European **J**ournal of **Chem**istry

Theoretical investigation of a few selected compounds as potent anti-tubercular agents and molecular docking evaluation: A multi-linear regression approach

Shola Elijah Adeniji [1,*], Abdulwahab Isiaka [1], Kalen Ephraim Audu [2] and Olajumoke Bosede Adalumo [3]

[1] *Department of Chemistry, Ahmadu Bello University, Zaria, Kaduna State 810107, Nigeria*
shola4343@gmail.com (S.E.A.), femiagie@gmail.com (A.I.)
[2] *Department of Biology, Ahmadu Bello University, Zaria, Kaduna State 810107, Nigeria*
kalenephraim14@gmail.com (K.E.A.)
[3] *Department of Science Education, Ajasin University, Akungba Akoko, Ondo State 342211, Nigeria*
olajumokeadalumo@gmail.com (O.B.A.)

* *Corresponding author at: Department of Chemistry, Ahmadu Bello University, Zaria, Kaduna State 810107, Nigeria.*
 e-mail: shola4343@gmail.com (S.E. Adeniji).

RESEARCH ARTICLE

KEYWORDS

ABSTRACT

Emergence of multi-drug resistant strains of *Mycobacterium tuberculosis* to the available drugs has demanded for the development of more potent anti-tubercular agents with efficient pharmacological activities. Time consumed and expenses in discovering and synthesizing new drug targets with improved biological activity have been a major challenge toward the treatment of multi-drug resistance strain *M. tuberculosis*. 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; AATS7s, VR1-Dzi, VR1-Dzs, SpMin7-Bhe and RDF110i. The internal validation test for the derived model was found to have correlation coefficient (R^2) of 0.8875, adjusted correlation coefficient (R^2_{adj}) value of 0.8234 and leave one out cross validation coefficient (Q_{cv}^2) value of 0.8012 while the external validation test was found to have (R^2_{test}) of 0.7961 and Y-randomization Coefficient (cRp^2) of 0.6832. Molecular docking shows that ligand 13 of 2,4-disubstituted quinoline derivatives have promising higher binding score of -18.8 kcal/mol compared to the recommended drugs; isoniazid -14.6 kcal/mol. The proposed QSAR model and molecular docking studies will provides valuable approach for the modification of the lead compound, designing and synthesis more potent anti-tubercular agents.

Journal website: www.eurjchem.com

1. Introduction

Over the years, tuberculosis has been a serious threat to mankind which is caused by specie of bacteria known as *Mycobacterium Tuberculosis* (TB). World Health Organization in 2018, 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]. Tuberculosis may infect any part of the body, but most commonly occurs in the lungs (known as pulmonary tuberculosis). Extra-pulmonary TB occurs when tuberculosis develops outside of the lungs, although extra-pulmonary TB may coexist with pulmonary TB. 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). 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.*

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 [3-7] 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

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

S/N*	Molecular structure	Observed activity (pA)	Calculated activity (pA)	Residual
1 [a]	2-(2-(4-Methoxybenzylidene)hydrazinyl)-N-phenylquinoline-4-carboxamide	9.4979	9.731930	-0.234030
2	2-(2-(4-Methoxybenzylidene)hydrazinyl)-N-phenylquinoline-4-carboxamide	6.9772	6.896778	0.080422
3 [a]	N-Benzyl-2-(2-(pyridin-3-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.2608	6.510442	0.750358
4	N-Benzyl-2-(2-(furan-2-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.1707	6.972982	0.197718
5 [a]	N-Benzyl-2-(2-(thiophen-2-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.4233	7.152527	0.270773
6	2-(2-(Anthracen-9-ylmethylene)hydrazinyl)-N-benzylquinoline-4-carboxamide	7.2838	6.985668	0.298132
7 [a]	N-Benzyl-2-(2-((4-methoxynaphthalen-1-yl)methylene)hydrazinyl)quinoline-4-carboxamide	7.1472	7.478650	-0.531450
8 [a]	N-Benzyl-2-(2-(2-methylpropylidene)hydrazinyl)quinoline-4-carboxamide	7.6035	7.712630	-0.109130
9 [a]	N-Benzyl-2-(2-propylidenehydrazinyl)quinoline-4-carboxamide	7.2938	6.495725	0.898075
10	N-Benzyl-2-(2-(4-methoxybenzylidene)hydrazinyl)quinoline-4-carboxamide	7.2630	7.786450	-0.623450
11	N-(5-Phenylpentyl)-2-(2-(pyridin-4-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.4772	7.411826	0.065374
12	2-(2-(Furan-2-ylmethylene)hydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	7.0807	7.172820	-0.092120
13	N-Benzyl-2-(2-benzylidenehydrazinyl)quinoline-4-carboxamide	9.6090	9.627790	-0.018790
14	N-(5-Phenylpentyl)-2-(2-(thiophen-2-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.2747	7.224153	0.050547
15 [a]	2-(2-(Anthracen-9-ylmethylene)hydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	7.4091	7.674090	-0.264990
16	2-(2-((4-Methoxynaphthalen-1-yl)methylene)hydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	7.7412	7.318700	0.422500
17	2-(2-(2-Methylpropylidene)hydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	7.6688	7.273758	0.395042
18	2-(2-Benzylidenehydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	6.2688	6.325600	-0.056800
19	2-(2-(4-Methoxybenzylidene)hydrazinyl)-N-(5-phenylpentyl)quinoline-4-carboxamide	7.6970	7.737650	-0.040650
20	(2-(2-(4-Methoxybenzylidene)hydrazinyl)quinolin-4-yl)(morpholino)methanone	6.8414	6.809542	0.031858
21	(4-Methylpiperazin-1-yl)(2-(2-(pyridin-4-ylmethylene)hydrazinyl)quinolin-4-yl)methanone	7.3673	7.357741	0.009559
22	(2-(2-(Furan-2-ylmethylene)hydrazinyl)quinolin-4-yl)(4-methylpiperazin-1-yl)methanone	7.1891	7.392020	-0.202920
23 [a]	(2-(2-((4-Methoxynaphthalen-1-yl)methylene)hydrazinyl)quinolin-4-yl)(4-methylpiperazin-1-yl)methanone	7.2022	7.500520	-0.298320
24	(4-Methylpiperazin-1-yl)(2-(2-(2-methylpropylidene)hydrazinyl)quinolin-4-yl)methanone	7.7696	7.486908	0.282692
25	(2-(2-Benzylidenehydrazinyl)quinolin-4-yl)(4-methylpiperazin-1-yl)methanone	6.7716	6.752730	-0.481130
26	(2-(2-(4-Methoxybenzylidene)hydrazinyl)quinolin-4-yl)(4-methylpiperazin-1-yl)methanone	7.4420	7.492240	-0.050240
27	N-Phenyl-2-(2-(thiophen-2-ylmethylene)hydrazinyl)quinoline-4-carboxamide	7.3209	7.025132	0.295768

* Superscript a represent the test set.

predictability and molecular docking study against *M-tuberculosis*.

2. Experimental

2.1. Data set

The molecules comprising the derivatives of 2,4-disubstituted quinoline reported as anti-*mycobacterium tuberculosis* that were used in this study were obtained from the literature [8]. The biological activities of these compounds and the list of the compounds were presented in Table 1. The observed structures and the biological activities of these compounds were presented in Table 1.

2.2. Geometrical optimization

Spartan 14 software version 1.1.4 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 basic set [3,7].

2.3. Calculation of 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 were calculated with the aid of PaDEL descriptor software version 2.20 [9] 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 [9] were subjected to normalization using Equation (1) [8].

$$D = \frac{d_1 - d_{min}}{d_{max} - d_{min}} \qquad (1)$$

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 so as to remove redundant descriptors.

2.5. Splitting of data 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 [3].

2.6. Building of QSAR Models and internal validation

The QSAR models were built by adopting the Genetic Function Approximation (GFA) technique incorporated in the Material Studio software version 8.0 [https://www.3ds biovia.com/products/collaborative-science/biovia-materials-studio/] to select the optimum descriptors for the training set. Meanwhile, Multi-linear regression Approach (MLR) [7] 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. 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 Equation (2) was used define applicability domain space ±3 for outlier molecule [10,11].

$$h_i = M_i(M^T M)^{-1} M_i^T \qquad (2)$$

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 Equation (3) is the limit boundary to check for an influential molecule.

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

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 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 [11,12]. Random shuffling of the data was performed on the training set following the principle laid by [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 Equation (4) must be ≥ 0.5 in order to establish the robustness of the model.

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

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

2.9. Assertion 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 for any QSAR model [11,13-15] 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 used in the study was obtained from protein data bank with PDB code 31FZ [16]. 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 [https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/]. The prepared receptor was transported into the Pyrx software in order to make it a macro molecule.

2.10.2. Preparation of the optimum ligand

The optimum ligand was geometry optimized with Spartan 14 software at Density Functional Theory (DFT) level in order to attain the most stable conformer of the inhibitor at minimum energy [4]. The optimized molecule/ligand was saved as a pdb file in a folder. The optimized structure was then saved in PDB file format and transported into the Pyrx software in order to make the inhibitor as a micro molecule (ligand) [17].

2.10.3. Receptor-ligand docking with PyRx virtual screening software

PyRx [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. In this study, AutoDock Vina was only used to carry out the molecular docking. In order to perform protein-ligand docking, both the ligand and receptor (DNA gyrase) was converted from pdb files to pdbqt (protein data bank, partial charge and atom type) files (Vina input file format). The conversion of pdb files to pdbqt files (Vina input file format) was done by launching the PyRx virtual screening software in order compute the Binding Score (kcal/mol). The more the negative the binding score, the better the orientation of the ligand in the binding site of DNA gyrase. The docked results were compiled, visualized and analyzed using Discovery Studio Visualizer [17].

3. Results and discussion

3.1. QSAR studies

Optimum QSAR model for predicting the derivatives of 2,4-disubstituted quinoline against *M. tuberculosis* was successfully achieved by adopting the combination of computational and theoretical method. Data set comprises of 27 compounds was partitioned into 19 training set and 7 test set using Kennard and Stone algorithm method [4]. The 20 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.

The observed activities reported in literature and the calculated activities calculated for all the anti-tubercular compounds were presented in Table 1. The difference between the observed activities and calculated 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 selected by Genetic Function Approximation (GFA) approach were reported in Table 2.

Various statistical analyses were conducted on the calculated descriptors in order to check the validity of the built model as reported in Table 3. 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 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 3 affirm that the descriptors were significantly orthogonal to each order since there is no inter-correlation between them.

Table 2. List of some descriptors used in the QSAR optimization model.

No	Descriptors symbols	Name of descriptor(s)	Class
1	AATS7s	Average Broto-Moreau autocorrelation - lag 7/weighted by I-state	2D
2	VR1-Dzi	Randic-like eigenvector-based index from Barysz matrix/weighted by first ionization potential	2D
3	VR1-Dzs	Randic-like eigenvector-based index from Barysz matrix/weighted by I-state	2D
4	SpMin7-Bhe	Smallest absolute eigenvalue of Burden modified matrix - n7/weighted by relative Sanderson electronegativities	2D
5	RDF110i	Radial distribution function - 110 / weighted by relative I-state	3D

Table 3. Statistical parameters that influence the model.

Descriptor	Standard regression coefficient (b_j)	Mean effect (ME)	p-value (Confidence interval)	VIF	Standard error
AATS7s	-0.4082	-0.4178	2.38×10^{-5}	1.3429	0.0067
VR1-Dzi	0.2248	0.2233	6.19×10^{-4}	3.8139	0.0298
VR1-Dzs	0.3393	0.3478	4.34×10^{-6}	1.9125	0.0984
SpMin7-Bhe	-0.6830	-0.6922	2.72×10^{-5}	1.6912	0.0049
RDF110i	0.8528	0.8847	5.02×10^{-3}	2.2013	0.0054

Table 4. Pearson's correlation coefficient for the descriptor used in the QSAR model.

Inter-correlation	AATS5e	VR1-Dzs	SpMin7-Bhe	TDB9e	RDF110i
AATS7s	1.0000				
VR1-Dzi	0.4112	1.0000			
VR1-Dzs	0.3315	0.2833	1.0000		
SpMin7-Bhe	0.1395	-0.5232	-0.3092	1.0000	
RDF110i	0.0719	-0.0831	0.0384	0.0943	1.0000

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 3 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 was found to be (p < 0.05) as presented in Table 3. 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 [17]. 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 descriptor. The correlation coefficient between each descriptors reported in Table 4 were all < ±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 model is reliable and robust were presented in Table 5. These results were all in full agreement with general validation criteria resented in Table 5 to truly indorse that the stability and robustness of the model is valid.

3.2. Model built

The coefficient of Y-Randomization (cR_p^2) with significant value of 0.6832 greater than threshold value of 0.5 reported in Table 6 provide a reasonable supports that the model built is robust and not just by chance.

pBA = - 6.631974301 × AATS7s + 0.001749220 × VR1_Dzi
+ 0.060901621× VR1_Dzs - 6.088397140× SpMin7_Bhe
+ 0.097016191× RDF90i + 21.24098010   (5)

3.3. Mechanistic information of the selected descriptors in built

AATS7s is Average Broto-Moreau autocorrelation – lag 7/ weighted by I-state auto-correlation descriptor. The negative mean effect of this descriptor indicates that the inhibitory activity will decrease with hydrogen bonds of path length 3. VR1-Dzi is Randic-like eigenvector-based index from Barysz matrix/weighted by first ionization potential while VR1-Dzs is Randic-like eigenvector-based index from Barysz matrix/weighted by I-state. From the model generated in this study, these descriptors have positive coefficient and positive mean effect value. SpMin7-Bhe descriptors have been proposed as chemical structure descriptors derived from a new representation of molecular structure. The Sign of the coefficient of this descriptor is negative implying that groups having more branching are diminishes the activities of the active compounds toward Mycobacterium tuberculosis. RDF90i is 3D radial distribution function at 2.5 inter-atomic distance weighted by atomic masses. RDF90i with positive mean effect (MF) indicates positive impact on the activity.

The graphical representation to show the degree of correlation between the calculated activities and observed activities of the training and test set were shown in Figure 1 The correlation coefficient (R^2) value of 0.8875 and 0.7961 for both the training set and test set shows that there is a high correlation existing between the calculated activities and observed activities of the training and test set which were also in agreement with the accepted QSAR threshold values reported in Table 5.

The residual plot shown in Figure 2 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.

The Williams plot to show the Applicability Domain space (AD) is shown in Figure 3. 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. Also all the compounds fall within the defined warning leverage (h* = 0.95). Therefore no compound is said to be an influential molecule.

3.4. Docking studies

3.4.1. Evaluation binding score

Elucidation of interaction type and the binding mode between the inhibitory compound and target (DNA gyrase) was achieved via molecular docking studies.

Table 5. Validation parameters for each model using Multi-linear Regression (MLR).

No	Validation parameters	Formula	Threshold	Model
Internal validation				
1	Friedman LOF	$\dfrac{SEE}{\left(1-\dfrac{C+d\times p}{M}\right)^2}$	Significantly low	0.0484
2	R-squared	$1-\left[\dfrac{\Sigma\left(Y_{exp-Y_{pred}}\right)^2}{\Sigma\left(Y_{exp-\overline{y}_{training}}\right)^2}\right]$	$R^2 > 6$	0.8823
3	Adjusted R-squared	$\dfrac{R^2-P(n-1)}{n-p+1}$	$R^2_{adj} > 0.6$	0.8234
4	Cross validated R-squared (Q^2_{cv})	$1-\left[\dfrac{\Sigma\left(Y_{pred-Y_{exp}}\right)^2}{\Sigma\left(Y_{exp-\overline{Y}_{training}}\right)^2}\right]$	$Q^2 > 6$	0.8012
5	Significant Regression			Yes
6	Critical SOR F-value (95%)	$\dfrac{\Sigma\left(Y_{pred-Y_{exp}}\right)^2}{p}/\dfrac{\Sigma\left(Y_{pred-Y_{exp}}\right)^2}{N-p-1}$	$F_{(test)} > 2.09$	3.4528
7	Replicate points		0	0
8	Computed observed error		0	0
9	Min expt. error for non-significant LOF (95%)		Significantly low	0.0983
Model randomization				
10	Average of the correlation coefficient for randomized data ($\overline{R}_r$)		$\overline{R} < 0.5$	0.4543
11	Average of determination coefficient for randomized data ($\overline{R}^2_r$)		$\overline{R}^2_r < 0.5$	0.2841
12	Average of leave one out cross-validated determination coefficient for randomized data ($\overline{Q}^2_r$)	$Q^2_{cv}=1-\left[\dfrac{\Sigma\left(Y_{pred-Y_{exp}}\right)^2}{\Sigma\left(Y_{exp-\overline{Y}_{training}}\right)^2}\right]$	$\overline{Q}^2_r < 0.5$	-1.4322
13	Coefficient for Y-randomization (cR^2_p)	$R^2\times\left(1-\sqrt{\left|R^2-\overline{R}^2_r\right|}\right)$	$cR^2_p > 0.6$	0.6832
External validation				
14	Slope of the plot of Observed activity against Calculated activity values at zero intercept (K)	$\dfrac{\Delta Y_{Obs}}{\Delta Y_{cal}}$	$0.85 < k < 1.15$	1.019
15	Slope of the plot of Calculated against Observed activity at zero intercept (k′)	$\dfrac{\Delta Y_{Obs}}{\Delta Y_{cal}}$	$0.85 < k < 1.15$	0.8034
16	$/r^2_0 - r'^2_0/$		< 0.3	0.0152
17	$\dfrac{r^2-r^2_0}{r^2}$		< 0.1	0.0021
18	$\dfrac{r^2-r'^2_0}{r^2}$		< 0.1	0.0543
19	R^2_{test}	$1-\dfrac{\Sigma\left(Y_{ext}-\hat{Y}_{ext}\right)^2}{\Sigma\left(Y_{ext}-\overline{Y}\right)^2}$	$R^2_{pred} > 0.6$	0.7954

Table 6. Y-Randomization parameters test.

Model	R	R^2	Q^2
Original	0.8823	0.8234	0.8012
Random 1	0.5451	0.2647	-1.0792
Random 2	0.4764	0.2696	-0.2065
Random 3	0.8381	0.4868	0.0145
Random 4	0.5734	0.3321	-0.0964
Random 5	0.3525	0.1261	-0.8451
Random 6	0.6987	0.2545	0.0231
Random 7	0.4447	0.1635	-0.9035
Random 8	0.5351	0.2688	-0.6712
Random 9	0.4466	0.2656	-0.6872
Random 10	0.6956	0.3963	-0.0026
Random models parameters			
Average r	0.4543		
Average r^2	0.2841		
Average Q^2	-1.4322		
cRp^2	0.6832		

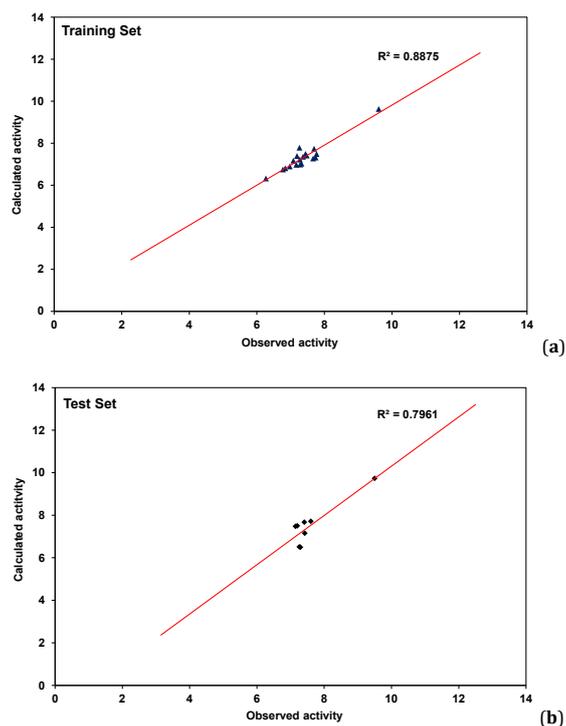**Figure 1.** (**a**) The plot of calculated activity against observed activity of training set, (**b**) The plot of calculated activity against observed activity of test set.
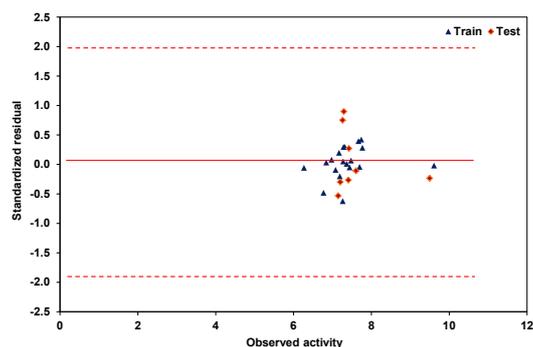

Figure 2. Plot of standardized residual activity versus observed activity.
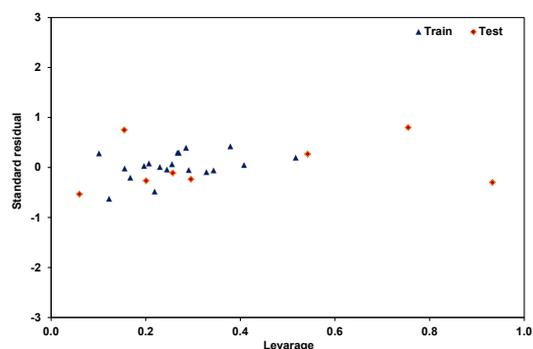

Figure 3. The Williams plot of the standardized residuals versus the leverage value.

The docking results clearly show that the binding score for the optimum compound number 13 correlates with its activity value. For target enzyme, binding score was found to be -18.8 kcal/mol as reported in Table 7. The binding score of recommended drugs; isoniazid (-14.6 kcal/mol) was found to be lesser than the binding score of the compound 13. This indicate that this compounds could serve as better anti-tubercular drug and can be improve by structure base design.

3.4.2. Determination of bond type and bond length

Discovery Studio Visualizer software was used to show the bond length and type of interaction (bond) that exist between the optimum compound and the target site.

Table 7. Binding score, hydrogen bond and hydrophobic interaction of the optimum ligands with *M. tuberculosis* target (DNA gyrase).

Ligand	Binding score (BA) Kcal/mol	Target	Hydrogen bond hydrophobic interaction		
			Amino acid	Bond length (Å)	Amino acid
13	-18.8	DNA gyrase	ARG98	3.3701	VAL97, PRO124, VAL97, ASP94, ASP122, PRO123
			SER118	2.8704	
			GLY120	1.9128	
			GLY120	3.2821	
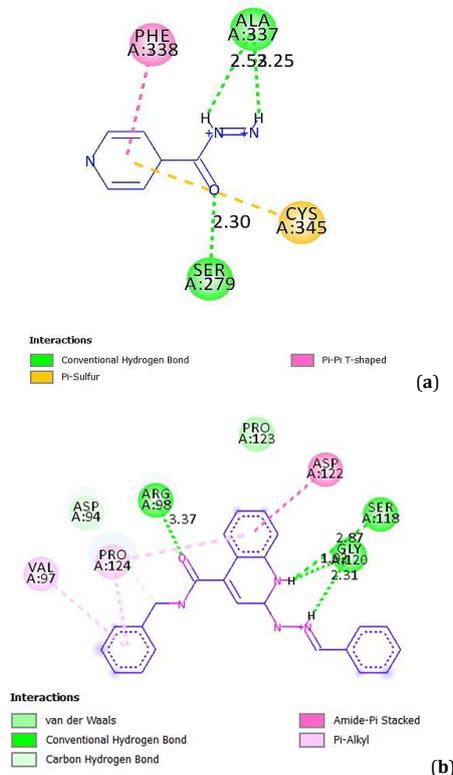
Figure 4. (**a**) The 2D interactions between DNA gyrase and ligand 13 of 2,4-diquiloline derivatives, (**b**) The 2D interactions between DNA gyrase and isoniazid.
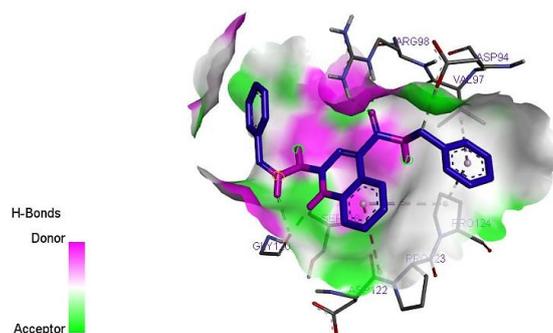
Figure 5. H-bond interaction between the ligand 13 of 2,4-disubstituted quinoline derivatives and *M. tuberculosis* target (DNA gyrase).

The 3D and 2D interaction of ligand 13 was shown in Figure 4. Ligand 13 formed four hydrogen bonds (3.3701, 2.8704, 1.9128 and 3.2821 Å) with ARG98, SER118, GLY120 and GLY120 of the target while hydrophobic interactions were observed with VAL97, PRO124, VAL97, ASP94, ASP122 and PRO123 of the target site.

Ligand 13 formed a total of four hydrogen bonds with target site of DNA gyrase. The C=O of the ligand acts as hydrogen acceptor and formed one hydrogen bond with ARG98 of the target. The N-H group (hydropyridine) of the ligand acts as hydrogen donor and formed two hydrogen bonds with GLY120 of the target. The N-H group (hydrazine) of the ligand also acts as hydrogen donor and formed a

hydrogen bond with SER118 of the target site. The hydrogen bond formations alongside with the hydrophobic interaction provide evidence that ligand 13 of 2,4-disubstituted quinolone derivatives is a potent inhibitors against DNA gyrase receptor. Elucidations of hydrogen donor and hydrogen acceptor region were shown in Figure 5.

3.4.3. Determination of bond type and bond length between recommended anti-tubercular drug (isoniazid) and DNA gyrase

The 2D interaction of the recommended anti-tubercular drugs (isoniazid) with the DNA gyrase target site were in

Figure 4. Isoniazid formed three hydrogen bonds (2.29943, 2.52954 and 2.24657 Å) with SER279, ALA337 and ALA337 while hydrophobic bonds were observed with CYS345 of the target site. It is an evidence that increase in number of hydrogen bonds observed in ligand 13 of 2, 4-disubstituted quinoline derivatives account for its higher binding score -18.8 kcal/mol compared to the recommended drugs; "Isoniazid - 14.6 kcal/mol".

4. Conclusion

QSAR generated models was able to predict the activity of 2,4-disubstituted derivatives as a potent anti-tubercular agent and molecular docking studies carry out help to understand and elucidate the interaction between the inhibitor compounds and the target site of *M. tuberculosis* (DNA gyrase). The model derived was subjected to internal and external validation test to confirm that the built QSAR model is significant, robust, and reliable. From the results, it is concluded that 2,4-disubstituted quinoline derivatives can be modelled using molecular descriptors; AATS7s, VR1-Dzi, VR1-Dzs, SpMin7-Bhe and RDF110i. Molecular docking simulation shows that ligand 13 of 2,4-disubstituted quinoline derivatives have promising higher binding score of -18.8 kcal/mol compared to the recommended drugs; Isoniazid -14.6 kcal/mol. The built QSAR model and docking studies could serve as a vital tool for pharmaceutical as well as medicinal chemists to design and synthesis novel anti-tubercular drugs with better activities against *M. tuberculosis*.

Disclosure statement [DS]

Conflict of interests: The authors declare that they have no conflict of interest.
Author contributions: All authors contributed equally to this work.
Ethical approval: All ethical guidelines have been adhered.
Sample availability: Samples of the compounds are obtained from the literature.

ORCID [iD]

Shola Elijah Adeniji
[iD] http://orcid.org/0000-0002-7750-8174
Abdulwahab Isiaka
[iD] http://orcid.org/0000-0001-7826-2279
Kalen Ephraim Audu
[iD] http://orcid.org/0000-0002-2403-8893
Olajumoke Bosede Adalumo
[iD] http://orcid.org/0000-0003-0609-977X

References

[1]. World Health Organization, Tuberculosis, Retrieved Feb 01, 2020, from http://www.who.int/news-room/fact-sheets/detail/tuberculosis.
[2]. Hansch, C.; Kurup, A.; Garg, R.; Gao, H. *Chem. Rev.* **2001**, *101*, 619-672.
[3]. Adeniji, S. E.; Uba, S.; Uzairu, A. *J. Pathog.* **2018**, 1018694.
[4]. Ogadimma, A. I.; Adamu, U. *Open Access Lib. J.* **2016**, *3*, 1-13.
[5]. Eric, G. M.; Uzairu, A.; Mamza, P. A. *J. Sci. Res. Rep.* **2016**, *10*, 1-15.
[6]. Adeniji, S. E.; Uba, S.; Uzairu, A. *Phys. Chem. Res.* **2018**, *6*, 479-492.
[7]. Adeniji, S. E.; Uba, S.; Uzairu, A. *J. King Saud University-Sci.* **2020**, *32(1)*, 67-74.
[8]. Nayyar, A.; Jain, R. *Indian J. Chem. B* **2008**, *47*, 117-128.
[9]. Yap, C. W. *J. Comput. Chem.* **2011**, *32*, 1466-1474.
[10]. Singh, P. *J. Curr. Chem. Pharm. Sci.* **2013**, *3(1)*, 64-79.
[11]. Adeniji, S. E.; Uba, S.; Uzairu, A.; Arthur, D. E. *Adv. Preventive Med.* **2019**, 5173786, 1-18.
[12]. Tropsha, A.; Gramatica, P.; Gombar, V. K. *Mol. Inform.* **2003**, *22*, 69-77.
[13]. Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P. K.; Kar, S.; Das, R. N. *J. Comput. Chem.* **2013**, *34*, 1071-1082.
[14]. Veerasamy, R.; Rajak, H.; Jain, A.; Sivadasan, S.; Varghese, C. P.; Agrawal, R. K. *Int. J. Drug Des. Discover.* **2011**, *2(3)*, 511-519.
[15]. Adeniji, S. E.; Uba, S.; Uzairu, A. *Netw. Model. Anal. Health Inform. Bioinforma.* **2020**, *9(8)*, 1-18.
[16]. Piton, J.; Petrella, S.; Delarue, M.; Andre-Leroux, G.; Jarlier, V.; Aubry, A.; Mayer, C. Retrieved Feb 01, 2020, from https://www. rcsb.org/structure/3IFZ
[17]. Adeniji, S. E.; Uba, S.; Uzairu, A. *Future. J. Pharm. Sci.* **2018**, *4*, 284-295.