Egypt. J. Chem. **Vol. 63,** No. 12 pp. 4703 - 4717 (2020)

Egyptian Journal of Chemistry

http://ejchem.journals.ekb.eg/

Density Functional Theory (DFT) Investigation and Molecular Docking Simulation of 1, 2, 4-triazole Derivatives as Potent Inhibitors Against a Receptor (DNA Gyrase)

Shola Elijah Adeniji

Department of Chemistry, Ahmadu Bello University, Zaria-Nigeria

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.* The developed model in this study was achieved using Density Functional Theory (DFT) optimization approach validated. Molecular docking studies was as well carried in order to show the interactions and binding modes between the ligand and the receptor (DNA gyrase). The lead compound (compound 8) with higher anti-tubercular activity was observed with prominent binding affinity of -12.3 kcal/mol. Therefore, compound 8 could serve as a template structure to designed compounds with more efficient activities. The outcome of this research is recommended for pharmaceutical and medicinal chemists to design and synthesis more potent compounds with prominent anti-tubercular activities using the model designed in this study.

Keywords: QSAR; Docking; Ligan; receptor

1. Introduction

Multi-drug resistance strain *Mycobacterium tuberculosis* (TB) has pose a challenge toward the treatment of tuberculosis in the global community. 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 patient suffering from tuberculosis [2]. The resistances of the *M. tuberculosis* toward the current drugs led to development of new approach that is fast and precise which could able to predict the biological activity for the new compounds against *M. tuberculosis* [3].

DNA gyrase is a type II topoisomerase found in all bacteria. It plays an important role in DNA replication using ATP energy. This makes it a good target for antibacterial chemotherapy. DNA gyrase generates negative supercoils for the entire bacterial chromosome. This relaxes the positive supercoils that transcription generates ahead of the translocating RNA polymerase which results in a condensed chromosome for proper partitioning during cell division [4]. It is a tetramer composed of two A subunits, where the DNA binding domain is located, and two B subunits with ATPase activity, which catalyze reactions that transiently cleave two DNA strands by a process dependent on ATP hydrolysis. The GyrA and GyrB break and reunite DNA which aids the DNA replication. With this function, either the DNA domain (GyrA) or ATP binding cavities (GyrB) can be blocked by inhibitors for the termination of the DNA replication. [5].

*Corresponding author e-mail: *Author email: shola4343@gmail.com, ORCID ID: http://orcid.org/0000-0002-7750-8174, Tel: +2347060511720*
Receive Date: 15 December 2019, Revise Date: 11 February 2020, Accept Date: 24 June 2020
DOI: 10.21608/EJCHEM.2020.21024.2253

It has been established that heterocyclic compounds play an important role in designing new class of structural entities for medicinal applications [6]. Among pharmacologically important heterocyclic compounds, triazole and its derivatives are attracted considerable attention in fields, such as medicinal and agrochemical research as well as in the material sciences due to their unique structure and properties [6]. Triazole, also known as pyrrodiazole is one of the classes of organic heterocyclic compounds containing a five membered diunsaturated ring structure composed of three nitrogen atoms and two carbon atoms at non-adjacent positions. In the recent past, triazole nucleus has gathered an immense attention among chemists as well as biologists as it is one of the key building elements due to their chemotherapeutical values [7]. Triazole and its derivatives are important class of bioactive molecules in the field of drugs and pharmaceuticals. They exhibit significant wide range of pharmacological activities such as anti-microbial [8], analgesic [9], anti-neoplastic [10], anti-malarial [11]. Among other heterocyclic derivatives, triazole compounds were reported as most promising candidates towards anti-TB activity [12].

Meanwhile, a theoretical approach; quantitative structure activity relationships (QSARs) is one of the most widely used computational method which helps in drug designing and prediction of drugs activities [3]. 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 [3] have established relationship between some anti-tubercular inhibitor's like quinolone, chalcone, pyrrole and 7-methyijuglone 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. Therefore, this study aimed to establish a valid QSAR model that could correlate the structures of 1, 2, 4-triazole derivatives, molecular docking simulation and to design new potent compounds with better anti-tubercular activities against *Mycobacterium tuberculosis*. The aim of this research was to carry out Density Functional Theory (DFT) investigation and molecular docking simulation of 1, 2, 4-triazole derivatives as potent inhibitors against a receptor (DNA gyrase).

2. Materials and method

2.1 Data collection

Thirty (30) 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 [13]. The general structure of 1, 2, 4-triazole derivatives and experimental activities of these compounds were presented in Figure 1 and Table 1 respectively.

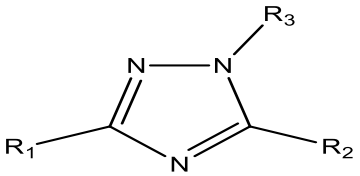


Fig.1. General structure of 1, 2, 4-triazole derivatives

2.2 Molecular 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 [13].

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 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 Equation 1 [14].

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

Where dmax and dmin 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 [14].

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. 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 model was built by adopting the Genetic Function Approximation (GFA) technique incorporated in the Material Studio software version 8.0 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 Equation 2 was used define applicability domain space ±3 for outlier molecule [3].

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

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.

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

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 [3].

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 [2, 15, 16]. Random shuffling of the data was performed on the training set following the principle laid by [16, 17]. 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 $(\mathrm{c}R_p^2)$ shown in Equation 4 must be ≥ 0.5 in order to establish the robustness of the model.

$$\mathrm{c}R_p^2 = R \times [R^2 - (R_r)^2]^2 \qquad \text{(Eq.4)}$$

Where, $\mathrm{c}R_p^2$ 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, 15–18] in order to affirm the reliability, fitting, stability, robustness and predictability of the developed models.

2.10 Docking studies

2.10.1 Preparation of the receptor (DNA gyrase)

The DNA gyrase receptor was gotten from protein data bank with PDB code 31FZ in form crystal structure [19, 20]. Foreign bound substances such as ligands, solvent molecules and cofactors associated with the receptor were removed with the aid of Discovery Studio Visualizer software. The prepared receptor was saved in an input format (PDB) which is recognized by Discovery Studio Visualizer and Pyrx software. In the Pyrx software, the prepared receptor was transported in order to become a macro molecule [2, 18] as shown in Figure 2.

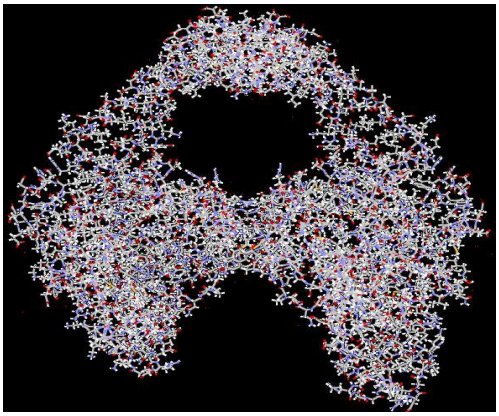


Fig. 2. (A) Crystal structure of DNA gyrase.

2.10.2 Preparation of the ligands

All the studied compounds (1, 2, 4-triazole derivatives.) were optimized in order to have a stable conformer at a minima energy with the aid of Spartan 14 software at Density Functional Theory (DFT) level. After optimization, the ligand was saved as a PDB format, imported into the Pyrx software in order to become micro molecules (ligands) [2, 18].

2.10.3 Receptor-ligand complex docking

Molecular docking of the receptor with the ligands were carried out utilizing the PyRx virtual screening software. The software comprises the combination of several software's such as AutoDock Vina 4.2, Open Babel and Mayavi, etc. In order to perform receptor-ligand docking, the ligands and the receptor (DNA gyrase) were converted from pdb format to pdbqt (protein data bank, partial charge and atom type) format. The conversion of pdb format to pdbqt format (Vina input format) was done by launching the PyRx virtual screening software in order compute the Binding Affinity (kcal/mol). The more the negative the binding affinity, the better the orientation of the ligand in the binding site of DNA gyrase. The docked results were visualized, analyzed and compiled with the aid of Discovery Studio Visualizer software [2, 18].

3. Results and discussion

A theoretical approach was employed to derive a QSAR model for predicting the activities of 1, 2, 4 Triazole analogues against *Mycobacterium tuberculosis*. Kennard-Stone algorithm approach employed [21, 22] in this research was able to divide the studied compounds which comprises of 30 compounds into a training set of 21 compounds while the remaining 9 compound serve as the test set. The model generated was built on the basis of the training set while validation of the model was accessed by the test set

The combination of the utmost 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 technique in generating the QSAR model. GFA-MLR led to selection of four descriptors and four QSAR models as shown below. However, based on the internal and external validation parameters reported in Table 2 for a valid QSAR model, model 1 with robust and optimum predictive ability was selected as best model. These parameters were in agreement with the threshold value reported in Table 3 which actually confirmed the robustness and stability of the model.

Model 1

pBA = -1.935644981 (AATS7s) + 2.133204345 (nHBint3) - 1.484532032 (minHCsatu) + 0.964401389 (Vi) + 5.463456786

Model 2

pBA = 2.133597608 (nHBint3) - 1.313467537 (nHBint7) + 6.924567876 (E1i) + 0.972490084 (Vi) + 2.469905087

The QSAR model generated in this research was compared with the models obtained in the literature [3] as shown below;

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

N_{train} = *35*, R^2 = 0.9142, R_{adj} = 0.8851, Q^2_{cv} = 0.8324 and the external validation for the test set was found to be R^2pred = 0.7494 [3]

From the above models the validation parameters reported in this work and those reported in the literature were all in agreement with parameters presented in Table 3 which actually confirmed the robustness of the model generated.

The observed activities, predicted activities of the inhibitors, and the residual values for each compound were reported in Table 4. The low residual values between observed activities and predicted activities

indicate that the model generated has a high predictive ability.

The names and symbols of each descriptors selected by GFA approach were presented in Table 5. The combination of the selected descriptors (2D and 3D) reported in model 1 indicates that these types of descriptors are able characterize and give better information on the structure of the anti-tubercular molecules.

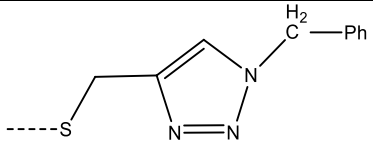
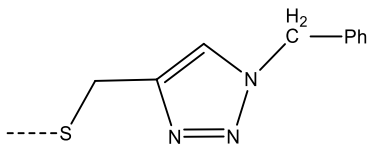
The Person correlation coefficients calculated for the descriptors in the model was reported in Table 6. The low correlation coefficients that exist between each descriptor in the model imply that there exists no significant inter-correlation between each descriptor [21, 22]. Statistical parameters calculated for the selected descriptors reported in Model 1 were repented in Table 6. The descriptors were subjected to Variance Inflation Factor (VIF) in order to check for orthogonality. Meanwhile, the VIF values for each descriptor shown in Table 6 were less than 4 which confirm that the descriptors were statistically significant and orthogonal. The mean effect (ME) reported in Table 6 give a vital information on the effect of each descriptor and the degree of contribution in the developed model. The signs and the magnitude on the mean effects values indicate direction and strength with which each descriptor is influencing the activity of each compound. Table 6 represents the P-values of each of the descriptors in the model at 95% confidence level. Therefore the null hypothesis that says there is no association between the descriptors influencing the model 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. More alos the value of Y-randomization Coefficient (cR_p^2) computed to be 0.6379 which is greater than 0.5 shows that the built model is robust and not obtained by chance.

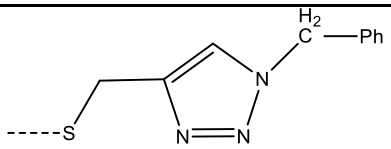
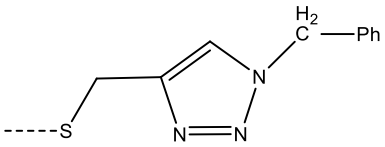
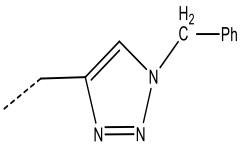
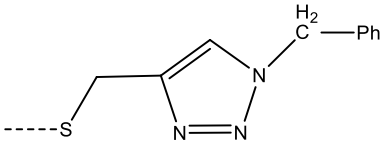
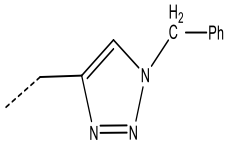
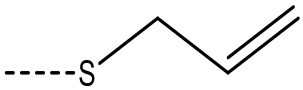
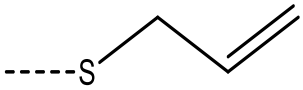
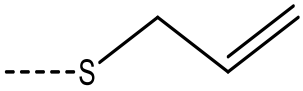
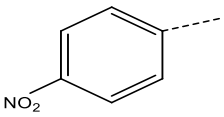
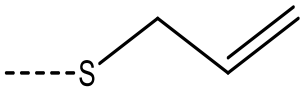
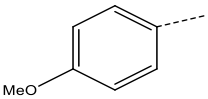
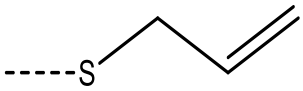
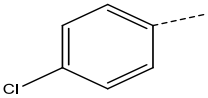
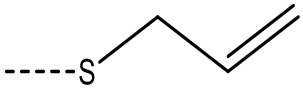
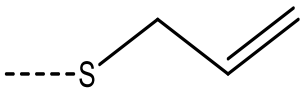
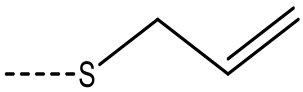
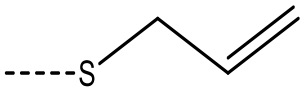
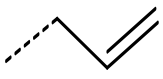
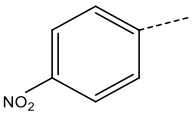
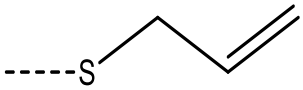
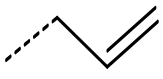
The graph of predicted activities plotted against observed activities of the training and test set are presented in Figure 3 and 4. The correlation coefficient (R^2) value of 0.8718 for the training set and (R^2) value of 0.7330 for the test set recorded in this work was found to in line with accepted QSAR threshold values reported in Table 3. This affirms the stability, reliability and predictive power of the built model. The plot of residual activity against observed activities shown in Figure 5 designates that there exist no computational inaccuracy in the derived QSAR model as the range of residuals values fall within an accepted limit of ±2 on residual activity axis.

The standardized residuals activities plotted against the leverage value known as The Williams plot is shown in Figure 6. The plotted graph clearly shows that all the compounds falls within limit boundary ±3 of standardized cross-validated residual. Hence, it can be infer that no outlier is observed in the data set. Moreover, all the compounds except compounds 3 and 6 were found to be lower than the warning leverage ($h^* = 0.71$). Therefore compounds 3 and 6 are said to be an influential molecules.

Molecular docking was carried out between the targets (DNA gyrase) of *M. tuberculosis* and 1, 2, 4-triazole derivatives. Nine (9) inhibitor ligands (compounds 1, 2, 8, 11, 16, 19, 22, 23 and 24) with better activity were selected and docked with the DNA gyrase in order to elucidate the interaction and the binding mode. The binding affinity values for these ligands ranges from (-6.2 to -12.3 kcal/mol) as reported in Table 8. The ligand (compound 8) with best activity was selected for visualization purpose utilizing Discovery Studio Visualizer as shown in Figure 7 and 8 below. Ligand 8 formed four hydrogen.

Table 1: Chemical structure and experimental activity of the inhibitory compounds

Molecule	R_1	R_2	R_3	Activity (pBA)
1	H		H	8.0250
2	methyl		H	8.0345

Molecule	R_1	R_2	R_3	Activity (pBA)
3	isobutyl		H	8.7064
4	Methyl			5.7441
5	isobutyl			5.9258
6	H		H	6.1667
7	methyl		H	5.8765
8	isobutyl		H	6.4171
9			H	5.9413
10			H	6.6397
11			H	8.0899
12	H			6.5267
13	Methyl			5.7405
14	isobutyl			5.6533
15				6.1923

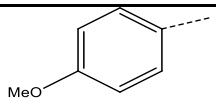
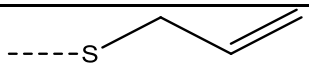
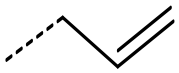
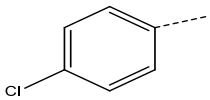
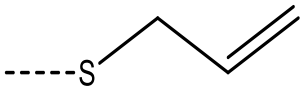
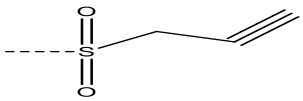
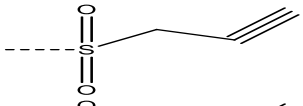
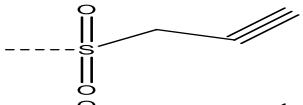
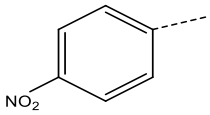
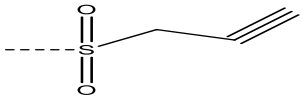
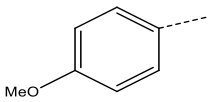
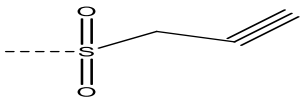
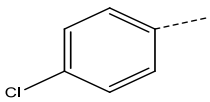
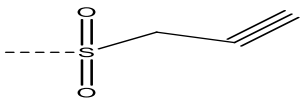
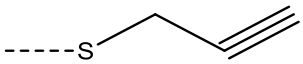
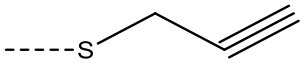
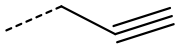
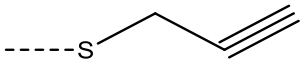
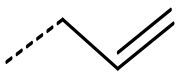
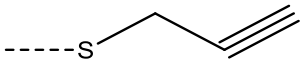
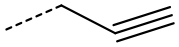
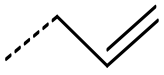
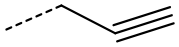
Molecule	R_1	R_2	R_3	Activity (pBA)
16				7.3233
17				6.0097
18	H		H	6.0928
19	methyl		H	7.3279
20	isobutyl		H	6.8568
21			H	6.2234
22			H	7.0079
23			H	7.314
24	isobutyl			7.0854
25	isobutyl			7.2615
26	H			5.2346
27	Methyl			6.4218
28	Br	Br		5.1016
29	Br	Br		6.1213
30	Br	Br		5.4406

Table 2: QSAR internal and external validation parameters for each model using Genetic Function Approximation

S/NO		Model 1	Model 2
1	Friedman LOF	0.1412	0.1544
2	R-squared	0.8718	0.75183
3	Adjusted R-squared	0.8019	0.7318

4	Cross validated (R-squared (Q^2_{cv})	0.7716	0.7216
5	Significant Regression	Yes	Yes
6	Significance of regression F-value	78.3421	75.3487
7	Critical SOR F-value (95%)	2.8281	2.8514
12	R^2 test	0.7330	0.6517

Table 3: Recommended values value for the validation parameters for a given QSAR model

Validation Parameter	Formula	Threshold	comment	Reference
	Internal validation			
$\mathbf{R^2}$	$\frac{\left[\sum\left\{(Y-\bar{Y})\times\left(\hat{Y}-\bar{\hat{Y}}\right)\right\}\right]^2}{\sum(Y-\hat{Y})^2\times\sum\left(\hat{Y}-\bar{\hat{Y}}\right)^2}$	$R^2 > 0.6$	passed	[2,16, 17]
$\mathbf{R^2_{adj}}$	$\frac{(N-1)\times R^2 - p}{N-1-p}$	$R^2_{adj} > 0.6$	passed	[2,16, 17]
$\mathbf{Q^2}$	$1-\frac{\sum(Y-\hat{Y}_{loo})^2}{\sum(Y-\bar{Y})^2}$	$Q^2 > 0.6$	Passed	[2,16, 17]
$\mathbf{F_{(4,15)}}$	$\frac{\sum(Y-\bar{Y})^2}{p}\Big/\frac{\sum(Y-\hat{Y})^2}{N-p-1}$	$F_{(test)} > 2.09$	Passed	[2,16]
	Random model			
$\bar{R}_r$	an average of the correlation coefficient for randomized data	$\bar{R} < 0.5$	passed	[16,17]
$\bar{R}^2_r$	an average of determination coefficient for randomized data	$\bar{R}^2_r < 0.5$	Passed	[16,17]
$\bar{Q}^2_r$	an average of leave one out cross-validated determination coefficient for randomized data	$\bar{Q}^2_r < 0.5$	Passed	[16,17]
$^cR^2_p$	$R^2\times\left(1-\sqrt{\lvert R^2-\bar{R}^2_r\rvert}\right)$	$^cR^2_p > 0.6$	Passed	[2, 15]
	External validation			
$\mathbf{R^2_{Pred}}$	$1-\frac{\sum(Y_{ext}-\hat{Y}_{ext})^2}{\sum(Y_{ext}-\bar{Y})^2}$	$R^2_{pred} > 0.6$	Passed	[2,15–18]

Table 4: Observed, Predicted and Residual values for the inhibitory compounds

Molecule	Observed Activity	Predicted Activity	Residual	Leverage
1[a]	8.025	8.0710	-0.0460	0.3935
2[a]	8.0345	8.3953	-0.3608	0.1349
3[a]	6.497	6.0502	-1.5532	0.7265
4	5.7441	5.9529	-0.2088	0.14
5	5.9258	5.7392	0.1866	0.1350
6[a]	6.1667	6.5066	-0.3399	0.8313
7	5.8765	5.9417	-0.0652	0.2689
8[a]	8.7332	7.7333	1.6999	0.1409
9	5.9413	5.8350	0.1063	0.1299

Molecule	Observed Activity	Predicted Activity	Residual	Leverage
10	6.6397	6.8240	-0.1843	0.1255
11	8.0899	7.8412	0.2487	0.2047
12	6.5267	6.3712	0.1555	0.0923
13	5.7405	5.6656	0.0749	0.1098
14	5.6533	5.9570	-0.3037	0.4072
15	6.1923	6.3825	-0.1902	0.2646
16 [a]	7.3233	7.1734	0.1499	0.2508
17	6.0097	5.9077	0.1020	0.4827
18	6.0928	6.2428	-0.1500	0.3137
19	7.3279	7.2634	0.0645	0.3053
20	6.8568	6.9664	-0.1096	0.1123
21	6.2234	6.5734	-0.3500	0.3298
22	7.0079	7.4413	-0.4334	0.4612
23	7.314	7.3090	0.0050	0.3071
24	7.0854	7.6564	-0.5710	0.1059
25 [a]	7.2615	7.0523	0.2092	0.5672
26	5.2346	5.7764	-0.5418	0.2996
27 [a]	6.4218	6.2843	0.1375	0.3914
28	5.1016	5.8498	-0.7482	0.3071
29 [a]	6.1213	6.3744	-0.2531	0.7359
30	5.4406	5.8245	-0.3839	0.5672

Where superscript **a** represent the test set

Table 5: List of some descriptors used in the QSAR optimization model

S/NO	Descriptors symbols	Name of descriptor(s)	Class
1	AATS7s	Average Broto-Moreau autocorrelation - lag 7 / weighted by I-state	2D
2	nHBint7	Count of E-State descriptors of strength for potential Hydrogen Bonds of path length 7	2D
3	minHCsatu	Minimum atom-type H E-State: H on C sp3 bonded to unsaturated C	2D
4	Vi	V total size index / weighted by relative first ionization potential	3D

Table 6: Pearson's correlation and statistics for descriptor used in the QSAR model

	Inter- correlation			Statistics		
AATS7s	nHBint3	minHCsatu	Vi	P- Value (Confidence interval)	VIF	Mean Effect (ME)
AATS7s	1			3.34E-05	1.0083	-0.6324
nHBint3	-0.543	1		0.00003	1.0945	0.8130
minHCsatu	0.0519	-0.2913	1	2.8E-04	2.612	-0.7209
Vi	0.7041	0.0934	0.283317	0.00008	1.8103	0.7204

Table 7: Y- Randomization Parameters test for Model 1

Model	R	R^2	Q^2
Original	0.762302	0.726026	0.895386
Random 1	0.387394	0.150074	-0.28301
Random 2	0.534646	0.285847	-0.15518
Random 3	0.357333	0.127687	-0.43633
Random 4	0.509588	0.25968	-0.08884
Random 5	0.231807	0.053735	-0.60188
Random 6	0.140884	0.019848	-0.61556
Random 7	0.513288	0.263465	-0.11043
Random 8	0.548099	0.300412	-0.062
Random 9	0.36673	0.134491	-0.25601
Random 10	0.505524	0.255554	-0.12398
Random Models Parameters			
Average r :	0.409529		
Average r^2 :	0.185079		
Average Q^2 :	-0.27332		
cRp^2 :	0.637983		

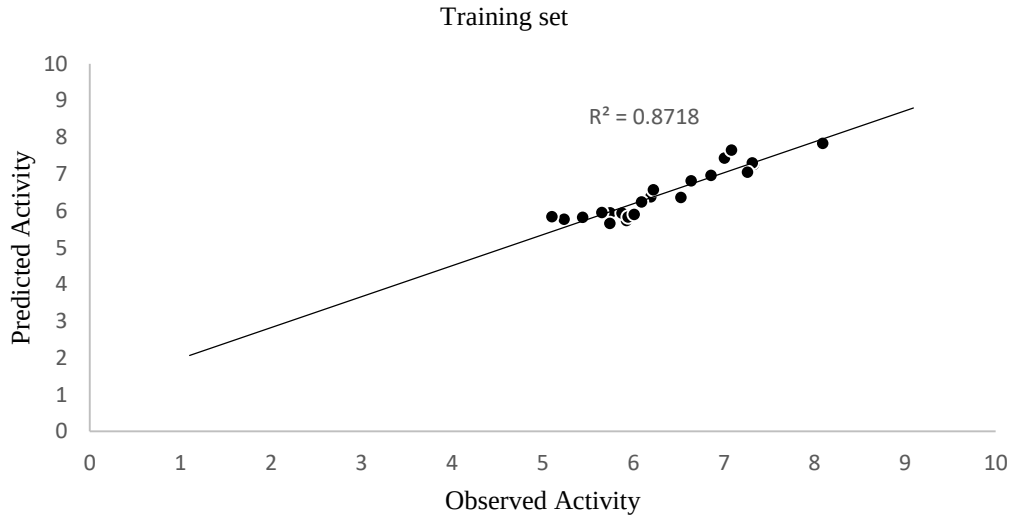


Fig. 3. Plot of predicted activity against observed activity of training set

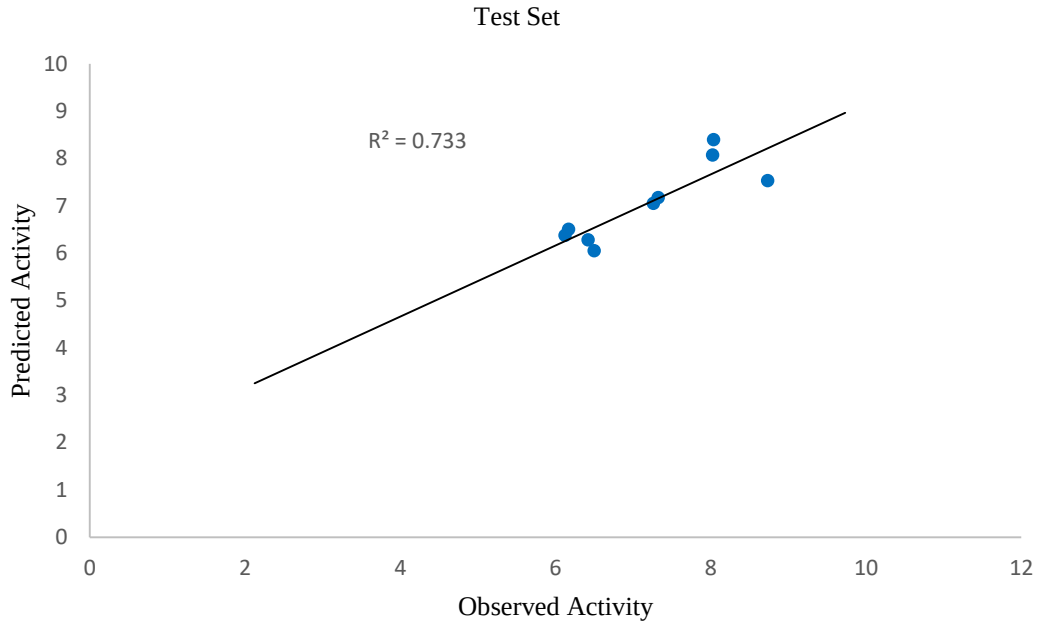


Fig. 4. Plot of predicted activity against observed activity of test set

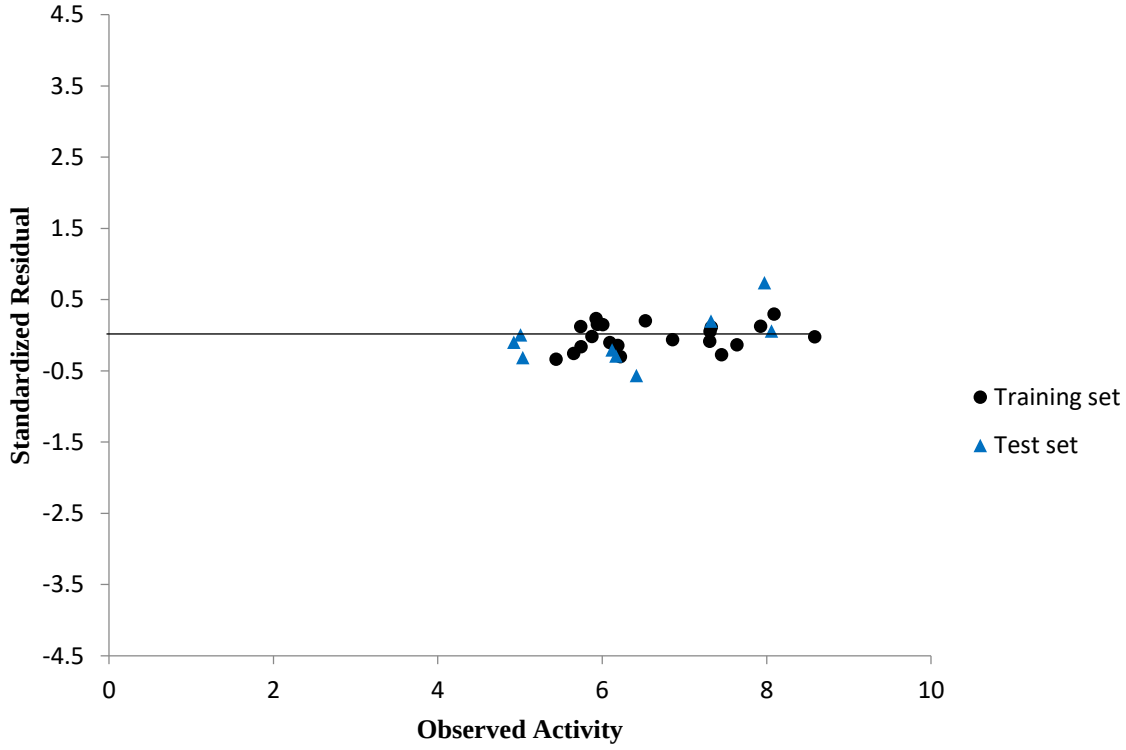


Fig.5. Plot of standardized residual versus observed activity

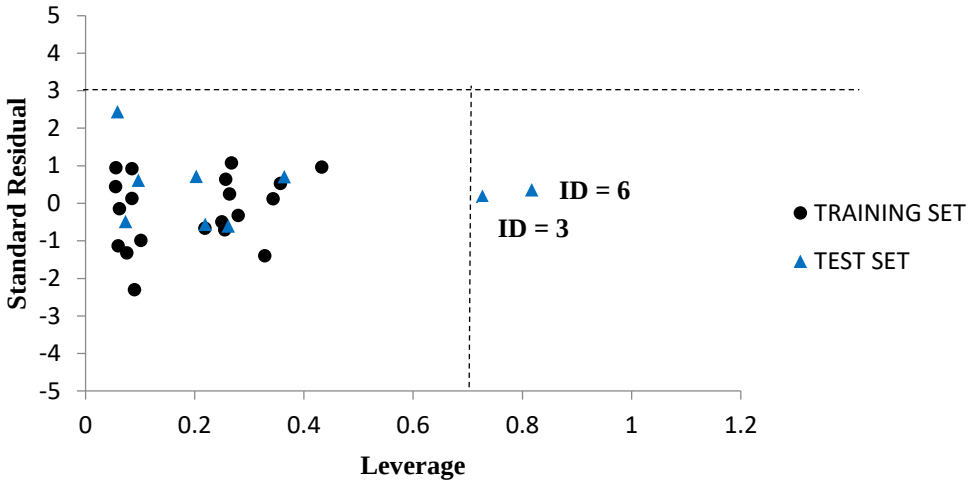


Fig. 6. Plot of the standardized residuals versus the leverage value

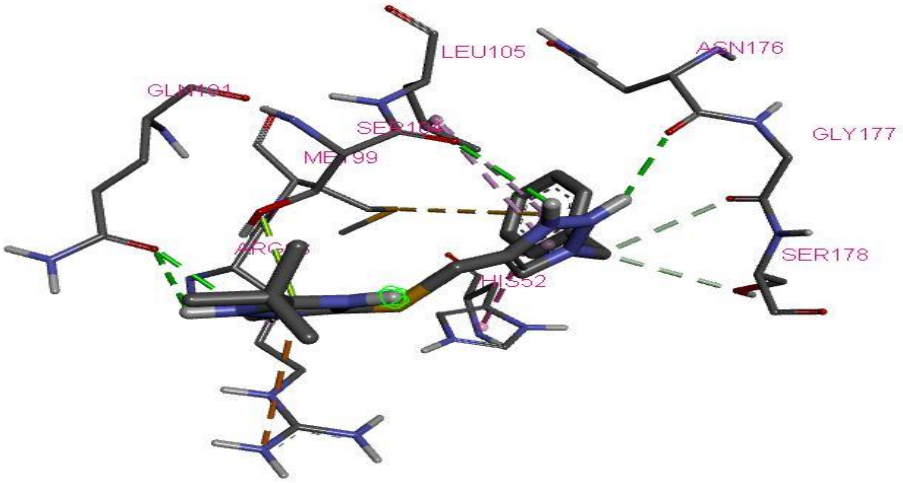


Fig. 7. 3D interactions between DNA gyrase and Ligand 8.

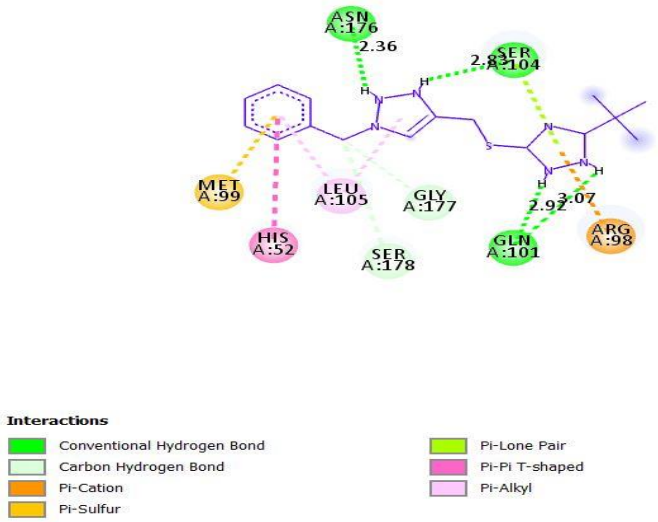


Fig.8. 2D interactions between DNA gyrase and Ligand 8.

Table 7: Binding Affinity, Hydrogen bond interaction and hydrophobic interaction formed between the ligands and the active site of DNA gyrase.

Ligand	Binding Affinity (BA) Kcal/mol	Target	Hydrogen bond		Hydrophobic
			Amino acid	Bond length (A^o)	
1	-9.5	DNA gyrase	SER118 SER118 GLY120	2.41232 2.47934 2.18555	PRO124
2	-9.3	DNA gyrase	GLN101 SER118 SER118	2.6091 2.33129 2.69125	TRP103, GLN277, VAL278, PRO124
8	-12.3	DNA gyrase	SER104 ASN176 GLN101 GLN101	2.82662 2.35700 3.07242 2.91737	LEU105, ARG98, HIS52, LEU105
11	-9.1	DNA gyrase	PRO119 SER118 GLY120	2.0904 2.20816 2.58818	TRP103, TRP103, HIS280 , VAL97, PRO124, PRO119 A, PRO124
16	-84	DNA gyrase	GLY112 PRO108	2.4735 2.96424	HIS490 , PRO108, PRO102, PRO108
19	-8.3	DNA gyrase	GLY17 ASN176	2.05316 2.79531	LEU105
22	-6.2	DNA gyrase	ALA100	2.66283	PRO102, PRO108
23	-8.1	DNA gyrase	TRP103 SER118	2.40711 2.48976	TYR93, PRO124, VAL97, PRO124
24	-6.6	DNA gyrase	ASP94	2.62857	PRO124, PRO124

bonds (2.82662, 2.35700, 3.07242 and 2.91737) with SER104, ASN176, GLN101 and GLN101of the target. In addition, it also formed hydrophobic bond with LEU105, ARG98, HIS52 and LEU105 of the target site. Hydrogen bond interaction between the ligand 3 and DNA gyrase target of *M. tuberculosis*. A total of four hydrogen bonds were formed. The N-H group of triazole ligand as hydrogen donor and formed two hydrogen bonds with SER104 and ASN176 of the target. While the N-H group triazolidine of the ligand also acts as hydrogen donor and formed two hydrogen bonds with GLN101 of the target. The hydrogen bond formation alongside with the hydrophobic interaction provides an evidence [21, 22] that ligand 8 of the inhibitor compound is potent against DNA gyrase receptor.

4. Conclusion

A theoretical approach was employed in this study to selected molecular descriptors to derive a model that could be used to correlate the structure of 1, 2, 4-triazole derivatives as potent inhibitors against *Mycobacterium tuberculosis* and their respective biological activities. 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 1, 2, 4-triazole derivatives can be modeled using molecular descriptors; AATS7s, nHBint3, minHCsatu and Vi. Molecular docking simulation revealed that compounds; 1, 2, 8, 11, 16, 19, 22, 23 and 24 with better activity have higher bind affinity ranging from

(-6.2 and -12.3kcal/mol). However, the lead compound (compound 8) with higher anti-tubercular activity have prominent higher binding affinity of -12.3kcal/mol which indicates that the compounds could serve as a template structure to design compounds with more efficient activities. The outcome of this research and the propose QSAR model develop can be recommended for pharmaceutical and medicinal chemists to design, synthesis and also carry out an in-vivo and in-vitro screening in order to substantiate the computational findings.

Conflicts of interest

There are no conflicts to declare.

References

1. W.H. Organization, others, Tuberculosis Fact Sheet (No. 104) 2000, Site Accessed Www Who Intmediacentrefactsheetswho104enindex Html. (2016).

2. S.E. Adeniji, S. Uba, A. Uzairu, Multi-linear regression model, molecular binding interactions and ligand-based design of some compounds against *Mycobacterium tuberculosis*, Network Modeling Analysis in Health Informatics and Bioinformatics, 9 (2020) 1-18.

3. S.E. Adeniji, S. Uba, A. Uzairu, Geometrical and Topological Descriptors for Activities Modeling of some Potent Inhibitors against Mycobacterium Tuberculosis: A Genetic Functional Approach. Egptian Journal of chemistry, 62 (2019) 1635 - 1647

4. C. W. James, DNA entanglement and the action of the DNA Topoisomerases, Cold pring Harbor Laboratory Press, Cold Spring Harbor, NY. (2009). 245.

5. Y.Y. Huang, J.Y. Deng, J. Gu, Z.P. Zhang, A. Maxwell, L.J. Bi, Y.Y. Chen, Y.F. Zhou, Z.N. Yu, X.E. Zhang, The key DNA-binding residues in the Cterminal domain of *Mycobacterium tuberculosis* DNA gyrase A subunit (GyrA). Nucleic Acids Reseasch, 34 (2006) 5650–5659.

6. Y. Zhang, K. Post-Martens, S. Denkin, New drug candidates and therapeutic targets for tuberculosis therapy. Drug Discovery Today, 11 (2006) 21-27.

7. B.S. Holla, M. Mahalinga, M.S. Karthikeyen, B. Poojary, P.M. Akberali, N.S. Kumari, Synthesis, characterization and anti-microbial activity of some substituted 1, 2, 3-triazoles. European Journal of Medical Chemistry, 40 (2005) 1173-1178.

8. E.A. Sherement, R.I. Tomanov, E.V. Trukhin, V.M. Berestovitskaya, Synthesis of 4- Aryl-5-nitro-1,2,3-triazoles. Russian Journal of Organic Chemistry, 24 (2004) 594-595.

9. H.N. Hafez, H.A. Abbas, A.R. El-Gazzar, Synthesis and evaluation of analgesic, anti-inflammatory and ulcerogenic activities of some triazolo- and 2- pyrazolylpyrido[2,3-d]-pyrimidines. Acta Pharmacy, 58 (2008) 359-378.

10. L.P. Guan, Q.H. Jin G.R. Tian, K.Y. Chai, Z.S. Quan, Synthesis of some quinoline-2 (1H)-one and 1, 2, 4-triazolo[4, 3 -a] quinoline derivatives as potent anticonvulsants. Journal of Pharmarcy and Science, 10 (2007) 254-262.

11. R. Gujjar, A. Marwaha, J. White, L. White, S. Creason, D.M. Shackleford, J. Baldwin, W.N. Charman, Identification of a metabolically stable triazolopyrimidine-based dihydroorotate dehydrogenase inhibitor with activity in mice. Journal of Medicinal Chemistry 52 (2009) 1864-1872.

12. N.B. Patel, I.H. Khan, S.D. Rajani, Pharmacological evaluation and characterizations of newly synthesized 1,2,4-triazoles. *European Journal of Medicinal Chemistry.* 45 (2010).

13. https://patents.justia.com/patent/8865910.

14. S.E. Adeniji, O.B Adalumo , FO. Ekoja Anti-tubercular modelling, molecular docking simulation and insight toward computational design of novel compounds as potent antagonist against DNA gyrase receptor, Medicine in Microecology 5 (2020) 100020.

15. K. Roy, P. Chakraborty, I. Mitra, P.K. Ojha, S. Kar, R.N. Das, 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 (2013) 1071–1082. QSAR models-strategies and importance, Int. J. Drug Des. Discov. 3 (2011) 511–519.

16. S.E. Adeniji, S. Uba, A. Uzairu, Theoretical modeling for predicting the activities of some active compounds as potent inhibitors against *Mycobacterium tuberculosis* using GFA-MLR approach, Journal of King Saud University – Science 32 (2020) 575–586.

17. A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, Mol. Inform. 22 (2003) 69–77.

18. S.E. Adeniji, S. Uba, A. Uzairu, Theoretical modeling and molecular docking simulation for investigating and evaluating some active compounds as potent anti-tubercular agents against MTB CYP121 receptor, Future Journal of Pharmaceutical sciences 4 (2018) 284–295

19. J. Piton, S. Petrella, M. Delarue, G. Andre´-Leroux, V. Jarlier , A. Aubry, C. Mayer, Structural Insights into the Quinolone Resistance Mechanism of *Mycobacterium tuberculosis* DNA Gyrase, PLoS ONE 5 (2010), e12245. doi:10.1371/journal.pone. 0012245.

20. J. Piton, S. Petrella, M. Delarue, G. Andre´-Leroux, V. Jarlier , A. Aubry, C. Mayer, Structural Insights into the Quinolone Resistance Mechanism of *Mycobacterium tuberculosis* DNA Gyrase, https://www.rcsb.org/structure/3IFZ.

21. S.E. Adeniji, Genetic functional algorithm model, docking studies and in silico design of novel proposed compounds against Mycobacterium tuberculosis, Egyptian Journal of Basic and Applied Sciences 7 (2020) 292–314.

22. S.E. Adeniji, D.E Arthur, M. Abdullahi. O.B Adalumo, Computational investigation, virtual docking simulation of 1, 2, 4-Triazole analogues and insillico design of new proposed agents against protein target (3IFZ) binding domain. Bulletin of the National Research Centre. (2020) 44:132