*Egypt. J. Chem.* **Vol.62,** No. 7, pp. 1235 - 1247 (2019)

**Egyptian Journal of Chemistry**
**http://ejchem.journals.ekb.eg/**

# Geometrical and Topological Descriptors for Activities Modeling of Some Potent Inhibitors against *Mycobacterium tuberculosis*: A Genetic Functional Approach

**Shola Elijah Adeniji[1*] Sani Uba and Adamu Uzairu**
*[*1]Department of Chemistry, Ahmadu Bello University (ABU) Zaria, Kaduna State, Nigeria.*

IMPROVEMENT on more potent anti-tuberculosis agents is as a result of emergence of multi-drug resistant strains of *M. tuberculosis*. Syntheses of novel compounds are usually carried out via trial approach with lots of errors which is time consuming and expensive. QSAR is a theoretical approach, which has the potential to reduce the aforementioned problem in discovering new potent drugs against *M. tuberculosis*. This approach was employed to develop multivariate QSAR model to correlate the chemical structures of the 1,2,4-triazole analogues with their observed activities using a theoretical approach. In order to build the robust QSAR model, the best descriptors that could efficiently predict the activities of the inhibitory agents were selected by employing Genetic Function Approximation (GFA) as a modeling tool. Correlation coefficient ($R^2$) of 0.9142, cross validation coefficient (Q_cv^2) value of 0.8324 and adjusted correlation coefficient ($R^2$adj) value of 0.8851 were the internal validation test conducted to access the derived model while ($R^2$test) of 0.7495 and Y-randomization Coefficient (cR_p^2) of 0.7334 were the external validation tests to confirm the robustness of the built model. The proposed QSAR model provides a valuable approach for modification of the lead compound, design and synthesizing more potent anti-tubercular agents.

**Keywords:** Applicability domain, Genetic function approximation, QSAR, Tuberculosis, Triazole.

## Introduction

Tuberculosis (TB) is the most deadly bacterial disease caused by species of bacteria known as *Mycobacterium tuberculosis*. In (2013), World Health Organization (WHO) estimated death of 1.5 million people, 9.0 million people living with tuberculosis and 360,000 people whom were HIV positive [1]. At present, pyrazinamide (PZA), para-amino salicylic acid (PAS), isoniazide (INH) and rifampicin (RMP) are the current drugs administered to patients suffering from tuberculosis.

The resistances of the *M. tuberculosis* toward the current drugs led to development of QASR approach that is fast and precise which could be able to predict the biological activities for the new compounds against *M. tuberculosis*.

Mean while, a theoretical approach; quantitative structure activities relationships (QSARs) is one of the most widely used computational methods which helps in drug design and prediction of drugs activities [2]. QSAR model is a mathematical linear equation which relates the molecular structures of the compounds and their biological activities. In this research, a data set of 1,2,4-triazole derivatives which had been synthesized and evaluated as anti - *Mycobacterium tuberculosis* [3] have been selected for QSAR study. Few researchers [4–7] have established relationship between some anti-

*Corresponding author Email: shola4343@gmail.com , Tel.: 07060511720
Received 20/11/2018; Accepted 20/01/2019
DOI: 10.21608/ejchem.2019.6296.1531

tubercular inhibitor's like quinolone, chalcone, pyrrole and 7-methyijuglone using QSAR approach. Therefore, this study is aimed to relate the structures of 1,2,4-triazole derivatives with their respective biological activities by establishing a valid QSAR model for purpose of predicting their respective activities and for further design of more promising anti-tubercular agents.

## Materials and Methods

*Data set*

The derivatives of 1,2,4-triazole as potent anti-mycobacterium tuberculosis that were used in this research were selected from the literature [8]. Biological activities of 1,2,4- derivatives as potent anti- anti-mycobacterium tubercular agents wereinitially expressed in percentage (%) and then converted to logarithm unit using equation (1) in order to increase the linearity and approach normal distribution of theactivities values. The chemical structures alongside with their IUPAC names and biological activities of these compounds were presented in Table 1.

$$1 \quad pBA = \log\left[\left(\frac{Molecular\ \ weig\ ht_{(g/mol)}}{Dose_{(g/ml)}}\right)\left(\frac{\text{percentage (\%)}}{100-\text{percentage (\%)}}\right)\right] \qquad \text{(Eq.1)} \qquad (5)$$

pBA = log

*Structure optimization*

In order for the molecules to attain a stable conformer at a minimal energy, all the molecules were geometrically optimized with the aid of Spartan 14 V1.1.4 by employing Molecular Mechanics Force Field (MMFF) count to remove strain energy and later subjected to Density Functional Theory (DFT) by utilizing the (B3LYP) basic set (5).

*Molecular descriptor calculation*

Descriptor is a mathematical logic that describes the properties of a molecule based on the correlation between the structure of the compound and its biological activities. Descriptors calculation for all the inhibitory compounds was achieved using PaDEL-Descriptor software V2.20. A total of 1876 molecular descriptors were calculated.

*Normalization of data and pretreatment*

The values for the calculated descriptors' were normalized using Equation 2 so that each variable will have the same prospect at the inception so as to sway the model [9].

$$Y = \qquad \mathbf{Y} = \frac{Y_1 - Y_{min}}{Y_{max} - Y_{min}} \qquad \text{(Eq.2)}$$

Where $Y_1$ is the descriptor value for each molecule, $Y_{min}$and $Y_{max}$ are the minimum and maximum value for each descriptors column of Y. After successful normalization of the data, the data were further subjected to pretreatment using data pretreatment software obtained from DTC lab in order to remove noise and redundant data.

*Data division into training and test set*

Kennard and Stone's algorithm approach was employed in this study to divide the data set into a training set and a test compounds in proportion of 70 to 30%. The training set was used to develop the QSAR model while the test was used to confirm the developed model.

*Development of the model*

Multi-linear regression approach (MLR) is a strategy used to develop the QSAR. MLR display a direct relationship between the dependent variable Y and independent variable X (descriptors). In MLR analysis, the mean of the dependent variable Y relies on X (Descriptors). MLR equationis used to incorporate more than one independent variable (Descriptors) with a single response variable (pMIC).

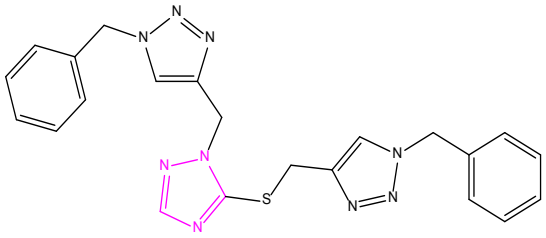
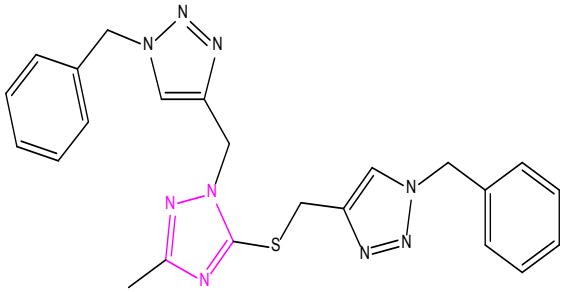
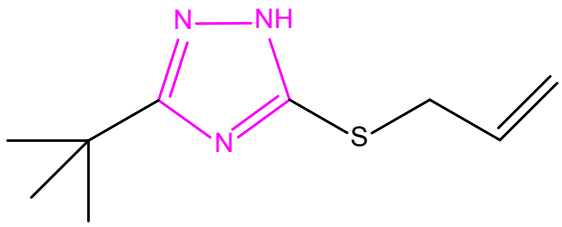
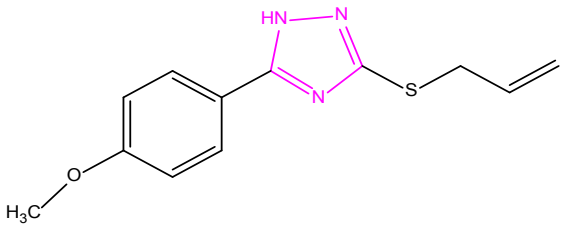
$$Y = k_1x_1 + + k_2x_2 + + k_3x_3 + + C \qquad \text{(Eq.3)}$$

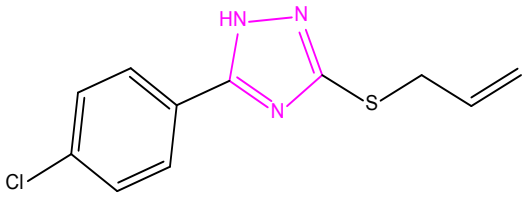
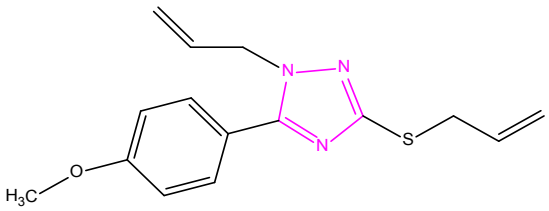
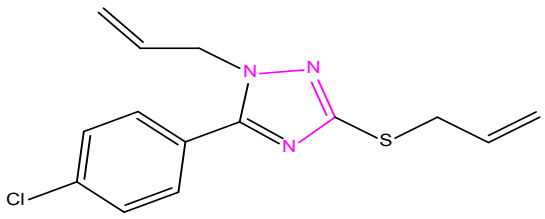
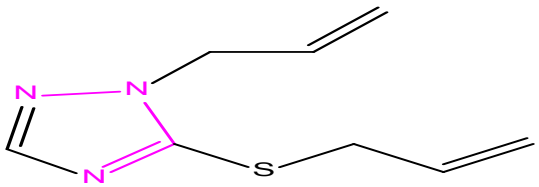
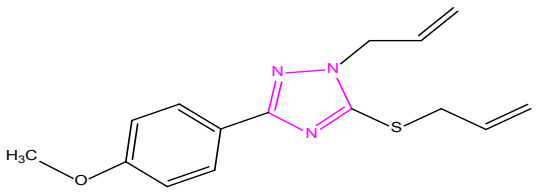
Where Y represent the dependent variable, x represent the independent variables, 'k's are regression coefficients for each '*x*'s and 'C' is a regression intercept.

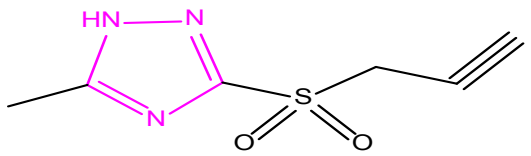
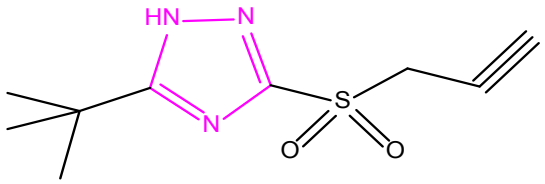
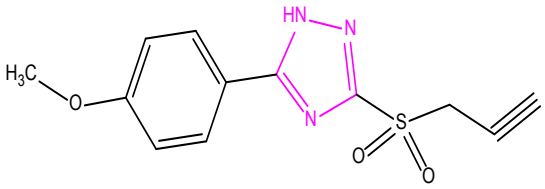
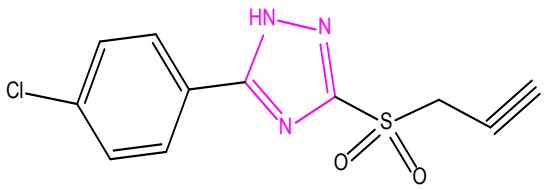
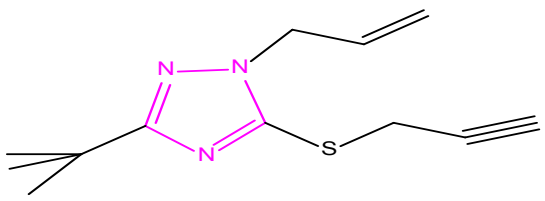
*Generation of QSAR model and validation*

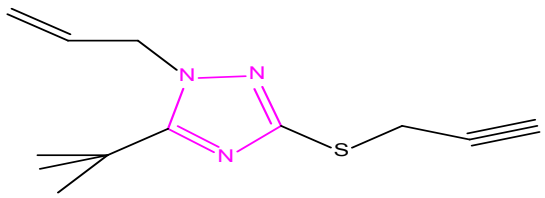
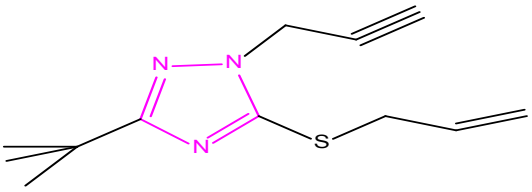
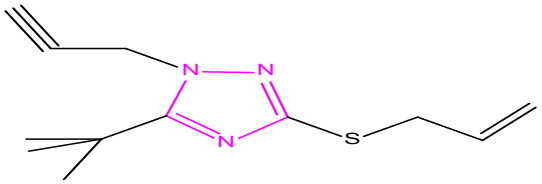
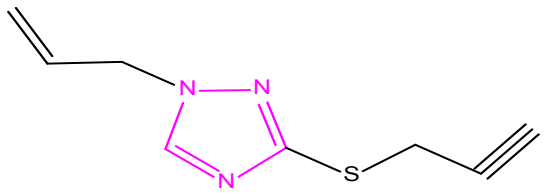
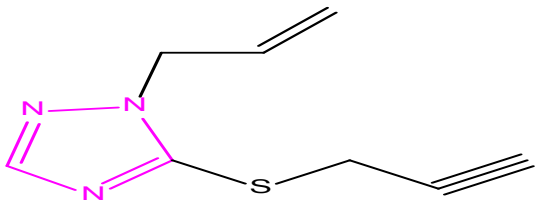
The combinations of the optimum descriptors for the training set were obtained from the descriptor pool using the Genetic Function Approximation technique. Their anti-mycobacterium tuberculosis activities were placed as the last column in their respective spread

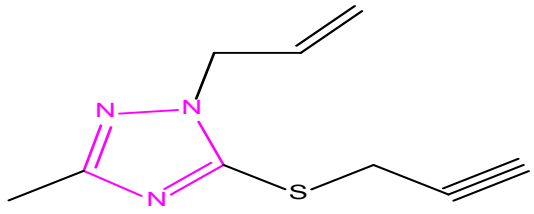
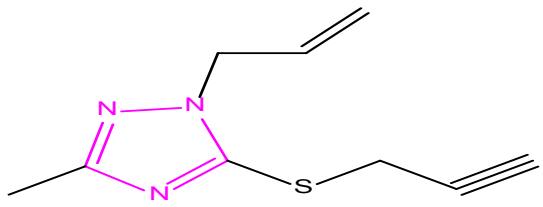
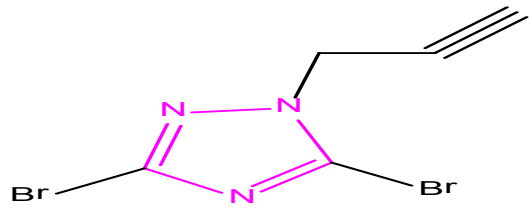
**TABLE 1. Molecular structures of 1,2,4 triazole derivatives and their activities against mycobacterium tuberculosis.**

| S/N | Molecules | Observed Activities (pBA) | Calculated Activities (pBA) | Residual | Leverage |
|---|---|---|---|---|---|
| 1[a] |  1-benzyl-4-(((1-((1-benzyl-1H-1,2,3-triazol-4-yl) methyl)-1H-1,2,4-triazol-5-yl)thio)methyl)-1H-1,2,3-triazole | 6.3456 | 6.37977 | -0.03417 | 0.186966 |
| 2 |  1-benzyl-4-(((1-((1-benzyl-1H-1,2,3-triazol-4-yl) methyl)-3-methyl-1H-1,2,4-triazol-5-yl)thio)methyl)-1H-1,2,3-triazole | 7.4134 | 7.49781 | -0.08441 | 0.267393 |
| 3[a] |  5-(allylthio)-3-(tert-butyl)-1H-1,2,4-triazole | 6.4171 | 6.410504 | 0.006596 | 0.832612 |
| 4 |  3-(allylthio)-5-(4-methoxyphenyl)-1H-1,2,4-triazole | 7.6397 | 7.592776 | 0.046924 | 0.15548 |

| | | | | | |
|---|---|---|---|---|---|
| **5** |  | 8.0899 | 8.37959 | -0.28969 | 0.328411 |
| | 3-(allylthio)-5-(4-chlorophenyl)-1H-1,2,4-triazole | | | | |
| **6** [a] |  | 7.366 | 7.64835 | -0.28235 | 0.085176 |
| | 1-allyl-3-(allylthio)-5-(4-methoxyphenyl)-1H-1,2,4-triazole | | | | |
| **7** |  | 7.0123 | 7.01666 | -0.00436 | 0.343511 |
| | 1-allyl-3-(allylthio)-5-(4-chlorophenyl)-1H-1,2,4-triazole | | | | |
| **8** [a] |  | 6.5267 | 6.289043 | 0.237657 | 0.089973 |
| | 1-allyl-5-(allylthio)-1H-1,2,4-triazole | | | | |
| **9** [a] |  | 7.3233 | 7.60012 | -0.27682 | 0.067538 |
| | 1-allyl-5-(allylthio)-3-(4-methoxyphenyl)-1H-1,2,4-triazole | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **10** |  | 7.3279 | 7.127765 | 0.200135 | 0.101346 |
| | 5-methyl-3-(prop-2-yn-1-ylsulfonyl)-1H-1,2,4-triazole | | | | |
| **11** |  | 6.8568 | 7.04696 | -0.19016 | 0.218861 |
| | 5-(tert-butyl)-3-(prop-2-yn-1-ylsulfonyl)-1H-1,2,4-triazole | | | | |
| **12** |  | 7.3079 | 7.3622 | -0.0543 | 0.079898 |
| | 5-(4-methoxyphenyl)-3-(prop-2-yn-1-ylsulfonyl)-1H-1,2,4-triazole | | | | |
| **13** |  | 7.314 | 7.277527 | 0.036473 | 0.154686 |
| | 5-(4-chlorophenyl)-3-(prop-2-yn-1-ylsulfonyl)-1H-1,2,4-triazole | | | | |
| **14** |  | 8.5854 | 8.6647 | -0.0793 | 0.357197 |
| | 1-allyl-3-(tert-butyl)-5-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **15** |  | 8.0615 | 7.569 | 0.4925 | 0.214607 |
| | 1-allyl-5-(tert-butyl)-3-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | | | | |
| **16** |  | 8.0615 | 7.79949 | 0.26201 | 0.263698 |
| | 5-(allylthio)-3-(tert-butyl)-1-(prop-2-yn-1-yl)-1H-1,2,4-triazole | | | | |
| **17** |  | 6.8494 | 6.60166 | 0.24774 | 0.255295 |
| | 3-(allylthio)-5-(tert-butyl)-1-(prop-2-yn-1-yl)-1H-1,2,4-triazole | | | | |
| **18** [a] |  | 7.9432 | 7.906989 | 0.036211 | 0.409976 |
| | 1-allyl-3-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | | | | |
| **19** |  | 7.4535 | 7.57474 | -0.12124 | 0.25708 |
| | 1-allyl-5-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **20** |  1-allyl-3-methyl-5-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | 7.9759 | 7.966669 | 0.009231 | 0.337231 |
| **21** |  1-allyl-3-methyl-5-(prop-2-yn-1-ylthio)-1H-1,2,4-triazole | 7.9759 | 8.17805 | -0.20215 | 0.249575 |
| **22** [a] |  3,5-dibromo-1-(prop-2-yn-1-yl)-1H-1,2,4-triazole | 7.9294 | 7.437563 | 0.491837 | 0.577201 |

Where superscript **a** represent the test set

sheets in Microsoft Excel 2010 which were later imported into the Material Studio software version 8.0 to generate the QSAR model by employing Multi-linear regression Approach (MLR) and to evaluate the internal validation parameters.

*Evaluation of outlier and influential molecule (applicability domain approach)*

The applicability domain approach was employed to determination of outlier and influential molecule. Any compound outside the applicability domain space of ± 3 is said to be an outlier. To define and describe the applicability domain of the built QSAR models, the leverage *hi* approach was employed and defined as; (10).

$$hi = X_i\left(X^T X\right)^{-1} X_i^T \qquad \text{(Eq.4)}$$

*Xi* is training set matrix of *i*. *X* is the n× k descriptor matrix of the training set compound and $X^T$ is the transpose of the training set (*X*). $X_i^T$ is the transpose matrix $X_i$ used to build the model. The warning leverage $h^*$is the limit values to check for influential molecule. The warning leverage $h^*$ is defined as;

$$h^* = 3\frac{(j\ +\ 1)}{m} \qquad \text{(Eq.5)}$$

Where *j* is the number of descriptors in the build model and *m* is the number of compounds that made up the training set.

*Assessment of Y-randomization*

Y-Randomization test is a confirmatory test to show that the developed QSAR model created is reliable, strong, robust and not gotten by chance. This test was performed on the training set data previously described by [11]. Multi-linear regression (MLR) models were generated

by randomly shuffling the dependent variable (activities data) while keeping the independent variables (descriptors) unaltered. It is expected that the developed QSAR model should have significantly low $R^2$ and $Q^2$ values for numbers of trials in order to ascertain that the developed QSAR models is robust. Y-randomization Coefficient (c $R_{\downarrow}p^{\uparrow}{}_2$) is another important parameter which should be more than 0.5 for passing this test [12]
c

(Eq.6)

Where,

c$R^2_{\bar{p}}$ is Y-randomization Coefficient, R is correlation coefficient for *y*-Randomization and Rr is average 'R' of random models.

*Excellence measure of the model*

The fitting ability, stability, reliability, predictive and robustness of the developed models were evaluated by internal and external validation parameters. The validation parameters were compared with the accepted threshold value for any QSAR model (10) shown in Table 2.

## Results and Discussion

A theoretical approach was to derive a QSAR model for forecasting the activities of inhibitory analogues against *Mycobacterium tuberculosis*. The model generated was built on the basis of the training set while validation of the model was accessed by the test set

The best descriptors that could better predict the activities of the inhibitory compounds were selected with the approach of Genetic Function Algorithm (GFA) while multi-linear Regression (MLR) method was used as modeling techniquein generating the QSAR model. GFA-MLR led to selection of three descriptors and three QSAR models.

The threshold values reported in Table 2 to actually confirm the robustness of the model were all in agreement with validation parameter reported for the model obtained. The observed activities, calculated activities of the inhibitors, the residual values and the leverage value for each compound were reported in Table 1.

The names and symbols of each descriptors selected by GFA approach were presented in Table 3. The combination of the selected descriptors reported in the model indicates the information each descriptor gives on the structure of the compound and the influence each descriptor plays in characterizing the anti-tubercular molecules.

Statistics and correlation matrix of the selected descriptors that were reported in Model 1 were represented in Table 4. The low correlation coefficients that exists between each descriptor in the model imply that the inter-correlation between each descriptor is insignificant. The descriptors were subjected to Variance Inflation Factor (VIF) in order to check for "or thogonally". Meanwhile, the VIF values for each descriptor shown in Table 4 were less than 4 which confirm that the descriptors were statistically significant and orthogonal.

The mean effect (ME) values reported in Table 4 which give vital information on the effect of each descriptor and the degree of contribution in the developed model (13). The signs and the magnitude on the mean effects values indicate direction and strength by which the activities were influenced. Table 4 represents the P-values of the respective descriptors at 95% confidence level. Therefore the null hypothesis that says there is no association between the descriptors and the activities of the molecules is rejected thus; the alternative hypothesis that says there is a relationship between the descriptors used in generating the model and the activities of the compounds at $p < 0.05$ is accepted.

Assessment of Y- Randomization was also conducted and reported in Table 5. The parameter (c$R_{\downarrow}p^{\uparrow}{}_2$) value of 0.733262 greater than 0.5 gives clear evidence that the model generated is valid and not inferred by chance.

The graph of calculated activities plotted against observed activities of the training and test set are presented in Fig. 1 and 2. The correlation coefficient ($R^2$) value of 0.9142 for the training set and ($R^2$) value of 0.7495 for the test set recorded in this work was found to in line with accepted QSAR threshold values reported in Table 3. This affirms the stability, reliability and predictive power of the built model. The plot of residual activities against observed activities shown in Fig. 3 designates that there exists no computational inaccuracy in the derived QSAR model as the range of residuals values fall within an accepted limit of $\pm 2$ on residual activities axis.

The standardized residuals activities plotted against the leverage value known as The Williams plot is shown in Fig. 4. The plotted graph clearly shows none of the molecule is found outside the domain boundary $\pm 3$. Hence, it can be infer that

**TABLE 2. Validation parameters for each model using Multi-linear Regression (MLR).**

| S/NO | Validation Parameters | Formula | Threshold Value (14) | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|---|---|
| **Internal Validation** | | | | | | |
| 1 | Friedman LOF | $\frac{SEE}{\left(1-\frac{C+d\times p}{M}\right)^2}$ | | 0.03562 | 0.03653 | 0.03931 |
| 2 | R-squared | $1-\left[\frac{\sum\left(Y_{exp}-Y_{pred}\right)^2}{\sum\left(Y_{exp}-\bar{Y}_{training}\right)^2}\right]$ | $R^2 > 0.6$ | 0.9142 | 0.8832 | 0.8565 |
| 3 | Adjusted R-squared | $\frac{R^2-P(n-1)}{n-p+1}$ | $R^2_{adj} > 0.6$ | 0.8851 | 0.8488 | 0.8124 |
| 4 | Cross validated R-squared ($Q^2_{cv}$) | $1-\left[\frac{\sum\left(Y_{pred}-Y_{exp}\right)^2}{\sum\left(Y_{exp}-\bar{Y}_{training}\right)^2}\right]$ | $Q^2 > 0.6$ | 0.8324 | 0.8031 | 0.7820 |
| 5 | Significant Regression | | | Yes | Yes | Yes |

| S/NO | Validation Parameters | Formula | Threshold Value (14) | Model 1 | Model 2 | Model 3 |
|---|---|---|---|---|---|---|
| 6 | Critical SOR F-value (95%) | $\frac{\sum\left(Y_{pred}-Y_{exp}\right)^2}{p} \Big/ \frac{\sum\left(Y_{pred}-Y_{exp}\right)^2}{N-p-1}$ | $F_{(test)} > 2.09$ | 3.6832 | 3.6932 | 3.7233 |
| 7 | Replicate points | | | 0 | 0 | 0 |
| 8 | Computed observed error | | | 0 | 0 | 0 |
| 9 | Min expt. error for non-significant LOF (95%) | | | 0.06432 | 0.05632 | 0.07632 |
| **Model Randomization** | | | | | | |
| 10 | Average of the correlation coefficient for randomized data ($\bar{R}_r$) | | $\bar{R} < 0.5$ | 0.3719 | 0.3287 | 0.4321 |
| 11 | Average of determination coefficient for randomized data ($\bar{R}^2_r$) | | $\bar{R}^2_r < 0.5$ | 0.1645 | 0.1356 | 0.2353 |
| 12 | Average of leave one out cross-validated determination coefficient for randomized data ($\bar{Q}^2_r$) | | $\bar{Q}^2_r < 0.5$ | -1.2524 | -1.4321 | -1.3734 |
| 13 | Coefficient for Y-randomization ($cR^2_p$) | $R^2 \times \left(1-\sqrt{\left\|R^2-\bar{R}^2_r\right\|}\right)$ | $^cR^2_p > 0.6$ | 0.7332 | 0.6432 | 0.6284 |
| **External validation** | | | | | | |
| 14 | $R^2_{test}$ | $1-\frac{\sum\left(Y_{ext}-\bar{Y}_{ext}\right)^2}{\sum\left(Y_{ext}-\bar{Y}\right)^2}$ | $R^2_{pred} > 0.6$ | 0.7494<br>0.7112 | | 0.6532 |

Key: SSE is the sum of squares of errors, C is the number of terms in the model, d is a user-defined smoothing parameter, p is the total number of descriptors contained in the model and M is the number of data in the training set. Yexp, Ypred and ¯Y_training are the .experimental activity, the predicted activity and the mean experimental activity of the samples in the training set respectively

Model 1

pBA = - 0.37456543543 (AATS5e) + 2.087643542 (minHCsatu) + 0.293436327 (RDF90s) + 3.02312046

Model 2

pBA = - 0.3285458991* AATS7s + 0.024550934 (TDB9e) - 0.117941052 (RDF110i) - 9.645640119

Model 3

pBA = - 0.335632223* AATS7s + 0.021034761 (TDB9e) - 0.129647108 * RDF30i + 8.992978173

**TABLE 3. Descriptors used in the QSAR optimization model.**

| S/NO | Descriptors symbols | Name of descriptor (s) | Class |
|---|---|---|---|
| **1** | AATS5e | Average Broto-Moreau autocorrelation - lag 5 / weighted by I-state auto-correlation | 2D |
| **2** | minHCsatu | Minimum atom-type H E-State: H on C sp3 bonded to unsaturated C | 2D |
| **3** | RDF90s | Radial distribution function - 110 / weighted by relative I-state | 3D |

**TABLE 4. Statistical analysis and Pearson's correlation .**

| **Inter- correlation** | | | **Statistics** | | |
|---|---|---|---|---|---|
| **AATS7s** | TDB9e | RDF90i | P- Value (Confidence interval) | VIF | Mean Effect (ME) |
| **AATS7s** | 1 | | 0.00014 | 2.4313 | -0.4322 |
| **TDB9e** | -0.18343 | 1 | 0.00073 | 2.2322 | 0.5356 |
| **RDF90i** | 0.43432 | -0.23298 | 0.00051 | 1.0132 | 0.2084 |

**TABLE 5. Y- Randomization Parameters test for model 1.**

| **Model** | **R** | **R^2** | **Q^2** |
|---|---|---|---|
| **Original** | 0.85791 | 0.736009 | 0.361481 |
| **Random 1** | 0.263469 | 0.069416 | -0.42957 |
| **Random 2** | 0.634931 | 0.403137 | -3.21615 |
| **Random 3** | 0.44027 | 0.193838 | -1.71176 |
| **Random 4** | 0.45403 | 0.206144 | -0.7079 |
| **Random 5** | 0.642442 | 0.412732 | -4.71577 |
| **Random 6** | 0.116309 | 0.013528 | -0.3569 |
| **Random 7** | 0.24943 | 0.062215 | -0.2046 |
| **Random 8** | 0.296007 | 0.08762 | -0.42455 |
| **Random 9** | 0.270977 | 0.073429 | -0.37515 |
| **Random 10** | 0.351074 | 0.123253 | -0.38131 |

| **Random Models Parameters** | |
|---|---|
| **Average r :** | 0.371894 |
| **Average r^2 :** | 0.164531 |
| **Average Q^2 :** | -1.25236 |
| **cRp^2 :** | 0.733262 |

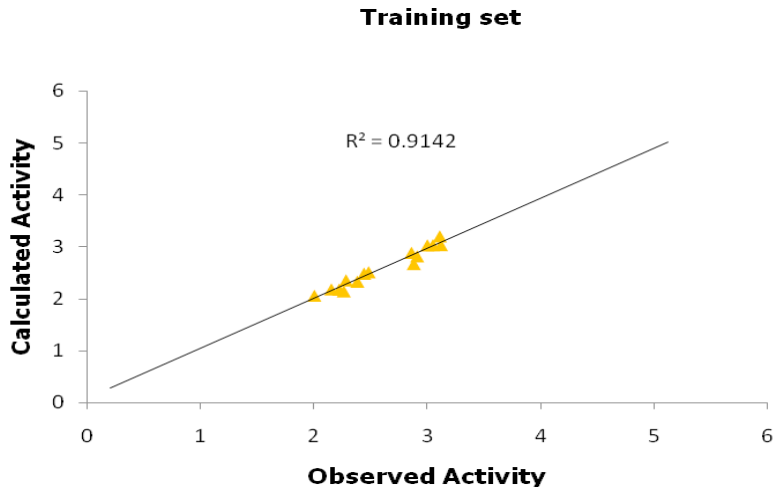



**Fig.1. Calculated activities vs Observed activities of the training set.**

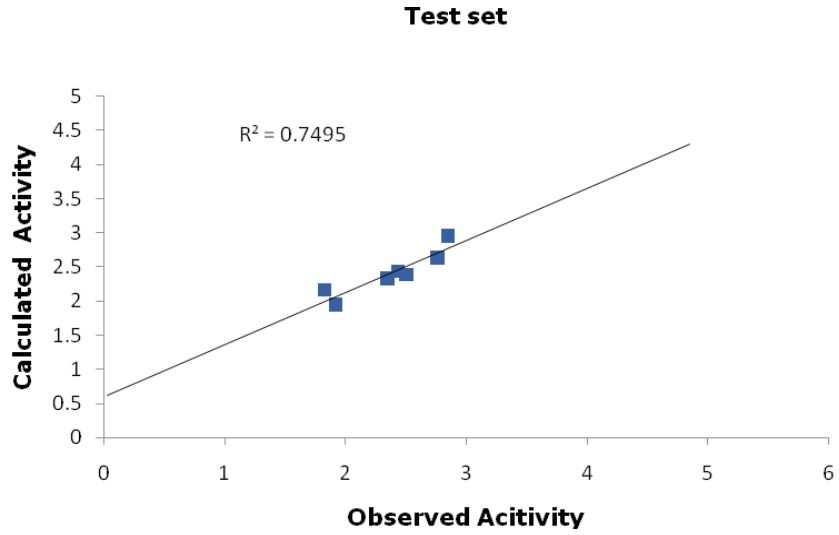



**Fig. 2. Calculated activities vs Observed activities of test set**

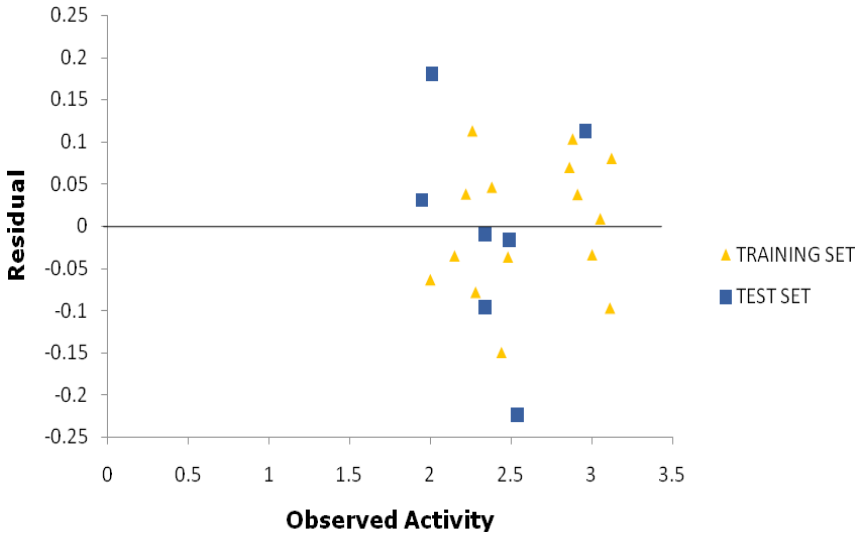



**Fig. 3. Residual activities vs Observed activities.**

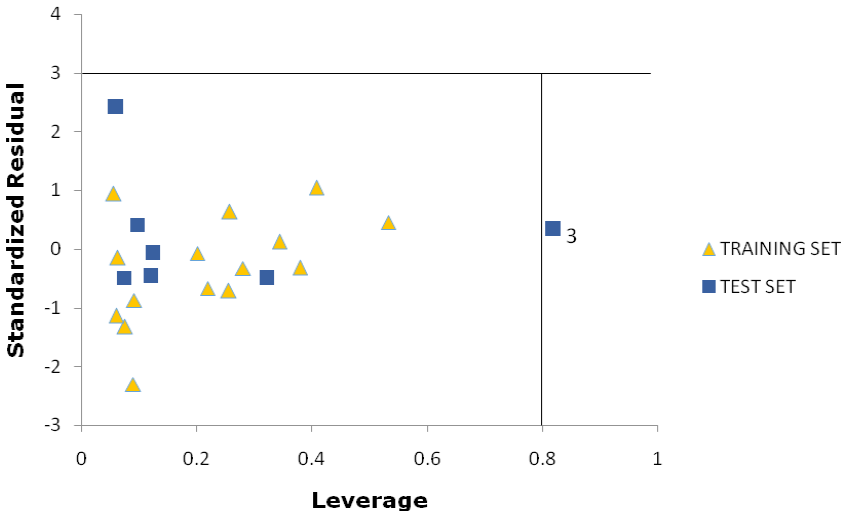



**Fig. 4. Standardized residual activities vs leverages values.**

no outlier is observed in the data set. However, the calculated warning leverage (h* = 0.80) is observed to be less than compound (number 3). Therefore the compound is an influential molecule.

## Conclusion

In this research, a theoretical approach was employed in this study to selected molecular descriptors to derivea model that could be used to correlate the structure of 1,2,4-triazole derivatives against *tuberculosis* and their respective biological activities. The model was found to be influenced with descriptor (AATS5e, min HCsatu and RDF90s). Meanwhile, the model derived was subjected to internal and external validation test to affirm that the built QSAR model is significant, robust, and reliable. The built QSAR model will bea vital tool for pharmaceutical as well as medicinal chemists to design and synthesis novel anti-tubercular drugs with better activities against *M. tuberculosis*.

## References

1. Organization WH, others. Tuberculosis Fact Sheet (No. 104) 2000. Site Accessed www.who Intmediacentrefactsheetswho104enindex Html. (2013)

2. Hansch C., Kurup A., Garg R., Gao H., Chem-bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms. *Chemical Review;* **101** (3), 619–672 (2001)

3. Adeniji S.E., Uba S. and Uzairu A., Theoretical modeling and molecular docking simulation for investigating and evaluating some active compounds as potent anti-tubercular agents against MTB CYP121 rerceptor. *Future Journal of Pharmaceutical Sciences;* **4** (2), 284-295 (2018)

4. Ogadimma, A.I. and Adamu, U., Analysis of Selected Chalcone Derivatives as Mycobacterium tuberculosis Inhibitors. *Scientific Research,* **3** (3), 1–13 (2016)

5. Adeniji S.E., Uba S. and Uzairu A., QSAR Modeling and Molecular Docking Analysis of Some Active Compounds against Mycobacterium tuberculosis Receptor (Mtb CYP121). *Hindawi Journal of Pathogen;* Article ID 1018694, (2018).

6. Adeniji SE, Uba S. and Uzairu A., 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** (3), 479–492 (2018).

7. Eric G.M., Uzairu A., Mamza P.A., A Quantitative Structure-activity Relationship (QSAR) Study of the Anti-tuberculosis Activity of Some Quinolones **10** (1), 2320-0227 (2016).

8. Sarkar D., Deshpande S.R, Maybhate S.P., Likhite A.P., Sarkar S., Khan A., 1, 2, 4-triazole derivatives and their anti-microbial activity. *Google Patents;* (2016).

9. Singh P., Quantitative Structure-Activity Relationship Study of Substituted-[1, 2, 4] Oxadiazoles as S1P1 Agonists. *Journal of Current Chemistry and Pharmaceutical Science,* **3** (1), 345 – 457 (2013).

10. Veerasamy R., Rajak H., Jain A., Sivadasan S., Varghese C.P., Agrawal R.K., Validation of QSAR models-strategies and importance. *International Journal of Drug Discovery;* **3** (2), 511–519 (2011).

11. Tropsha A., Gramatica P., Gombar V.K., The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *Molecular Information;* **22** (1), 69–77 (2003).

12. Adeniji S.E., Uba S., Uzairu A., Theoretical modeling for predicting the activities of some active compounds as potent inhibitors against Mycobacterium tuberculosis using GFA-MLR approach. *Journal King Saud University of Science.* Article in press (2018).

13. Adeniji S.E., Arthur D.E., Oluwaseye A., Computational modeling of 4-Phenoxynicotinamide and 4- Phenoxypyrimidine -5- carboxamide derivatives as potent anti-diabetic agent against TGR5 receptor. *Journal King Saud University of Science.* Article in press (2018).

14. Adeniji S.E., Uba S., Uzairu A., In silico study for investigating and predicting the activities of 1, 2, 4-triazole derivaties as potent anti-tubercular agents. *Journal of Engineering and Exact Science,* **4** (2), 0246–0254 (2018).