrase preparation was done by launching the Discovery Studio Visualizer software; The prepared receptor was then saved in PDB file format which is the recommended input format in Pyrx and Discovery Studio Visualizer software. The

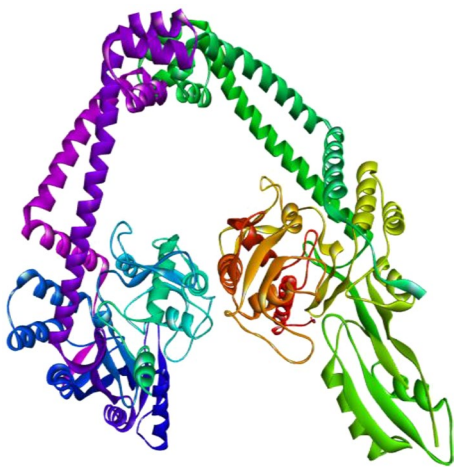


Fig. 1 Crystal structure of DNA gyrase

prepared receptor was transported into the Pyrx software in order to make it a macro molecule. [13].

2.10.2 Receptor (*DNA Gyrase*) Preparation

The crystal form of the target protein (DNA gyrase) was downloaded from protein data bank with PDB code 31FZ [14, 15]. All imported foreign substances such as solvent molecules, cofactors and ligands allied with the enzyme were disinterested using Discovery Studio Visualizer software [https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/]. Later on, the target protein was saved format (PDB) which is the recommend format for Pyrx software and Discovery Studio Visualizer. Thereafter, the target protein saved in PDB format was imported in the Pyrx software and converted as macro molecules [5, 16].

2.10.3 Ligand Preparation

The stable conformation of triazole derivatives at a minima energy were achieved with the aid of Spartan 14 software at Density Functional Theory (DFT) level which serve as an optimized tool. The optimized ligands were then saved as a PDB format which is the recommend format for the Pyrx software. Later on, the ligands saved in PDB format were imported in the Pyrx software and converted as micro molecules [5, 16].

2.10.4 Docking of Receptor and Ligand

Ligand-receptor interactions between triazole derivatives and the receptor (DNA gyrase) was carried out using molecular docking technique by employing the PyRx virtual screening software. The PyRx software [https://pyrx.sourceforge.io/], is an open source software for performing virtual screening. PyRx uses AutoDock Vina [http://vina.scripps.edu/] and AutoDock 4.2 [http://autodock.scripps.edu/] as docking softwares. Discovery Studio Visualizer software version 2016 [https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/] was used to visualized and analyzed the docked results. [5, 16].

3 Results and Discussion

3.1 QSAR Studies

Optimum QSAR model for predicting the derivatives of 1, 2, 4 Triazole against *M. tuberculosis* was successfully achieved by adopting the combination of computational and theoretical method. Data set comprises of 50 compounds was partitioned into 35 training set and 15 test set using Kennard and Stone algorithm method. The 35 training set compounds were used to derive QSAR model using Multi-linear regression technique which also served as data set for internal validation test while the external validation test for the derived model was conducted on the test set.

Model 1

$$\begin{aligned}\mathrm{pBA} = &-3.927401745 * \mathrm{MATS2s} + 4.730973152 \\ &* \mathrm{nHBint3} + 1.1035920582 * \mathrm{maxtsC} \\ &+ 0.310934301 * \mathrm{TDB9u} - 0.791306892 \\ &* \mathrm{RDF90i} - 4.281096493 * \mathrm{RDF110s} \\ &+ 8.840916286\end{aligned}$$

Model 2

$$\begin{aligned}\mathrm{pBA} = &-2.418520845 * \mathrm{MATS2s} + 1.783195320 \\ &* \mathrm{nHBint3} + 1.310849563 * \mathrm{maxtsC} \\ &+ 0.0280218642 * \mathrm{TDB9u} - 4.992450732 \\ &* \mathrm{RDF150p} - 4.59209513 * \mathrm{Ds} + 9.702350851\end{aligned}$$

Model 3

$$\begin{aligned}\mathrm{pBA} = &-6.934102832 * \mathrm{MATS2s} + 1.760023432 \\ &* \mathrm{nHBint3} + 4.803387356 * \mathrm{maxtsC} \\ &+ 2.7934152560 * \mathrm{TDB9u} + 0.950439041 \\ &* \mathrm{RDF90i} - 3.521095439 * \mathrm{De} \\ &+ 7.4873028922\end{aligned}$$

The experimental activities reported in literature, the predicted activities calculated for all the anti-tubercular compounds, the leverage values and the residual values were presented in Table 1. The difference between the experimental activities and predicted activities is the residual values which were observed to be significant low. The low residual value indicates that the model built has a good predictive ability.

The optimum (2D and 3D) descriptors that efficiently describe the anti-tubercular compounds in relation to their biological activities were selected by GFA approach. The characterization and relative information on the molecular structure of the anti-tubercular agent illustrated by the descriptors were reported in numerical value as shown in Table 2. Meanwhile, for the purpose of reproducibility all the calculated descriptors for the both the training and test set in model 1 were presented in Table 3.

Various statistical analysis were conducted on the calculated descriptors in order to check the validity of the built model as reported in Table 4. Variance inflation factor (VIF) was evaluated for all the descriptors in order to determine the degree of correlation between each the descriptor. Generally, VIF value equal to 1 or falls with 1 and 5 signify

Table 2 Name of selected descriptors used in the QSAR model 1

S/NO	Descriptors symbols	Name of descriptor	Class
1	MATS2s	Moran autocorrelation–lag 2/weighted by I-state	2D
2	nHBint3	Count of E-State descriptors of strength for potential Hydrogen Bonds of path length 3	2D
3	maxtsC	Maximum atom-type E-State:#C⁻	2D
4	TDB9u	3D topological distance based autocorrelation - lag 9/unweighted	3D
5	RDF90i	Radial distribution function - 090/weighted by relative first ionization potential	3D
6	RDF110s	Radial distribution function - 110/weighted by relative I-state	3D

non-existence of inter-correlation among the descriptors. However, if the VIF value is greater than 10, it signify that the model developed is unstable hence, the model should be re-checked if necessary. Regarding the VIF values for each the descriptors which were found to be less than 5 as reported in Table 4 affirm that the descriptors were significantly orthogonal to each order since there is no inter-correlation between them.

The degree of contribution that each descriptor plays in the built model was evaluated by determining the standard regression coefficient (b_j^s) and mean effect (ME). The magnitude and signs for b_j^s and ME values reported in Table 4 indicate strength and direction with which each descriptor influence the activity model. The relationship between the descriptors and biological activity of each compound was determined by one way Analysis of variance (ANOVA). The probability value of each of the descriptor at 95% confidence level were found to be ($p < 0.05$) as presented in Table 4. Therefore this signify that the alternative hypothesis that says there is a direct relationship between the biological activity of each compound and the descriptor swaying the built model is accepted thus; null hypothesis proposing no direct relationship between biological activity of each compound and the descriptor swaying the built model is rejected. To further justify the validation of the descriptors in the activity model, Pearson correlation statistic was conducted to also check whether there is inter-correlation between each descriptors. The correlation coefficient between each descriptors reported in Table 5 were all $< \pm 0.8$. Hence this implies that all the descriptors were void of multicollinearity.

Validation results for both the external and internal assessment to assure that the built models are reliable and robust were presented in Table 6. These results were all in full agreement with general validation criteria resented in Table 6 to truly indorse that the stability and robustness of the model is valid. Reference to these validation results obtained, model one was selected and established to be the prime model which was used to predict the biological activities of 1, 2, 4 Triazole against *M. tuberculosis.*

The QSAR model generated in this research was compared with the models obtained in the literature [10] as shown below.

$$\begin{aligned} pBA = {} & -6.515153698 * AATS5e + 0.056593117 * VR1\ Dzs \\ & - 6.230058484 * SpMin7\ Bhe + 0.016884210 * TDB7e \\ & + 0.09232054RDF90i + 43.764308643 \end{aligned}$$

$R^2 = 0.9265$, Radj = 0.9045, $Qcv^2 = 0.8324$ and the external validation for the test set was found to be $R^2pred = 0.8034$ [10].

The validation factors reported in this work and those reported in the literature were all in agreement with the validation parameters presented in Table 6 which really inveterate that the model generated is predictive and robust.

The coefficient of Y- Randomization (c R_p^2) with significant value of 0.7849 greater than threshold value of 0.5 reported in Table 7 provide a reasonable supports that the model built is robust and not just by chance.

The graphical representation to show the degree of correlation between the predicted activities and experimental activities of the training and test set were shown in Fig. 2 and 3. The correlation coefficient (R^2) value of 0.9579 and 0.8657 for both the training set and test set shows that there is a high correlation existing between the predicted activities and experimental activities of the training and test set which were also in agreement with the accepted QSAR threshold values reported in Table 6.

The residual plot shown in Fig. 4 signify that there is no indication of computational incompetency and inaccuracy in the QSAR model derived as all the standard residual values for both training and test set were found within the defined boundary of ± 2 on the standard residual activity axis [2, 10, 17, 18].

The Williams plot to show the Applicability Domain space (AD) is shown in Fig. 5. It is observed that only compound 42 was found to exceed the warning leverage of ($h^* = 0.60$). Therefore it can be infer that this compound is an influential molecule. Moreover, it is also observed that all the compounds fall within the defined space of ± 3 which indicates that no compound is said to be outlier.

Table 3 Predicted descriptors for training set in generating model 1

Compound ID	Descriptors						Predicted activity
	AATS7s	NHBINT3	MinHCsatu	TDB9u	RDF90i	RDF110s	
Training set							
1	2.7856	0	0	66.7723	9.7018	1.4484	7.9218
3	2.5097	0	0	64.2822	14.7031	4.1443	7.8435
4	2.5521	0	0	61.7102	25.6093	5.8516	7.1969
7	2.7088	0	0	64.0155	29.9552	18.9220	4.7913
8	2.5844	0	0	63.5719	39.8984	13.5634	6.2865
10	2.7050	0	0	60.2370	38.3752	3.4893	7.1414
11	2.4646	0	0	56.9890	27.4708	4.5612	6.0728
12	3.3007	0	0	64.9270	37.3878	15.7905	5.6249
14	2.6852	0	0.0142	63.8144	35.4093	9.1629	6.3309
15	2.1371	1	0.4190	1.1330	1.1331	0.1201	6.5298
17	2.0826	0	0.4133	58.5298	3.5705	0.3163	6.1220
18	4.7247	0	0.4671	67.1607	7.6531	0.4786	6.0118
19	2.8469	1	0.4570	69.1436	5.2474	3.2908	7.5921
20	2.2622	1	0.4468	70.8652	5.4773	2.8170	7.8302
21	1.9786	0	0.4280	64.9950	2.1000	0.2178	6.1960
22	1.7518	0	0.4280	63.8383	5.1976	1.2078	6.4174
23	2.0300	0	0.4280	59.1007	4.2923	0.1201	0.2182
24	3.2326	0	0.4817	64.1520	7.9956	5.4128	0.1845
25	2.4553	0	0.4660	60.5533	19.7357	1.8402	0.0713
27	2.2621	0	0.4254	54.1164	2.5325	0.1201	6.5716
28	2.1489	0	0.4254	57.2739	2.8150	0.1201	6.0828
31	2.5525	0	0.4634	63.6313	7.8665	1.6895	7.4961
32	2.3422	0	0.4532	62.2568	4.2688	0.5199	6.3017
33	3.1371	0	0.7315	1.1330	1.1321	0.1201	5.9367
41	2.3417	1	0.4374	58.2436	5.0760	0.3445	7.7939
35	4.5038	0	0.7315	71.5465	3.4249	0.1201	6.8627
37	4.9471	0	0.7695	74.0294	9.1915	2.0424	7.2271
38	4.8209	0	0.7593	78.9183	9.4349	0.5006	7.4916
40	2.1894	1	0.5183	61.3878	3.4046	0.1220	7.0875
41	2.2942	0	0.7315	70.1844	1.9325	0.1201	8.0856
43	2.3236	1	0.5183	65.6485	1.6593	0.4707	4.8788
45	2.5329	1	0.5157	58.8727	2.4240	0.1201	7.8435
46	1.9641	1	0.5183	65.0038	3.7744	0.6002	5.0807
47	1.0769	1	0.6886	0	0	0	5.8174
49	1.0769	0	0.4013	0	0	0	5.5669
Test set							
2	2.5473	0	0.0142	67.3958	10.1285	2.7705	7.6812
5	2.4150	0	0.0142	62.4139	26.6485	9.5468	7.9274
6	2.5570	0	0.0142	64.7532	33.3413	13.9776	5.6835
9	2.7743	0	0.0142	57.7635	30.3393	2.2780	6.4919
13	2.7782	0	0.0142	64.2698	32.3992	14.4761	5.4700
16	1.8552	0	0.4133	59.9928	1.7205	0.1201	6.8631
26	2.1783	0	0.4558	59.1556	10.3734	0.7431	6.8337
29	2.0432	0	0.4254	57.4796	7.0129	0.3764	5.4662
30	3.5854	0	0.4791	64.6347	6.5749	2.7407	6.0292
36	6.7419	0	0.7852	79.9580	7.7638	6.3706	8.6353
39	2.2351	1	0.5157	59.3441	10.1783	0.1238	5.6835
42	2.5117	1	0.4370	61.0209	3.6973	0.3489	7.6612

Table 3 (continued)

Compound ID	Descriptors						Predicted activity
	AATS7s	NHBINT3	MinHCsatu	TDB9u	RDF90i	RDF110s	
44	2.9239	1	0.5157	56.3073	2.3942	0	4.8582
48	1.5835	0	0.6031	0	0	0	6.2878
50	2.5807	0	0	65.0644	8.7086	1.3482	7.366

Table 4 Statistical parameters that influence the model 1

Descriptors	Standard regression coefficient (b_j)	Mean effect (ME)	*P* Value (confidence interval)	VIF	Standard error
MATS2s	− 0.0427	− 0.2647	5.29E−5	2.3793	0.02107
nHBint3	1.2646	0.4817	6.52E−9	1.7641	0.03454
maxtsC	1.5752	0.9111	0.000148	2.1484	0.04119
TDB9u	0.9262	0.6805	0.000428	1.8355	0.01352
RDF90i	1.6853	0.6695	3.62E−6	3.583	0.03418
RDF110s	− 0.0883	− 0.4643	0.000295	2.5137	0.02515

Table 5 Pearson's correlation coefficient for the descriptor used in the QSAR model

Descriptors	AATS7s	NHBint3	MinHCsatu	TDB9u	RDF90i	RDF110s
AATS7s	1					
nHBint3	− 0.29825	1				
minHCsatu	0.191253	0.271336	1			
TDB9u	0.448447	− 0.19142	− 0.15015	1		
RDF90i	0.16936	− 0.37299	− 0.77991	0.270029	1	
RDF110s	0.119377	− 0.25277	− 0.67104	0.212725	0.081411	1

3.2 Molecular Docking Studies

3.2.1 Assessment of Binding Affinity

Binding affinity between the ligand and target enzyme is elucidated via docking studies. The outcomes of the docking studies evidently showed that the activity value of each docked ligand correlated with binding affinity which ranged from − 4.0 to − 21.9 kcal/mol presented in Table 8. Meanwhile, ligands 41 was observed with higher binding affinity of − 21.9 kcal/mol compared to the binding affinity of commended drugs; isoniazid (− 14.6 kcal/mol) and other derivatives. Hence, this gives an indicating that ligands 41 could serve as a better compound against tuberculosis.

3.2.2 Bond Type and Bond Length in the Ligand-Receptor Complex of Compound 41

The prominent ligands (compound 41) with highest binding affinity was viewed examined using Discovery Studio Visualizer software. The interaction of ligand 41 with target enzyme ''DNA gyrase'' is presented in Fig. 6. The interaction was observed with five hydrogen bonds (2.6234, 2.1123, 2.1922, 2.6012 and 2.6302Å) with GLN385, THR77, GLN385, ALA167 and ALA167 of the enzyme. The ligand the S=O of acts as H-bond acceptor with formation of two hydrogen bonds with GLN385 and THR77 of the target. More also, the ligand N–H group acts as H-bond donor with formation of three hydrogen bonds with GLN385, ALA167, and ALA167 of the enzyme. Meanwhile, the hydrophobic interactions were detected with VAL78 and PHE168 of the enzyme. The region of the H-bond and hydrophobic interaction of the ligand-receptor complex formed are presented shown in Figs. 7, 8. Therefore, the hydrophobic interactions and the H-bonds formation offer a significant evidence to proof that ligand 41 among its co-ligand has the highest efficiency against DNA gyrase receptor.

3.2.3 Bond Type and Bond Length in the Ligand-Receptor Complex of Isoniazid

The binding interaction in 2-Dimension of the target enzyme with the commended drug ''isoniazid''

Table 6 Validation parameters for each model using multi-linear regression (MLR)

S/NO	Validation Parameters	Formula	Threshold	Model 1	Model 2	Model 3
Internal Validation						
1	Friedman Lack of fit (LOF)	$\frac{SEE}{\left(1-\frac{w+qxj}{N}\right)^2}$	Significantly low	0.4232	0.4636	0.4671
2	R-squared	$1-\left[\frac{\sum\left(Y_{obs-Y_{pred}}\right)^2}{\sum\left(Y_{obs-\bar{Y}_{training}}\right)^2}\right]$	$R^2 > 0.6$	0.9368	0.918	0.8849
3	Adjusted R-squared	$\frac{R^2-P(N-1)}{N-p+1}$	$R^2_{adj} > 0.6$	0.896	0.8903	0.8635
4	Cross validated R-squared (Q^2_{cv})	$1-\left[\frac{\sum\left(Y_{pred-Y_{obs}}\right)^2}{\sum\left(Y_{obs-\bar{Y}_{training}}\right)^2}\right]$	$Q^2 > 0.6$	0.855	0.7923	0.7544
5	Significant Regression			Yes	Yes	Yes
6	Critical SOR F-value (95%)	$\frac{\sum\left(Y_{pred-Y_{obs}}\right)^2}{p} / \frac{\sum\left(Y_{pred-Y_{obs}}\right)^2}{N-p-1}$.	$F_{(test)} > 2.09$	57.21	53.46	51.89
Model Randomization						
10	Average of the correlation coefficient for randomized data ($\bar{R}_r$)			0.3634	0.3911	0.4018
11	Average of determination coefficient for randomized data ($\bar{R}^2_r$)		$\bar{R}^2_r < 0.5$	0.1831	0.2056	0.2319
12	Average of leave one out cross-validated determination coefficient for randomized data ($\bar{Q}^2_r$)		$\bar{Q}^2_r < 0.5$	− 0.2190	− 0.7395	− 1.8420
13	Coefficient for Y-randomization (c R^2_p)	$R^2 \times \left(1-\sqrt{\left\lvert R^2-\bar{R}^2_r\right\rvert}\right)$	$^c R^2_p > 0.6$	0.6849	0.6312	0.6145
External validation						
14	Slope of the plot of observed activity against predicted activity values at zero intercept **(K)**	$\frac{\Delta Y_{Obs}}{\Delta Y_{pred}}$	$0.85 < k < 1.15$	0.9975	1.0046	1.0041
15	Slope of the plot of predicted against observed activity at zero intercept **(k′)**		$0.85 < k < 1.15$	0.9988	0.9251	0.8492
16	$/r^2_0 - r'^2_0/$	$\frac{\Delta Y_{pred}}{\Delta Y_{Obs}}$	<0.3	0.2231	0.2994	0.3308
17	$\frac{r^2-r^2_0}{r^2}$		<0.1	0.0207	0.0535	0.1222
18	$\frac{r^2-r'^2_0}{r^2}$		<0.1	0.036	0.0706	0.1093
19	$\mathbf{R^2_{test}}$	$R^2_{test} = 1-\frac{\sum(Ypred_{test}-Y_{obs_{test}})^2}{\sum(Ypred_{test}-\bar{Y}_{training})^2}$	> 0.6	0.7925	0.7204	0.6939

Key: SEE is the Standard error of estimation, *w* is the total number of terms present in the built model except the constant term, *j* is the number of descriptors confined in the built model, *q* is a user-defined factor and N is the number of compounds of training set. Y_{obs}, $\bar{Y}_{training}$ and Y_{pred} are the observed activity, mean observed activity of the training compounds and the predicted activity respectively. r^2 is correlation coefficients of the plot of observed activity against predicted activity values, r^2_o is the correlation coefficients of the plot of observed activity against predicted activity values at zero intercept, r'^2_o is correlation coefficients of the plot of predicted activity against observed activity at zero intercept [2, 10–13]

is represented in Fig. 6. The amino acid; SER279 and ALA337 and ALA337 are the main binding site through which the target enzyme bonded with Isoniazid via the hydrogen bond length; 2.52954, 2.29943 and 2.24657Å. Meanwhile, the amino acid; CYS345 and PHE338 are the main binding site through which the target enzyme bonded with Isoniazid via the hydrophobic interactions. Based on the observations, increase in number of hydrogen bonds in ligand 41 of triazole derivatives provide a concrete evidence to support the claim that ligand 41 binds efficiently with the binding pocket of the receptor when compared to the commended drug ‘‘isoniazid’’.

Table 7 Y-randomization parameters test for model 1

Model	R	R^2	Q^2
Original	**0.9559**	**0.8949**	**0.9657**
Random 1	0.5528	0.2353	− 0.16
Random 2	0.7	0.471	− 0.0322
Random 3	0.5227	0.2129	− 0.3133
Random 4	0.675	0.4449	0.0342
Random 5	0.3972	0.3389	− 0.4789
Random 6	0.3063	0.205	− 0.4926
Random 7	0.6787	0.5487	0.0126
Random 8	0.7135	0.4856	0.061
Random 9	0.5321	0.2197	− 0.133
Random 10	0.6709	0.4408	− 0.001
Random models parameters			
Average r	0.3634		
Average r^2	0.1831		
Average Q^2	− 0.2190		
cRp^2	0.6925		

3.3 Discussion on Designed Compounds

3.3.1 Computational Design of New Hypothetical Compound

Ligand based design approach was used to designed new hypothetical compounds with improved activities against tuberculosis. The best compounds among the derivative was used as template structure to design the new compounds. The modification was done by deletion, substitution and insertion of active substituent(s) into the structure template i.e. compound 41 as shown in Fig. 9. Compound 41 was selected as the template structure due to fact that the compound falls within the defined Applicability Domain (AD) as presented in Fig. 5. The deletion, substitution and insertion of the substituent(s) was successfully made around the triazole and acetylene of the template structure at position 12 and 8 as presented in Fig. 9. The molecular descriptors; maxtsC, nHBint3, TDB9u and RDF90i in the built QSAR models designated that the activity of the compound is positively influenced. Modifications of the template structure at position 12 and 8 with alkyl group, H atom, and methoxy group result to derivation of sixteen new novel compounds with improved activities against tuberculosis as described in Table 9. In order to screen and ascertain the Applicability Domain (AD) space of the designed, leverage value for each of the compound designed was calculated. Meanwhile, the leverage value reported for all the designed compounds as presented in Table 9 asserted that all the compound designed falls within the warning leverage $h^* = 0.60$. Therefore, this implied that each of the compound designed was within defined Model AD space. Based on the calculated activity for the compound in Table 9, it's obviously seen that compounds 41p was observed with against tuberculosis. The prominent anti-tubercular observed in designed compound 41p was due to modification of the template at position 12 with alkyl group (CH_3) which releases electron to ring system through positive inductive effect (+I) and modification at position 8 with 1H-triazole. The substituents with +I effect connected to the structure template rise the electron density which make the triazole pharmacophore of the compound 41p more basic. Hence, this gives reasonable explanation for its high activities toward *Mycobacterium tuberculosis.*

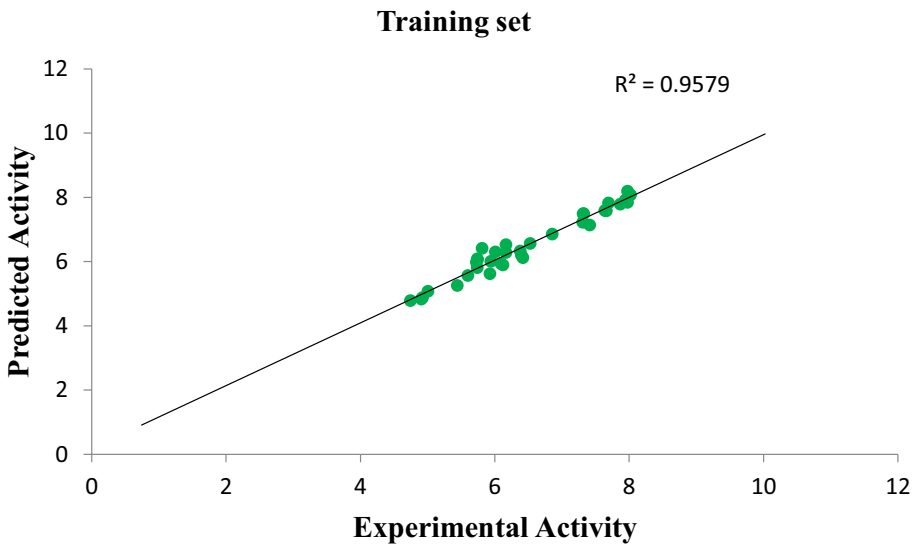


Fig. 2 Plot of predicted activity against observed activity of training set

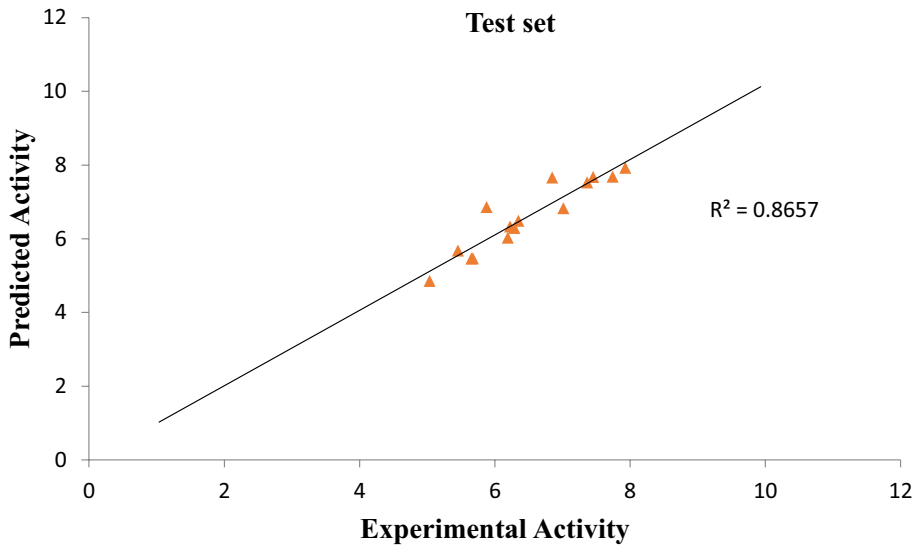


Fig. 3 Plot of Predicted activity against Observed activity of test set

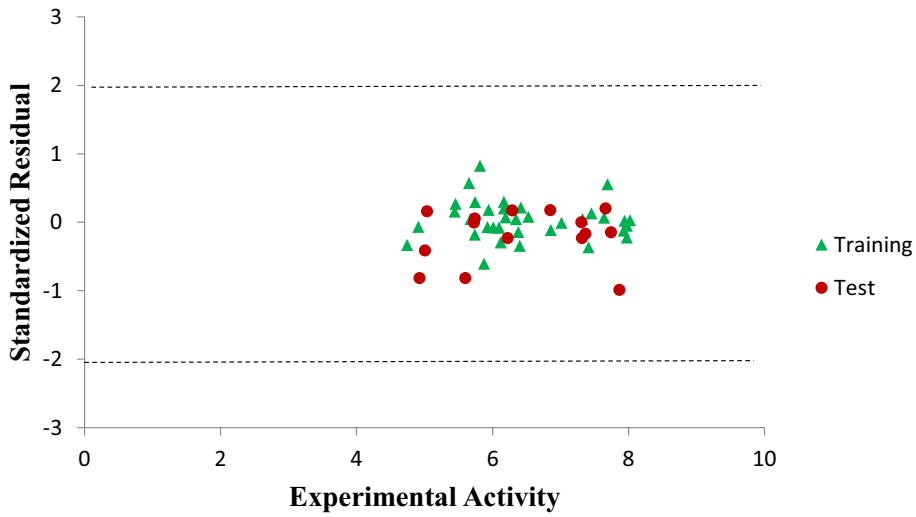


Fig. 4 Plot of standardized residual activity versus observed activity

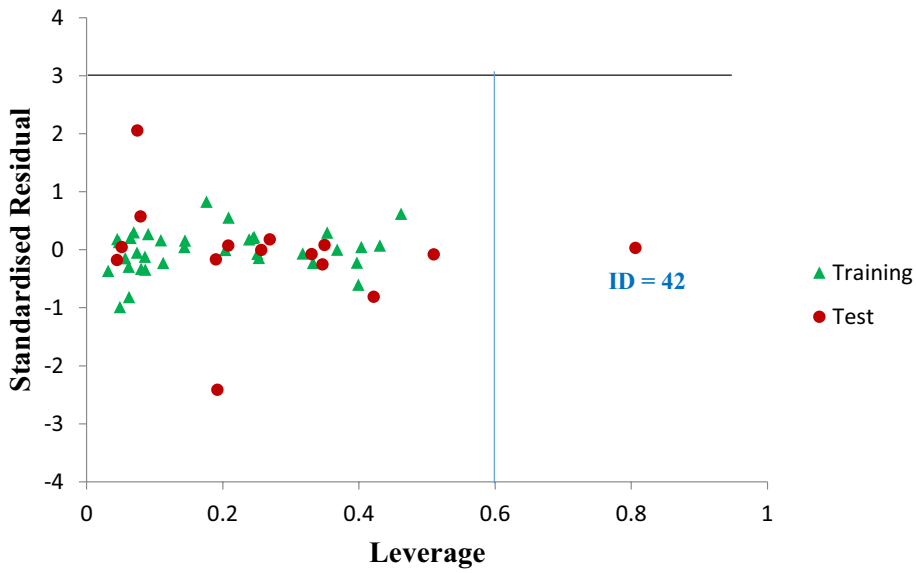


Fig. 5 The Williams plot of the standardized residuals versus the leverage value

Table 8 Molecular docking interactions between *M. tuberculosis* target (DNA gyrase) ligands (1,2,4-Triazole derivatives)

Ligand	Binding affinity (BA) Kcal/mol	Hydrogen bond interactions		Hydrophobic interaction
		Amino acid	Bond length (Å)	Amino acid
1	− 15.5	PRO95 HIS344 ALA232	3.0623 2.1455 1.2566	GLN 322, CYS345, GLN385, PHE 168, ALA176, TRP182, ARG72, VAL78
2	− 7.2	ALA169 GLN381	3.7564 1.3565	PHE280, ALA167, ALA233, THR77
3	− 14.6	GLN389 ALA169 VAL78	2.1734 2.3561 1.4464	ALA167, VAL228, ALA281, ARG386, LEU164, TRP182
4	− 14.8	LEU133 GLN341 ARG148	2.3562 2.3355 1.2443	PRO169, VAL228, VAL178, LEU164
5	− 13.7	ARG161 GLN383 ARG380	2.0061 2.3554 2.4672	PHE185, ALA167, VAL228, ASN74, CYS134
6	− 6.1	ALA113	1.2354	LEU176, VAL228, CYS143, PHE248
7	− 4.0	–	–	ALA167, TRP182, VAL78, CYS145, SER247
8	− 7.5	ALA147	2.4344	ALA212, TRP182, CYS221, PRO165
9	− 6.2	GLN315	1.3455	TRP182, ALA143, PHE168
10	− 12.3	THR78 GLN365	2.4566 2.4344	ALA236, VAL78, VAL228, LEU164
11	− 6.1	ASN78	1.3466	ALA233, PRO134, LA167, VAL78
12	− 6.5	GLN315	1.6457	VAL83, VAL83, TRP, 182LEU76
13	− 12.7	LEU113 TRP192	2.3431 3.0340	PRO346, ALA213, ALA165
14	− 7.8	ASN78	2.76681	LEU154,VAL84, ALA167
16	− 15.7	ASP112 PHE119 ALA116	2.3515 2.1544 2.6868	TYR213, PRO182
17	− 8.7	GLN345 CYS315	2.7344 2.4345	ALA233, VAL78, TRP182, VAL76
18	− 7.5	GLN395	2.54451	VAL83, PHE163, PRO285, ALA167,PRO215,
19	− 14.6	VAL74 ALA213 LEU76	2.1334 2.4888 2.4529	VAL78, SER217, PHE161, PRO285, THR238, ALA117
20	− 14.7	GLN355 ARG126 GLN125	2.5696 2.4581 2.0499	PHE241, PRO94, PRO54, VAL178, PRO119, PHE93, CYS345
21	− 8.6	ASN79 ASP212	3.0181 2.2843	VAL76, LEU207, LEU73, PRO215, VAL128
22	− 7.6	THR78	2.45441	PHE128,, PHE168, VAL78, TRP182, ALA137, TRP112
23	− 6.9	THR65	1.43632	CYS110, GLN185, ALA243
24	− 6.2	GLN325	2.1313	GLN315, ARG165, GLN385, VAL127, CYS234
25	− 4.1	–	–	PHE205, LEU113, GLY196, LEU207, ARG101
26	− 14.8	GLN115 ALA127 VAL92	2.2312 2.2332 2.5776	PRO102, ALA238, PHE148, TRP112, PHE230, ALA101, VAL82, LYS123, VAL78
27	− 9.2	ASP79 GLN315	3.3618 2.4892	CYS245, ALA267, PHE118, PRO346, ALA233 TRP112
28	− 7.5	VAL79	2.4331	PRO185, TRP182, ALA167, TRP132, PRO34, VAL27
29	− 7.7	ASN78	3.4509	LEU161, VAL98, VAL122
30	− 7.9	GLN325 LEU113	2.1721 2.2282	VAL68, ALA233, TRP48, PHE158

Table 8 (continued)

Ligand	Binding affinity (BA) Kcal/mol	Hydrogen bond interactions		Hydrophobic interaction
		Amino acid	Bond length (Å)	Amino acid
31	− 14.3	GLN315 CYS130	2.0342 2.1743	MET86, PHE215, LEU207, VAL77, ALA127, PRO94
32	− 7.8	VAL94	2.6435	LEU207, TYR103, VAL78, PRO112
33	− 8.2	GLN115 ARG172	2.5498 2.1843	TRP112, PHE230, VAL128, PHE220, ALA127
41	− 15.2	ALA213 PRO95 GLN75	2.3029 2.2831 2.2188	LEU134, GLY232, PHE178, VAL218
35	− 10.6	THR72 ALA137 GLN315	2.1436 2.3419 2.1300	ALA233, GLY212, VAL228, TRP142, LYS185, PHE118
36	− 14.6	PHE114 CYS114 GLY212	2.2222 2.2034 2.3752	LYS126, PHE168, VAL78, ALA117, PRO169, TRP112
37	− 14.2	GLU88 PRO114 ALA147	2.0607 2.3915 2.5475	ALA167, LEU113, VAL78, PRO215, PHE148, ALA213
38	− 14	PRO74 ARG87 VAL95	2.4532 2.1064 2.5428	PRO265, TRP182, VAL112, TRP182, PHE148
39	− 10.7	VAL88 ASN76 GLN325	2.3683 2.0314 2.0280	PRO285, VAL218, LEU134,,ALA233, ALA233, VAL77
40	− 14.7	ALA137 GLN315 LEU127	2.2427 2.2374 2.5441	PHE138, VAL78, ALA167, VAL142, LEU143 CYS214
41	− 21.8	THR77 GLN385 ALA167 GLN385 ALA167	2.1123 2.6234 2.6012 2.1922 2.6302	ALA233, VAL78
42	− 15	GLN315 CYS134 ARG116	2.1512 2.2745 2.2572	VAL228, PHE168
43	− 4.1	–	–	ALA137, PHE213
44	− 6.2	VAL102	1.3493	PHE110, VAL71, ALA213
45	− 4.3		–	VAL79, PHE125, VAL158, PRO232, SER217
46	− 6.4	THR97	1.4282	VAL128, ALA237, TRP113, LEU134,
47	− 6	THR88	1.2493	ALA117, LEU124, VAL228, VAL63, TRP152
48	− 8.1	GLN185 SER227	2.1241 2.2425	PHE138, VAL78, CYS345, ALA233, PRO225
49	− 6.8	TRP122	1.7295	ALA137, VAL182,VAL78, PRO285, VAL79
50	− 14.1	ASP182 LYS126 GLN185	2.1272 2.1488 2.2392	LEU123, TRP112, ALA147, PRO255, VAL79
Ethambutol	− 5.7	ALA337	2.5983	–
Isoniazid	− 14.5	SER279 ALA337	2.2994 2.5295, 2.2466	PHE338, CYS345

3.3.2 Molecular Docking of Prominent Designed Ligands (Compound 41p)

The results of the docking studies affirmed the correlation between the activity values of ligand 41p and its binding affinity. The binding affinity value for ligand 41p was established to be − 24.3 kcal/mol as presented in Table 10 which were higher than the binding affinity value of the template 41 compound (− 21.9 kcal/mol) stated Table 8.

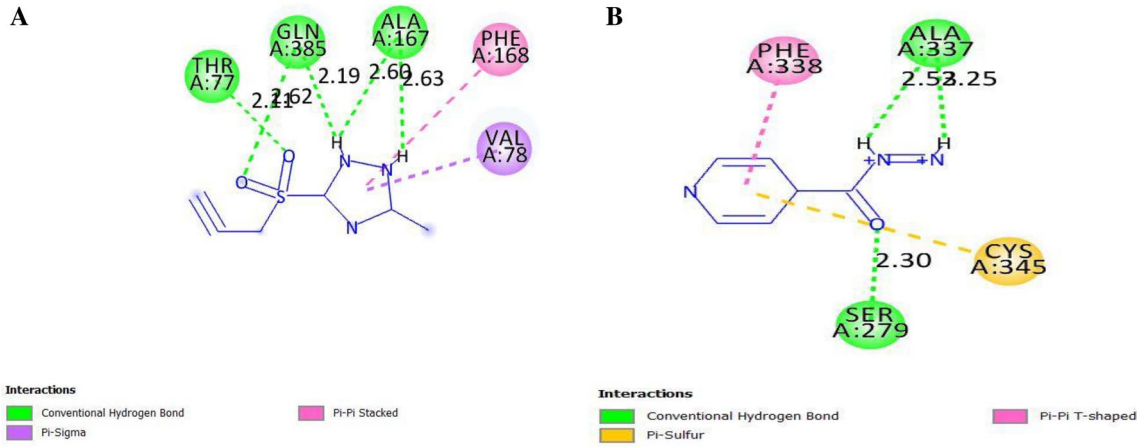


Fig. 6 **a** is the interactions between the ligand 41 and DNA gyrase. **b** is the interactions between Isoniazid and DNA gyrase

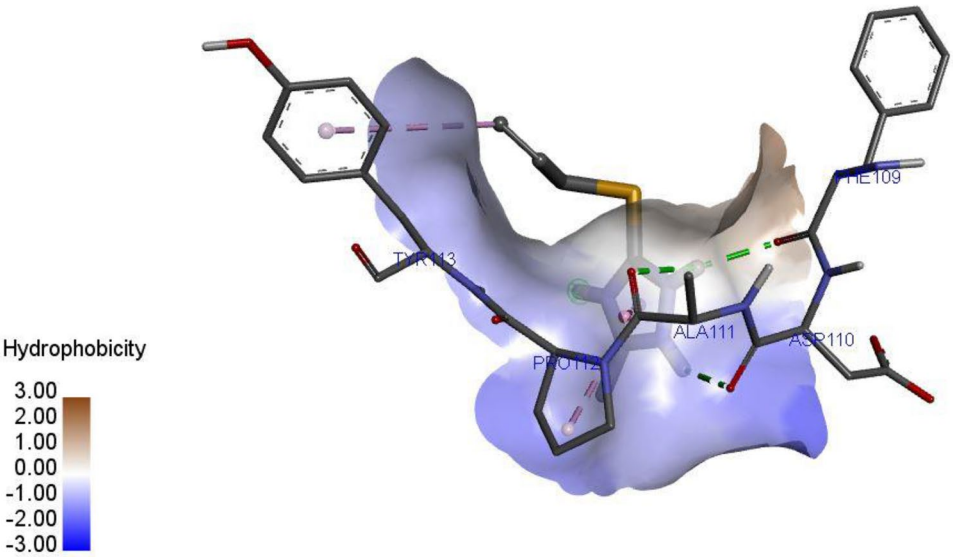


Fig. 7 Hydrophobic interactions between DNA gyrase and ligand 41

The interaction of the ligand 41p with DNA gyrase formed eight H-bonds with the enzyme. Six H-bonds formation was observed with of triazole (N–H group) of the ligand which acted as hydrogen donor with PRO B119, GLY B120, TRP B103 and VAL B278 of the target. Meanwhile, the ligand S=O group acted as hydrogen acceptor with formation of two hydrogen bonds with TRP B103 and SER B104 of the enzyme as presented in Fig. 10.

Number of H-bonding and distance has been reported to be the key reason influencing the binding affinity of receptor-ligand interaction [5, 16, 19]. Therefore, this reason provide structural insight to support the claim why designed compound 41p was able to binding efficiently in the binding pocket of the target enzyme Fig. 11.

4 Conclusion

Triazole derivatives was study using a theoretical method to select molecular descriptors to relate the structure of the derivatives against *M. tuberculosis*. The internal and external assessment confirmed that the built QSAR model is substantial, reliable and robust. Molecular descriptors; nHBint3, MATS2s, TDB9u, maxtsC, RDF110s and

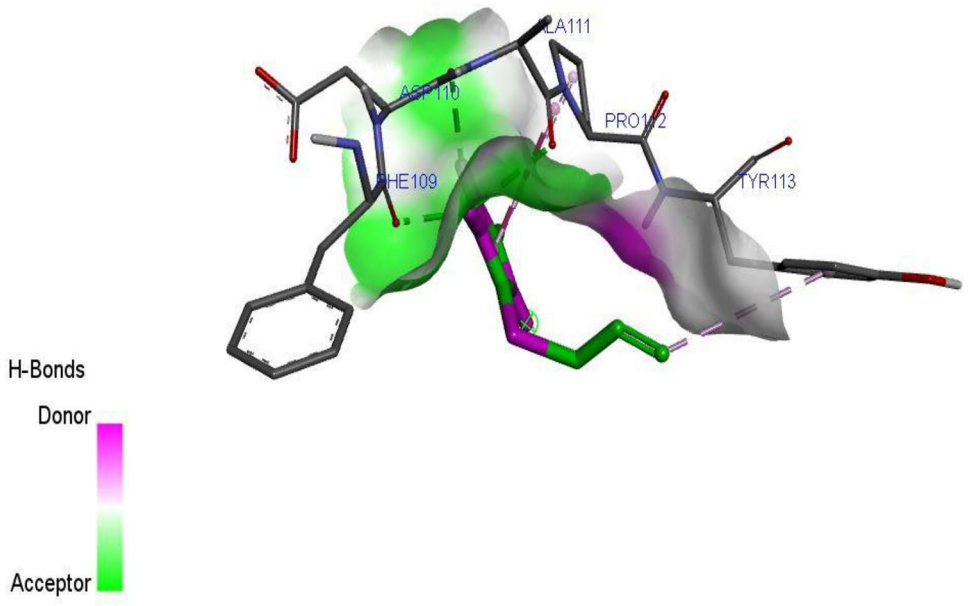


Fig. 8 H-bond interactions existing between DNA gyrase and ligand 41

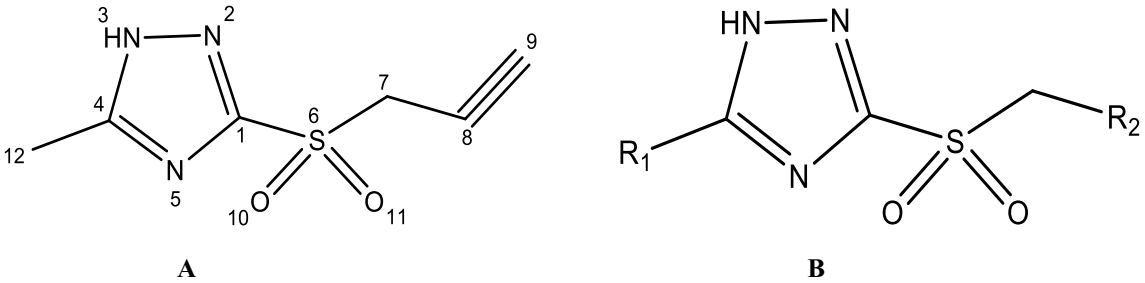


Fig. 9 **a** is the prime compound (41). **b** is the design template structure

RDF90i from the results have shown to be prominent descriptor needed to predict the biological activities of the studied compound. Furthermore, docking study indicates that compounds 41 of the derivatives with promising biological activity have the utmost binding energy of − 18.8 kcal/mol compared to the commended drugs; Isoniazid − 14.6 kcal/mol. Thereafter, compound 41 was used as a structure template to designed compounds with more efficient activities. Among hypothetical compounds designed; compounds 41p was experiential with highest activity against tuberculosis with more noticeable binding affinity of − 24.3 kcal/mol. The presumption of this research aid the medicinal chemists and pharmacist to design and synthesis a novel drug candidate against the tuberculosis. Moreover, in vitro and in vivo test could be carried out to validate the computational results.

Table 9 Compound designed, calculated descriptors and predicted activities

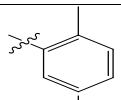
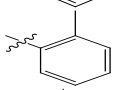
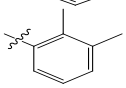
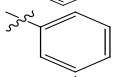
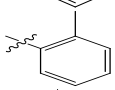
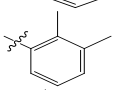
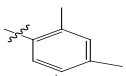
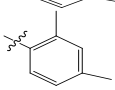
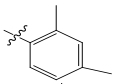
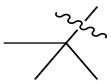
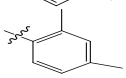
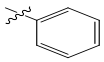
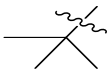
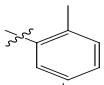
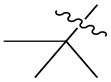
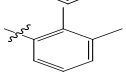
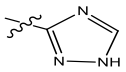
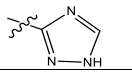
Compound ID	R_1	R_2	MATS2s	nHBint3	maxtsC	TDB9u	RDF90i	RDF110s	Predicted Activity (pBA)	Leverage
41a	H		3.5625	1	0.6436	67.5009	11.1572	2.7440	8.4166	0.5242
41b	CH_3		3.8701	1	0.3579	65.9916	24.2281	0.5357	8.5671	0.1995
41c	OCH_3		3.5078	1	0.3557	66.2984	22.5787	1.0094	8.4762	0.3201
41d	CH_3CH_2		4.2147	1	0.4277	67.0269	22.3770	0.8895	8.5522	0.3471
41e	CH_3CH_2		3.604	1	0.3397	62.2581	25.4329	0.9194	8.5824	0.3486
41f	CH_3CH_2		3.5081	1	0.4796	55.3564	21.4123	1.8581	8.5962	0.1411
41g	CH_3		3.4658	1	0.4066	55.3705	23.9847	0.9919	8.6421	0.4438
41h	CH_3CH_2		3.8857	1	0.4423	58.1968	22.8238	1.0226	8.5938	0.196
41i	OCH_3		3.2464	1	0.3575	61.0521	22.6697	0.9989	8.4879	0.1463
41j	H		3.8309	1	0.569	61.5988	28.2426	1.0205	8.5801	0.1449
41k			3.5883	1	0.5423	58.6879	21.7436	4.5788	8.6123	0.3584
41l			3.441	1	0.5782	60.9265	20.3561	5.5127	8.6122	0.3008
41m			3.5192	1	0.7376	56.5394	15.5796	6.5831	8.6091	0.4098
41n			3.5098	1	0.5598	60.7090	20.6710	4.9090	8.6073	0.4289
41o	H		3.2235	1	0.311	57.9263	28.8626	3.0143	8.5964	0.1886
41p	CH_3		3.2417	1	0.6268	62.5529	17.9120	4.7569	8.7619	0.4405

Table 10 Molecular docking interactions between *M. tuberculosis* target (DNA gyrase) ligands 41p

Ligand	Binding affinity (BA) Kcal/mol	Target	Hydrogen bond hydrophobic interaction		
			Amino acid	Bond length (Å)	Amino acid
41p	− 24.3	DNA gyrase	TRP B:103	2.31	TRP B:103
			SER B:104	2.74	
			GLY B:120	2.36, 2.40	
			PRO B:119	2.22	
			VAL B:278	2.79, 2.79	
			TRP B:103	2.68	

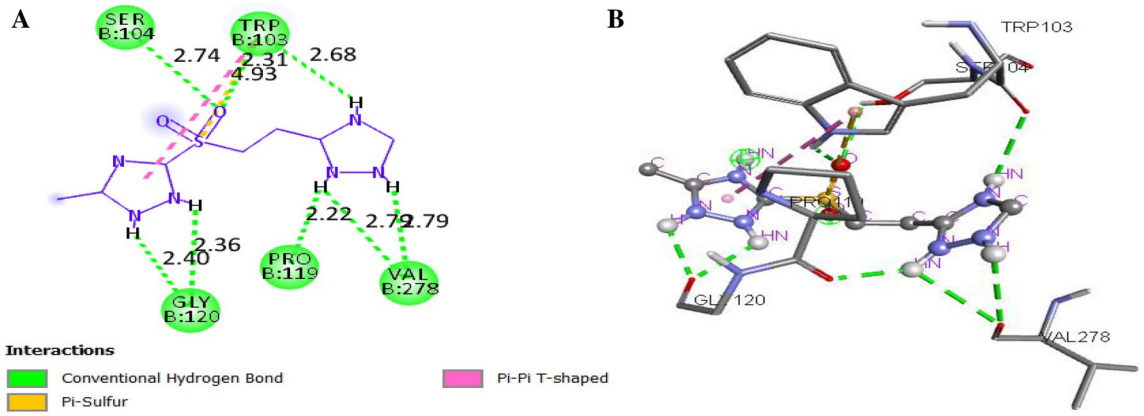


Fig. 10 **a** and **b** are the 2D and 3D interactions existing between ligand 41p and DNA gyrase

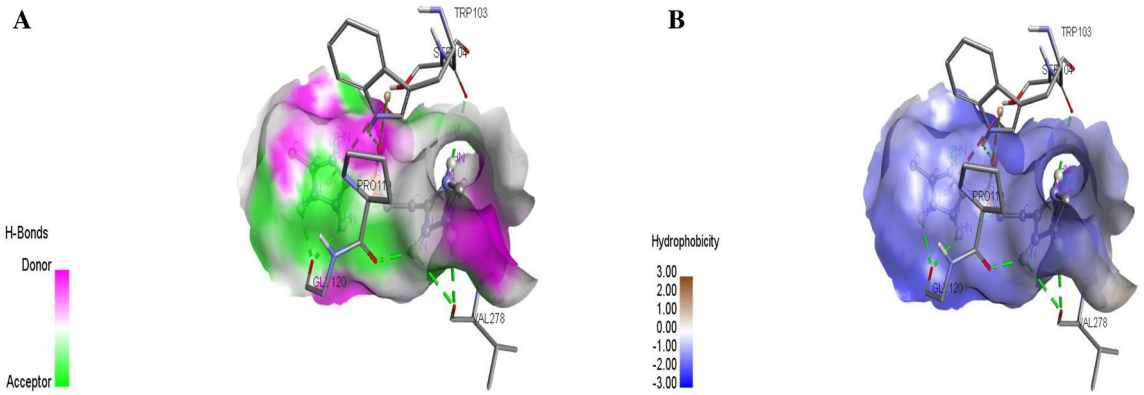


Fig. 11 **a** and **b** are the H-bond and hydrophobic interactions existing between ligand 41p and DNA gyrase

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. W.H. Organization, Others (2016) Tuberculosis Fact Sheet (No. 104) 2000, Site Accessed Www Who Intmediacentre-factsheetswho104enindex Html
2. Adeniji SE, Uba S, Uzairu A (2020) Theoretical modeling for predicting the activities of some active compounds as potent inhibitors against Mycobacterium tuberculosis using GFA-MLR approach. J King Saud Univ Sci 32:575–586
3. https://patents.justia.com/patent/8865910
4. Ogadimma AI, Adamu U (2016) Analysis of selected chalcone derivatives as Mycobacterium tuberculosis inhibitors. Open Access Libr J 3:1–13
5. Adeniji SE, Uba S, Uzairu A (2018) QSAR Modeling and Molecular Docking Analysis of Some Active Compounds against Mycobacterium tuberculosis Receptor (Mtb CYP121). J Pathog. Article ID 1018694
6. Adeniji SE, Uba S, Uzairu A (2018) A novel QSAR model for the evaluation and prediction of (E)-N'-benzylideneisonicotinohydrazide derivatives as the potent anti-mycobacterium tuberculosis antibodies using genetic function approach. Phys Chem Res 6:479–492
7. Eric GM, Uzairu A, Mamza PAA (2016) Quantitative structureactivity relationship (QSAR) study of the anti-tuberculosis activity of some quinolones. J Sci Res Rep 10:1–15
8. Singh P (2013) Quantitative structure-activity relationship study of substituted-[1,2,4] oxadiazoles as S1P1 agonists. J Curr Chem Pharm Sci 3:64–79
9. Arthur DE, Uzairu A, Mamza P et al (2018) In silico modelling of quantitative structure–activity relationship of multi-target anticancer compounds on k-562 cell line. Netw Model Anal Health Inform Bioinforma 7:11
10. Adeniji SE, Uba S, Uzairu A, ArthurDE (2019) A derived QSAR model for predicting some compounds as potent antagonist against mycobacterium tuberculosis: a theoretical approach. Adv Prev Med Article ID 5173786
11. Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. Mol Inform 22:69–77

12. Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN (2013) Some case studies on application of "rm2" metrics for judging quality of quantitative structure–activity relationship predictions: emphasis on scaling of response data. J Comput Chem 34:1071–1082
13. Veerasamy R, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK (2011) Validation of QSAR models-strategies and importance. Int J Drug Des Discov 3:511–519
14. Piton J, Petrella S, Delarue M, Andre´-Leroux G, Jarlier V, Aubry A, Mayer C (2010) Structural insights into the quinolone resistance mechanism of mycobacterium tuberculosis DNA gyrase. PLoS One 5:12245. https://doi.org/10.1371/journal.pone.0012245
15. http://www.rcsb.org/pdb/explore/litView.do?structureId=3IFZ
16. Adeniji SE, Uba S, Uzairu A (2018) Theoretical modeling and molecular docking simulation for investigating and evaluating some active compounds as potent anti-tubercular agents against MTB CYP121 receptor. Future J Pharm Sci 4:284–295
17. Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) In-silico activity prediction and docking studies of some 2, 9-disubstituted 8-phenylthio/phenylsulfinyl-9 h-purine derivatives as Anti-proliferative agents. Heliyon 6:e03158
18. Abdullahi M, Shallangwa GA, Uzairu A (2020) In silico QSAR and molecular docking simulation of some novel aryl sulfonamide derivatives as inhibitors of H5N1 influenza A virus subtype. Beni-Suef University. J Basic ApplSci 9:1–13
19. Patil R, Das S, Ashley S, Yadav L, Sudhakar A, Ashok KV (2010) Optimized hydrophobic interactions and hydrogen bonding at the target-ligand interface leads the pathways of drug-designing. PLoS One 8:1–10