



Research article

Theoretical design of novel antimalarial agents against *P. falciparum* strain, Dd₂ through the QSAR modeling of synthesized 2'-substituted triclosan derivativesZakari Ya'u Ibrahim^{*}, Adamu Uzairu, Gideon Shallangwa, Stephen Abechi

Department of Chemistry, Faculty of Physical Sciences, Ahmadu Bello University, P.M.B 1045, Zaria, Nigeria

ARTICLE INFO

Keywords:

Analytical chemistry
Theoretical chemistry
Pharmaceutical chemistry
QSAR
Molecular design
Antimalarial
2'-substituted triclosan
Molecular descriptors
GF-MLR
P. falciparum strain
Dd₂

ABSTRACT

In an attempt to design compounds with higher antimalarial activities, quantitative structure-activity relationship (QSAR) technique was utilized in the development of a molecular model for some synthesized 2'-substituted triclosan derivatives through a hybrid of the GA-MLR method. The model was found to have excellent statistical parameters ($R^2 = 0.8919$, $R^2_{Adj} = 0.8728$, $LOF = 0.2563$). The descriptors mean effect (MF) revealed BCUTw-11, which increases with an increase in molecular weight, to be the most contributive to the antimalarial activity. Consequently, compound **3**, with the highest activities ($pEC_{50} = 6.9586$) was deployed as the design template. The molecular weight of the template was increasing through substitutions of its atoms at several positions with heavier atoms/groups to increase the descriptor (BCUTw-11) value. Twelve (12) theoretical derivatives of the template were designed where six of the designed derivatives have better activity than the design template. The most active designed compound, **3L** was found to have the highest antimalarial activity ($pEC_{50} = 7.930$) than that of the design template.

1. Introduction

Malaria remained the most incessant and savage health challenges faced by almost half of the world population, common in sub-African, Asian, and South American countries, with children below the ages of five (5) and pregnant women mostly affected [1, 2]. Globally, more than 3 billion people were reported to be prone to malaria, with an estimated death rate of 500 thousand people [2].

Malaria is an infection induced by the genus *Plasmodium* and is passed on through the bite of Anopheles mosquitos. *Plasmodium falciparum* remains the cruelest strain of the five strains of the genus Plasmodium [3]. *P. falciparum* altered the surface of red blood cells once present in the human body through interceding parasite proteins [4]. The hemoglobin is ramshackled into amino-acids and heme by enzymes cysteine and aspartic proteinases [5]. The entire amino-acid constituents are assembled into parasite proteins, although only a fraction of heme is incorporated into parasite hemoproteins, the parasite enzymes detoxified the remaining heme [6].

The endured resistance of *Plasmodium falciparum* to the accustomed antimalarial drugs such as chloroquine, caused this drug to lose its efficacy. In addition to the global transformation of chloroquine-resistant

P. falciparum, resistance to different assortments of quinoline, antifolates, artemisinin, and inhibitors of electron transport, was also developed [7, 8]. Hence, the quest for an efficient antimalarial drug with higher pharmacological activity than the traditionally used antimalarial drugs remain the desired goal. In light of this, synthesized 2'-substituted triclosan derivatives found to be active against the multi-drug resistant *Plasmodium falciparum* strain Dd₂ [9] could provide an alternative application to the routine antimalarial drugs.

The α , β -unsaturated fatty acids double bonds bonded to the acyl carrier protein (ACP) in an NADH or NADPH based reaction were decreased by Enoyl-ACP (acyl carrier protein)-reductase (FabI) [10, 11]. Triclosan, reported to strikingly inhibit its FabI target [12, 13, 14, 15] binds straight to the *P. falciparum* FabI, which enhanced its affinity for the oxidized form of the co-factor NAD⁺ and therefore confining the protein in its NAD-bound form. This restricts the essential NADH binding step, thereby leading the fatty acid synthesis (FAS) as well as the cell growth to a pause [16]. Moreover, in 2002 the modeling analysis by Surolia and Co revealed the vacuum created from the substitution of methionine with alanine in the *P. falciparum* enzyme. This when exploited by the introduction of bulkier groups at carbon 4' of triclosan (2,4,

^{*} Corresponding author.

E-mail address: zakariyadibrahim@gmail.com (Z.Y. Ibrahim).

4'-trichloro-2'-hydroxydiphenylether) will improve the antimalarial efficacy [17].

The desire for an improve drugs with better antimalarial activities lead to the adoption of quantitative structure-activity relationship (QSAR) techniques, an essential process in the field of drug invention, improvement due to its time and cost-effectiveness [18] as well as curtail trial and error during new antimalarial drugs design [19]. The QSAR technique aid in correlating the biological activities of a series of compounds with calculated descriptors [20]. This technique conserves resources and hastens the design of new antimalarial drugs.

Several QSAR studies relating to the design of antimalarial drugs were reported [21, 22, 23, 24, 25] including research on the use of 3-Dimensional QSAR [26] in developing theoretical models for the prediction of antimalarial activity of triclosan derivatives. In 2010, Shah and Siddiqi constructed a 3-Dimensional QSAR model to predict the antimalarial activity of 63 triclosan derivatives against the *P. falciparum* enoyl acyl carrier protein reductase (PfENR) using the steric and electrostatic CoMFA (comparative molecular field analysis) potential fields, the steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor descriptors of CoMSIA (comparative molecular similarity indices analysis); the square correlation coefficient (R^2), F values, and the cross-validated R^2 (Q^2) are 0.96, 195.32, and 0.64 respectively for CoMFA; and 0.97, 209.46, and 0.54 respectively for CoMSIA were achieved using the partial least squares (PLS) analysis. However, no systematic QSAR studies was conducted on a series of synthesized 2'-substituted triclosan derivatives. This research is aimed at developing an excellent QSAR model for estimating the antimalarial activity of synthesized 2'-substituted triclosan derivatives, and utilizing the developed model in the design of novel derivatives of 2'-substituted triclosan with enhanced activity against *P. falciparum* based on the degree of contribution of each descriptor to the developed model.

2. Material and method

2.1. Data set

A data set of 28 synthesized 2'-substituted triclosan derivatives were extracted from PubChem as presented in the literature [9]. The chemical structures and antimalarial activity against the Dd₂ strain of *P. falciparum* are presented in Table 1. Their activities were expressed in logarithmic scale, pEC₅₀ (-LOG10 (EC₅₀)) which was consequently employed in QSAR analysis.

2.2. Descriptors calculation

All the molecular structures were drawn with a ChemDraw Ultra 12 software, and exported into the spartan'14 version 1.1.2 software and then optimized with DFT (DFT/B3LYP/6-31G*) in a vacuum, using the initial molecular geometry [27]. The twenty-eight (28) optimized Spartan 14 structures saved in SDF format, were then exported into PaDEL software where about 1,500 molecular descriptors ranging from Atom Count, Autocorrelation, BaryszMatrix, BCUT, Burden Modified Eigenvalues, Carbon Types, Electrotopological State Atom Type, Extended Topochemical Atom, Autocorrelation3D, and RDF descriptors were calculated [28]. The descriptors were pre-treated, where highly correlated, constant value/empty cell descriptors were all discarded before considering them for the QSAR modeling using the GA-MLR method.

2.3. Model development and selection

The genetic function algorithm (GFA) component of material Studio 8.0 software was used to develop the molecular models [29]. The descriptors as well as the activities of the compounds were imported into

the material Studio software. The compounds were randomly split in ratio 75:25 with 21 training compounds (75% of the data set) and 7 test compounds (covering the remaining 25%). Using the training set, the GFA method was utilized in the modeling by setting the mutation probability to be 0.1 and fixing the smoothing parameter as 0.5, where five (5) top QSAR equations were returned. The models were scored based on the values of the square of the coefficient of determination (R^2), internally cross-validated R^2 (Q^2), and the external validated R^2 (R^2_{pred}) [30, 31], where the model with the highest values of R^2 , Q^2 , and R^2_{pred} was selected as the best model.

2.4. QSAR model validation

2.4.1. Internal validation

The leave-one-out (LOO) cross-validation was the method employed to internally validate the developed model. This is done by excluding a training set data point, then the activities of the remaining data were used to construct a model which was used to test the activity of the excluded data. This process of data exclusion is repeated until the activities of all the training set were predicted. The coefficient of cross-validated R^2 (Q^2_{cv}) is evaluated using Eq. (1).

$$Q^2_{cv} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - Y_{mean})^2} \quad (1)$$

where, Y_i is the actual (experimental) activity and \hat{Y}_i is the predicted activity of the i th molecule in the training set, and Y_{mean} stands for the average activity of all the training set molecules [32, 33].

2.4.2. External validation

The external validation entails predicting the biological activities of some dataset separated from the training set (test set) applying the model. The model external validation was conducted by calculating the predicted correlation coefficient (R^2_{ext}) value for the test set using Eq. (2).

$$R^2_{ext} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - Y_{mean})^2} \quad (2)$$

where, Y_i is the actual (experimental) activity and \hat{Y}_i is the predicted activity of the i th molecule in the test set, and Y_{mean} stands for the average activity of all the training set molecules.

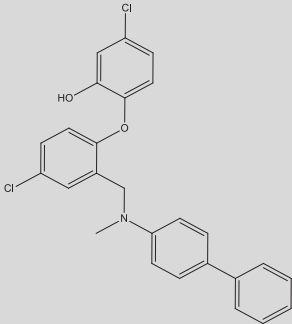
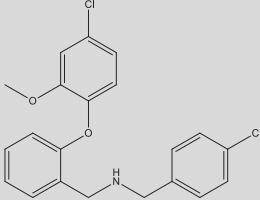
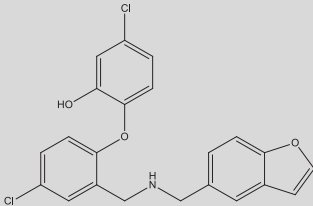
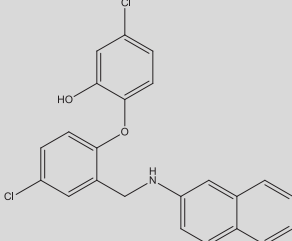
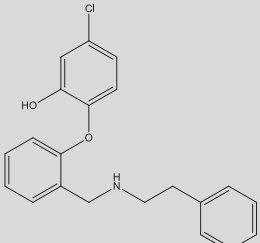
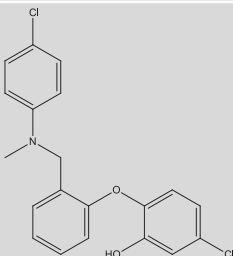
2.5. Y-randomization

The model robustness was analyzed to ascertain whether the model occurs by chance or not [34], through Y-randomization. The Y-randomization test was carried out by jostling the training set activities repeatedly. The jostled activities were then used to construct a model that is compared to the initial non-jostled activities model. The predictions of the jostled activity model are expected to be less than those of the non-jostled activity model with lower statistical parameters, otherwise, the model is not obtainable from data used.

2.6. Mean effect (MF)

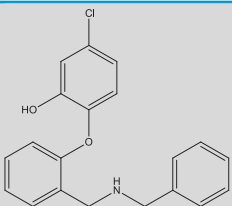
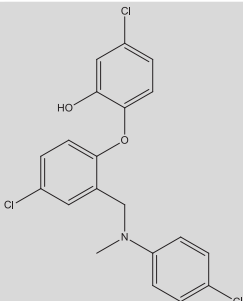
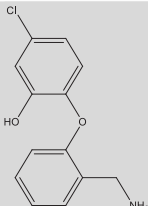
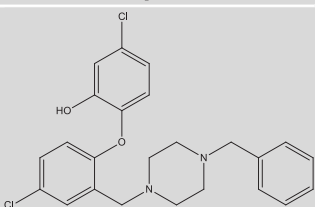
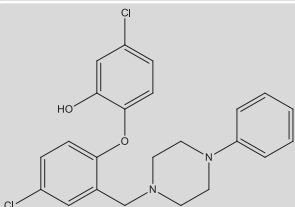
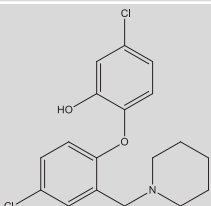
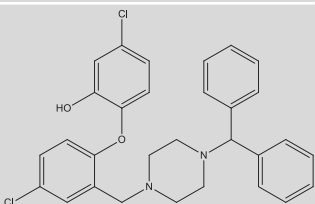
The model descriptors explanation helps in providing detailed information regarding the descriptors that contributed most to the antimalarial activity. The effective contribution of the model descriptor could be analyzed by evaluating their respective mean effect (MF). The variations of their contributions are reflected by the sign that accompanied the MF. The descriptor with the smallest mean effect, have the least contribute towards the activity, while significant contribution is obtained from descriptors with large MF. The mean effect is assessable from Eq. (3).

Table 1. Chemical structures, activities, and residuals of synthesized 2'-substituted triclosan derivatives against multi-drug resistant *Plasmodium falciparum* strain, Dd₂.

S/N	PubChem CID	Structures	EC ₅₀ (μM)	Experimental pEC ₅₀	Predicted pEC ₅₀	Residuals
1	44410210		0.15	6.824	6.680	0.144
2	44410295		24	4.620	4.630	-0.010
3	11495355		0.11	6.959	6.702	0.257
4	44410291		0.14	6.854	6.570	0.284
5	44410287		0.75	6.125	6.262	-0.136
6	44410286		0.3	6.523	6.589	-0.066

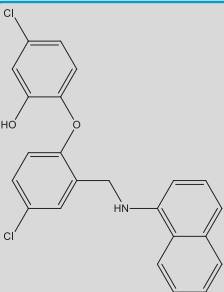
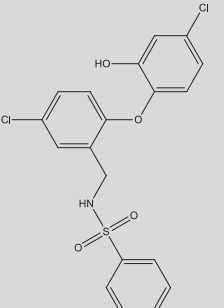
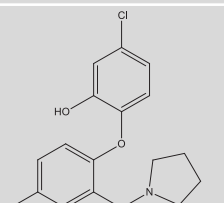
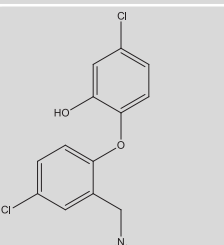
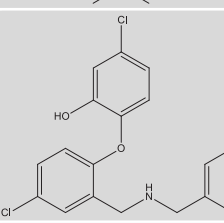
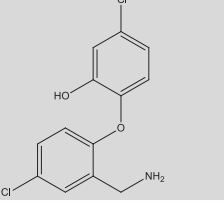
(continued on next page)

Table 1 (continued)

S/N	PubChem CID	Structures	EC ₅₀ (μM)	Experimental pEC ₅₀	Predicted pEC ₅₀	Residuals
7	44410284		0.42	6.377	5.888	0.489
8	44410256		0.27	6.569	6.701	-0.132
9	44410252		7.2	5.143	4.957	0.186
10	44410238		0.62	6.208	5.3424	0.866
11	44410234		15	4.824	5.306	-0.482
12	44410233		0.48	6.319	6.226	0.093
13	44410215		1.4	5.854	5.066	0.788

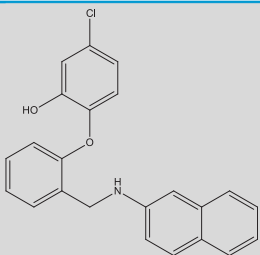
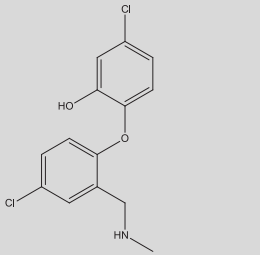
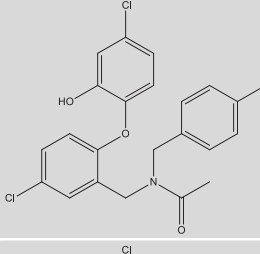
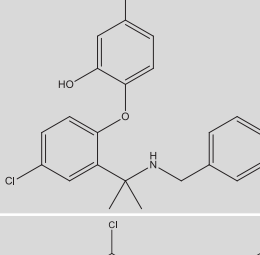
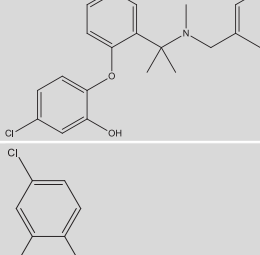
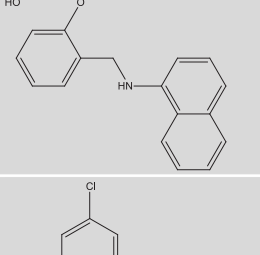
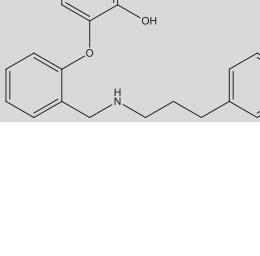
(continued on next page)

Table 1 (continued)

S/N	PubChem CID	Structures	EC ₅₀ (μM)	Experimental pEC ₅₀	Predicted pEC ₅₀	Residuals
14	44410211		2.2	5.658	6.323	-0.665
15	44410170		17	4.770	5.249	-0.479
16	44410138		0.45	6.347	6.275	0.072
17	44410134		0.22	6.658	6.962	-0.304
18	44410130		0.83	6.081	6.001	0.080
19	44410129		2.0	5.699	5.080	0.619

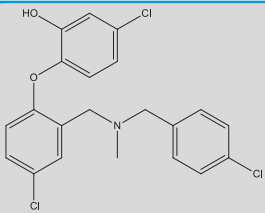
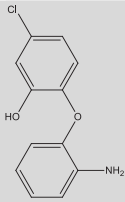
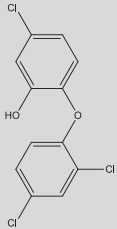
(continued on next page)

Table 1 (continued)

S/N	PubChem CID	Structures	EC ₅₀ (μM)	Experimental pEC ₅₀	Predicted pEC ₅₀	Residuals
20	44410128		0.33	6.481	6.464	0.017
21	44410098		0.47	6.328	5.867	0.461
22	44410097		5.9	5.229	5.105	0.124
23	44410094		0.35	6.456	6.035	0.421
24	11948630		0.38	6.420	6.763	-0.343
25	44410086		0.37	6.432	6.209	0.223
26	44410081		0.39	6.409	6.236	0.173

(continued on next page)

Table 1 (continued)

S/N	PubChem CID	Structures	EC ₅₀ (μM)	Experimental pEC ₅₀	Predicted pEC ₅₀	Residuals
27	44410066		0.19	6.721	6.673	0.048
28	44410251		140	3.854	3.763	0.091
Triclosan	5564		3.8	5.420	4.921	0.499

Test set.

$$\text{Mean Effect} = \frac{\beta_j \sum_i^n D_j}{\sum_j^m (\beta_j \sum_i^n D_j)} \quad (3)$$

where β_j is the coefficient of j , D_j is the value of each training set in the descriptor matrix and m is the tally of model descriptors and n is the tally of molecules used as training set [35].

2.7. Applicability domain (AD)

The ability of a QSAR model to screen chemical compounds is limited by the model's chemical space for only compounds within the domain, since no model can screen all chemical compound no matter how excellently robust and substantiated is the model [34]. The chemical space is measure using the applicability domain of the model which is the plot of the standardization residual activities of the data set against their leverage values. The leverages are the diagonals of a hat matrix (H) calculated from the expressions in Eq. (4) [36].

$$H_i = X_i (X_i^T X_i)^{-1} X_i^T \quad (4)$$

where, H_i , is the hat matrix of the training/test set, X_i is the original matrix of training/test set, and X_i^T is the transpose matrix of the training/test set.

2.8. Molecular design

The most contributive descriptor of the model provides a guide to design numerous hypothetical derivatives of 2'-substituted triclosan with better antimalarial activities using the compound that possess the highest antimalarial activity as the design template by substituting atom/group of atoms at different positions on the template. These are then optimized, descriptors calculated and their activities estimated.

3. Results and discussion

3.1. Regression model

To build the molecular model, 28 derivatives of synthesized 2'-substituted triclosan were split into a training set (21), for model construction and test set (7), for model validation. The GA-MLR studies result in the construction of five (5) molecular models from which the model with the best statistical parameters was selected as the best model shown below:

Regression Equation

$$\text{pEC}_{50} = -0.05474 * \text{ATSC3i} + 408.40626 * \text{BCUTw-11} - 8.87374 * \text{MATS3c} - 4894.23610$$

$$N = 21, R^2 = 0.8919, R^2_{\text{Adj}} = 0.8728, Q^2_{\text{cv}} = 0.8218, \text{LOF} = 0.2563, R^2_{\text{ext}} = 0.7489, N_{\text{ext}} = 7$$

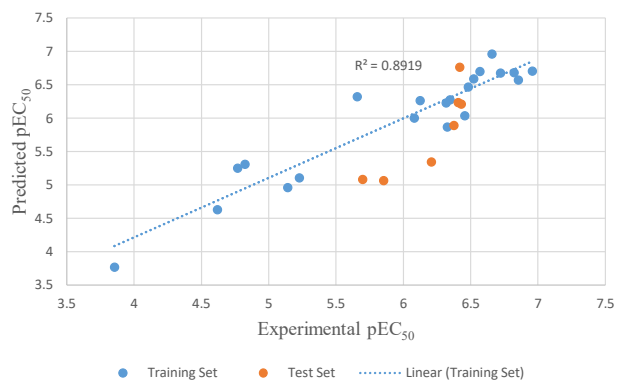
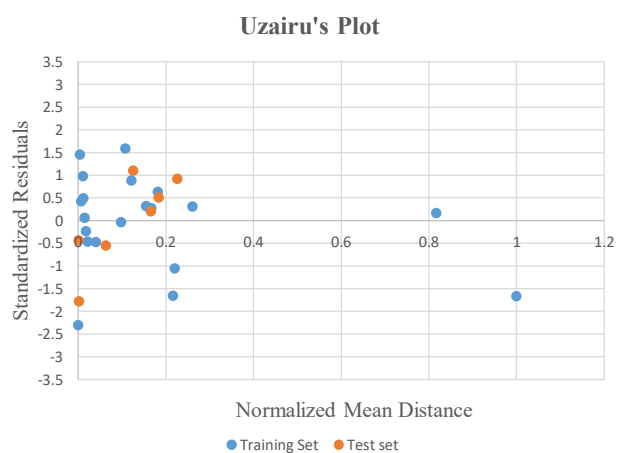
Where N represents the number of the training set, R^2 is the squared correlation coefficient, R^2_{Adj} is the adjusted squared correlation coefficient, Q^2_{cv} is the leave one out squared cross-validation coefficients, LOF is an experimental lack of fit, R^2_{ext} is the squared regression coefficient of the test set, and N_{ext} is the test number.

3.2. QSAR model selection

The three parametric regression equation was selected as the best model due to its highest squared regression coefficient of the test set R^2_{ext} ($R^2_{\text{ext}} = 0.7489$). The statistical parameters of our model, when compared to those of the 3D-QSAR studies of antimalarial activities of triclosan derivatives against PfENR [26] show great improvements over the 3D-QDAR statistical model. Given the 3D-QSAR's F-values, and the cross-validated R^2 (Q^2) as 195.32, and 0.64 respectively for CoMFA and

Table 2. List of the physiochemical descriptors that constitute the model.

S/N	Symbol	Names of descriptors	Class
1	ATSC3i	Centered Broto-Moreau autocorrelation - lag 3/weighted by first ionization potential	2D
2	BCUTw-11	nhigh lowest atom weighted BCUTS	2D
3	MATS3c	Moran autocorrelation - lag 3/weighted by charges	2D

**Figure 1.** Experimental activity against the predicted activity of training and test set.**Figure 2.** A scatter plot of standardized residuals and normalized mean distance.

209.46, and 0.54 respectively for CoMSIA, as against our model with F-value of 46.7427 and the cross-validated R^2 (Q^2) of 0.8218. Hence, our model can predict the antimalarial activity of triclosan derivatives better than the 3D-QSAR proposed model as evidence by the high R^2_{ext} . The selected model was used to predict the activity of the test set, with prediction results as given in Table 1. The definition of the descriptors in the model as well as their class were reported in Table 2. The small residual value for both the training and test set indicates a sound agreement between the predicted and the experimental values as can be observed from the display of the data set around the legend line drawn in Figure 1. Also,

careful observation of the scattered plot of the standardized residuals vs normalized mean distance otherwise known as Uzairu's plot (Figure 2) confirmed the predictive strength of the model.

3.3. QSAR model validation

The internal validation method involves the use of leave one out cross-validation coefficients (Q^2_{cv}), obtained by modeling the entire training set except one compound dispelled and the built model is used to predict the activity of the dispelled compound. The high value of $Q^2_{cv} = 0.8218$ is an indication of good internal validation. The external validation ensured splitting the data set into a model building (training) set model validating (test) set. The high square correlation coefficient of the test set, $R^2_{ext} = 0.7489$ shows the predictive strength of the model and hence its robustness. The correlation analysis carried out between the descriptors and the antimalarial activity, and between the descriptors themselves as can be seen in the correlation matrix (Table 3), shows a high correlation between the descriptors and the activity, pointing to the meaningful contribution of the descriptors to the antimalarial activity, on the other hand, the low correlation coefficients between the descriptors, shows noncollinearity between them. The variation inflation factors (VIF), explains the multi-collinearity existing between the model descriptors. The VIF is calculated from Eq. (5) expressed below:

$$VIF = 1 / (1 - R^2) \quad (5)$$

where R^2 is the squared correlation coefficient between the variables within the model.

The VIF values of the model's three descriptors as displayed in the correlation matrix (Table 3), revealed values are within, $1 < VIF \leq 5$ range. These values are within VIF criteria for model acceptance [37, 38]. Hence, the descriptors are orthogonal and the model is acceptable. The results of Y-randomization carried out after jostling the activity fifteen (15) times are presented in Table 4. The low values of R^2 and Q^2_{LOO} agree that the non-jostled model is significant and can be obtained through the GA-MLR method. The additional parameter of $cR^2_p > 0.5$ reestablish the power of the model and does not occur by chance. A plot of standardized residuals against the leverage values calculated for all compounds (Williams plot) gives room for the identification of outlier and model application limit. The Williams plot displayed in Figure 3, shows all data set except for compound 28, were within ± 3 standard deviation and threshold leverage, $h^* = 0.5714$. Compound 28 was considered as an outlier since it has a leverage value ($h = 0.7410$) greater than that of the threshold. The threshold leverage, h^* is obtained from Eq. (6).

$$h^* = \frac{3(p+1)}{n} \quad (6)$$

Table 3. Correlation matrix of the descriptors with their respective VIF.

	pEC ₅₀	ATSC3i	MATS3c	BCUTw-11	VIF
pEC ₅₀	1				
ATSC3i	-0.3531	1			1.1290
MATS3c	-0.5783	-0.0137	1		1.0588
BCUTw-11	0.6113	0.3321	-0.2265	1	1.1899

Table 4. MLR Y-Randomization test.

Model	R	R ²	Q ²
Original	0.9444	0.8919	0.8218
Random 1	0.4434	0.1966	-0.2429
Random 2	0.3750	0.1406	-0.4196
Random 3	0.5388	0.2903	-0.0780
Random 4	0.2063	0.0426	-1.0348
Random 5	0.4784	0.2289	-0.1524
Random 6	0.1016	0.0103	-0.3795
Random 7	0.2689	0.0723	-0.3045
Random 8	0.0431	0.0019	-0.8371
Random 9	0.4375	0.1914	-0.5121
Random 10	0.2145	0.0460	-1.0964
Random 11	0.3923	0.1539	-0.1617
Random 12	0.2540	0.0645	-0.3729
Random 13	0.3713	0.1378	-1.2731
Random 14	0.3372	0.1137	-0.3661
Random 15	0.4754	0.2260	-0.2301

Random Models Parameters	
Average r:	0.3292
Average r ² :	0.1278
Average Q ² :	-0.4974
cRp ² :	0.8359

where p, refers to the sum of the model's descriptor and n, the entire number of data set [39, 40].

3.4. Interpretation of descriptors

The descriptors present in the model ($pEC_{50} = -0.05474 \cdot ATSC3i + 408.40626 \cdot BCUTw-11 - 8.87374 \cdot MATS3c - 4894.23610$) are ATSC3i, BCUTw-11, and MATS3c. The descriptors ATSC3i and MATS3c both have a negative contribution towards the antimalarial activity as indicated in the model, while the contribution of BCUTw-11, is a positive one due to the positive regression coefficient. The relative contributions of each descriptor in the model were carried out in the mean effect (MF) analysis shown in Figure 4. In the figure, descriptor, MATS3c is the least contributive, followed by ATSC3i, with BCUTw-11, as the most contributive descriptor.

ATSC3i descriptor is the Centered Broto-Moreau autocorrelation of lag 3 weighted by first ionization potential. It estimates the correlation

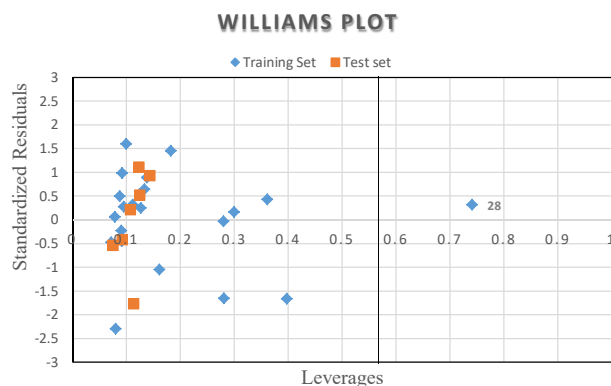


Figure 3. The plot of the standardized residuals against the leverages of the data set.

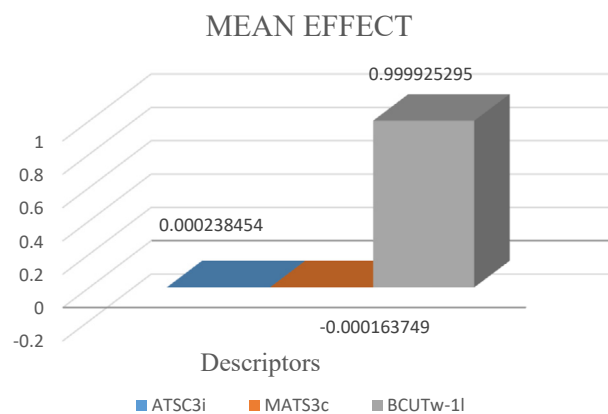


Figure 4. 3-D clustered column descriptors mean effect.

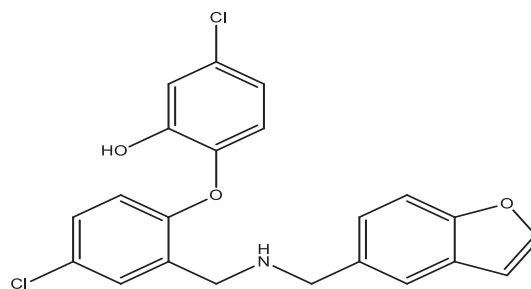


Figure 5. Design Template (Compound 3), 2-(2-(((benzofuran-5-ylmethyl)amino)methyl)-4-chlorophenoxy)-5-chlorophenol, with $pEC_{50} = 6.9586$.

among the first potential of ionization [41] partitioned by three bonds. The sign of its mean effect is positive, revealing an increase in the activity with an increase in the first ionization potential (ATSC3i) value of the molecule.

MATS3c descriptor is the Moran autocorrelation of lag 3 weighted by partial charges. It estimates correlation among charges partitioned by three bonds [42]. Like with ATSC3i, the negative sign of its mean effect indicates that the antimalarial activity decreases with an increase in the molecular branch.

BCUTw-11 is the Eigenvalue based descriptor. It illustrates a chemical diversity by atomic weight, partial charge, and polarizability. The value of the descriptor increases with increase in atomic weight [43]. The mean effect of BCUTw-11 has a positive sign, pointing to an increase in the activity value through increasing the atomic weight.

3.5. Molecular design

Compound 3, 2-(2-(((benzofuran-5-ylmethyl)amino)methyl)-4-chlorophenoxy)-5-chlorophenol that possess the highest activity ($pEC_{50} = 6.9586$), was utilized as the design template (Figure 5). The descriptor, BCUTw-11, was found to be the most influential descriptor, played an active role in the design. The descriptor values increase with an increase in atomic weight. Hence, the weight of the design template was increased by substituting several of its atoms at a different position with heavier atoms/groups. Twelve (12) derivatives of the template were designed, out of which six of the designed compounds (3B, 3C, 3D, 3I, 3J, and 3L) have better activities than the template compound as displayed in Table 5. Compound 3L, 5-(((5-chloro-2-(4-chloro-2-hydroxyphenoxy)benzyl)amino)methyl)-7-(chloromethyl)benzofuran-2-ol, was found to have the highest antimalarial activity ($pEC_{50} = 7.930$)

Table 5. Structures of the design template, designed derivatives of 2'-substituted triclosan and Triclosan along with their respective activities.

DESIGNED DERIVATIVES								Activities	
Compound	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Activities		
							EC ₅₀ (μM)	pEC ₅₀	
3 (Template)	H	H	H	H	H	H	0.110	6.959	
3A	H	H	H	OH	H	H	0.246	6.609	
3B	H	H	H	Cl	H	H	0.070	7.153	
3C	Cl	H	H	H	H	H	0.106	6.976	
3D	H	Cl	H	H	H	H	0.104	6.982	
3E	H	H	Cl	H	H	H	0.124	6.905	
3F	H	H	H	OCH ₃	H	H	0.249	6.603	
3G	H	H	H	H	H	Cl	0.112	6.951	
3H	H	H	H	H	H	OH	0.239	6.622	
3I	H	H	H	H	H	OH	0.046	7.339	
3J	H	H	H	OCH ₃	H	OH	0.061	7.217	
3K	H	H	H	H	CH ₂ OH	OH	0.141	6.850	
3L	H	H	H	H	CH ₂ Cl	OH	0.012	7.930	
Triclosan							3.800	5.420	

as reflected (Table 5), as such is the most active of the all designed derivatives.

4. Conclusion

The research aimed at developing a QSAR model utilizing 28 synthesized 2'-substituted triclosan derivatives. And its application in designing novel antimalarial drugs with enhanced activities activity against *P. falciparum*. In achieving the aim, descriptors, ATSC3i, BCUTw-11, and MATS3c were found to influence the antimalarial activities greatly out of several hundred calculated descriptors. The developed model was validated through different methods, and their descriptors mean effect calculated. The mean effect calculated revealed BCUTw-11 (a chemical diversity by atomic weight, partial charge, and polarizability) to be responsible for the antimalarial property of synthesized 2'-substituted triclosan derivatives and used to design twelve derivatives of 2'-substituted triclosan with enhanced activities. From which compound 5-(((5-chloro-2-(4-chloro-2-hydroxyphenoxy)benzyl)amino)methyl)-7-(chloromethyl)benzofuran-2-ol, was found to have the highest hypothetical activity, pEC₅₀ = 7.930. The designed compounds could be synthesis and clinically validated for antimalarial treatment.

Declaration

Author contribution statement

Zakari Ya'u Ibrahim: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Adamu Uzairu: Conceived and designed the experiments; Analyzed and interpreted the data.

Gideon Shallangwa: Performed the experiments; Contributed reagents, materials, analysis tools or data.

Stephen Abechi: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

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

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- [1] World Health Organization, World Malaria Report 2015, World Health Organization, Geneva, Switzerland, 2016.
- [2] S. Manohar, M. Tripathi, D.S. Rawat, 4-aminoquinoline based molecular hybrids as antimalarial: an overview, *Curr. Top. Med. Chem.* 14 (2014) 1706–1733.
- [3] T.N.C. Wells, P.L. Alonso, W.E. Gutteridge, New medicines to improve control and contribute to the eradication of malaria, *Nat. Rev. Drug Discov.* 8 (11) (2009) 879–891.
- [4] S. Bowman, D. Lawson, D. Basham, D. Brown, T. Chillingworth, C.M. Churcher, B.G. Barrell, The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*, *Nature* 400 (6744) (1999) 532–538.
- [5] A.V. Pandey, B.L. Tekwani, R.L. Singh, V.S. Chauhan, Artemisinin, an endoperoxide antimalarial, disrupts the hemoglobin catabolism and heme detoxification systems in malarial parasite, *J. Biol. Chem.* 274 (27) (1999) 19383–19388.

- [6] S. Kamchonwongpaisan, E. Samoff, S. Meshnick, Identification of hemoglobin degradation products in *Plasmodium falciparum*, *Mol. Biochem. Parasitol.* 86 (2) (1997) 179–186.
- [7] J.E. Hyde, Drug-resistant malaria, *Trends Parasitol.* 21 (11) (2005) 494–498.
- [8] A.M. Dondorp, F. Yi, P. Nosten, D. Das, A.P. Phyo, J. Tarning, N.J. White, Artemisinin resistance in *Plasmodium falciparum* malaria, *N. Engl. J. Med.* 361 (5) (2009) 455–467.
- [9] J.S. Freundlich, M. Yu, E. Lucumi, M. Kuo, H.-C. Tsai, J.-C. Valderramos, J.C. Sacchetti, Synthesis and biological activity of diaryl ether inhibitors of malarial enoyl acyl carrier protein reductase. Part 2: 2'-Substituted triclosan derivatives, *Bioorg. Med. Chem. Lett.* 16 (8) (2006) 2163–2169.
- [10] H. Bergler, P. Wallner, A.B. Leitinger, S. Fuchsichler, H. Aschauer, G. Kollenz, G. Hogenauer, F. Turnowsky, Protein EnvM is the NADH-dependent enoyl-ACP reductase (FabI) of *Escherichia coli*, *J. Biol. Chem.* 269 (1994) 5493–5496.
- [11] R.J. Heath, C.O. Rock, Enoyl-acyl carrier protein reductase (FabI) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*, *J. Biol. Chem.* 270 (1995) 26538–26542.
- [12] L.M. McMurry, M. Oethinger, S.B. Levy, Triclosan targets lipid synthesis, *Nature* 394 (1998) 531–532.
- [13] R.J. Heath, J.R. Rubin, D.R. Holland, E. Zhang, M.E. Snow, C.O. Rock, Mechanism of triclosan inhibition of bacterial fatty acid synthesis, *J. Biol. Chem.* 274 (1999) 11110–11114.
- [14] R.J. Heath, Y.T. Yu, M.A. Shapiro, E. Olson, C.O. Rock, Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis, *J. Biol. Chem.* 273 (1998) 30316–30320.
- [15] C.W. Levy, A. Roujeinikova, S. Sedelnikova, P.J. Baker, A.R. Stuitje, A.R. Slabas, D.W. Rice, J.B. Rafferty, Molecular basis of triclosan activity, *Nature* 398 (1999) 383–384.
- [16] K. Sugunam, A. Surolia, N. Surolia, Structural basis for triclosan and NAD binding to enoyl-ACP reductase of *Plasmodium falciparum*, *Biochem. Biophys. Res. Commun.* 283 (2001) 224–228.
- [17] N. Surolia, P.R. Satish, A. Surolia, Paradigm shifts in malaria parasite biochemistry and anti-malarial chemotherapy, *Bioessays* 24 (2002) 192–196.
- [18] S. Ekins, J. Mestres, B. Testa, In silico pharmacology for drug discovery: applications to targets and beyond, *Br. J. Pharmacol.* 152 (1) (2007) 21–37.
- [19] R. Hadanu, S. Idris, I.W. Sutapa, QSAR analysis of benzothiazole derivatives of antimalarial compounds based on AM1 semi-empirical method, *Indones J. Chem.* 15 (1) (2015) 86–92.
- [20] C. Hansch, A. Kurup, R. Garg, H. Gao, Chem-bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms, *Chem. Rev.* 101 (3) (2001) 619–672.
- [21] Z.Y. Ibrahim, A. Uzairu, G. Shallangwa, S. Abechi, QSAR and molecular docking based design of some indolyl-3-ethanone- α -thioethers derivatives as *Plasmodium falciparum* dihydroorotate dehydrogenase (PfDHODH) inhibitors, *SN Appl. Sci.* 2 (7) (2020).
- [22] C. Santos, J. Vieira, C. Lobato, L. Hage-Melim, R. Souto, C. Lima, J. Carvalho, A SAR and QSAR study of new artemisinin compounds with antimalarial activity, *Molecules* 19 (1) (2013) 367–399.
- [23] H. Liu, L. Qu, H. Gao, J. Wang, L. Han, B. Xiang, Study on the quantitative structure-activity relationship of C-10 substituted artemisinin (QHS)'s derivatives using rough set theory, *Sci. China Ser. B: Chem.* 51 (10) (2008) 937–945.
- [24] A. Najafi, S. Sobhanardakani, M. Marjani, Exploring QSAR for antimalarial activities and drug distribution within blood of a series of 4-aminoquinoline drugs using genetic-MLR, *J. Chem.* 2013 (2013) 1–12.
- [25] R. Hadanu, S. Mastjeh, M. Mustofa, E.N. Sholikhah, M.A. Wijayanti, I. Tahir, Quantitative structure-activity relationship analysis (qsar) of antimalarial 1,10-phenanthroline derivatives compounds, *Indones. J. Chem.* 7 (1) (2007) 72–77.
- [26] P. Shah, M.I. Siddiqi, 3D-QSAR studies on triclosan derivatives as *Plasmodium falciparum* enoyl acyl carrier reductase inhibitors, *SAR QSAR Environ. Res.* 21 (5-6) (2010) 527–545.
- [27] A. Oluwaseye, A. Uzairu, G.A. Shallangwa, S.E. Abechi, Quantum chemical descriptors in the QSAR studies of compounds active in maxima electroshock seizure test, *J. King Saud Univ. Sci.* (2018).
- [28] C.W. Yap, PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints, *J. Comput. Chem.* 32 (7) (2010) 1466–1474.
- [29] U. Abdulfatai, A. Uzairu, S. Uba, Quantitative structure activity relationship study of anticonvulsant activity of α -substituted acetamido-N-benzylacetamide derivatives, *Cogent Chem.* 2 (1) (2016).
- [30] D.B. Kell, B. Sonnleitner, GMP — good modelling practice: an essential component of good manufacturing practice, *Trends Biotechnol.* 13 (11) (1995) 481–492.
- [31] J. Aires-de-Sousa, M.C. Hemmer, J. Gasteiger, Prediction of ¹H NMR chemical shifts using neural networks, *Anal. Chem.* 74 (1) (2002) 80–90.
- [32] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Vol. 2, Montreal, 20-25 August, 1995, pp. 1137–1145.
- [33] G. Schüürmann, R.-U. Ebert, J. Chen, B. Wang, R. Kühne, External validation and prediction employing the predictive squared correlation coefficient test set activity mean vs training set activity mean, *J. Chem. Inf. Model.* 48 (11) (2008) 2140–2145.
- [34] A. Tropsha, P. Gramatica, V. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, *QSAR Comb. Sci.* 22 (1) (2003) 69–77.
- [35] A. Habibi-Yangjeh, M. Danandeh-Jenagharad, Application of a genetic algorithm and an artificial neural network for global prediction of the toxicity of phenols to *Tetrahymena pyriformis*, *Monatshfte Für Chemie - Chemical Monthly* 140 (11) (2009) 1279–1288.
- [36] N. Minovski, Š. Zuperl, V. Drgan, M. Novič, Assessment of applicability domain for multivariate counter-propagation artificial neural network predictive models by minimum Euclidean distance space analysis: a case study, *Anal. Chim. Acta* 759 (2013) 28–42.
- [37] S. Shapiro, B. Guggenheim, Inhibition of oral bacteria by phenolic compounds. Part 1. QSAR analysis using molecular connectivity, *Quant. Struct.-Act. Relat.* 17 (1998) 327–337.
- [38] M. Jaiswal, P.V. Khadikar, A. Scozzafava, C.T. Supuran, Carbonic anhydrase inhibitors: the first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides, *Bioorg. Med. Chem. Lett.* 14 (12) (2004) 3283–3290.
- [39] T.I. Netzeva, A.P. Worth, T. Aldenberg, R. Benigni, M.T.D. Cronin, P. Gramatica, C. Yang, Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships, *Alternatives Lab. Anim.* 33 (2) (2005) 155–173.
- [40] OECD, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models, Organisation for Economic Co-Operation and Development, Paris, France, 2007.
- [41] J.L. Baldim, B.G.V. de Alcântara, O. da S. Domingos, M.G. Soares, I.S. Caldas, R.D. Novaes, D.A. Chagas-Paula, The correlation between chemical structures and antioxidant, prooxidant, and antityrosinomatid properties of flavonoids, *Oxid. Med. Cell. Longev.* 2017 (2017) 1–12.
- [42] B.A. Thurston, A.L. Ferguson, Machine learning and molecular design of self-assembling -conjugated oligopeptides, *Mol. Simulat.* 44 (11) (2018) 930–945.
- [43] A.L. Hook, D.J. Scurr, J.C. Burley, R. Langer, D.G. Anderson, M.C. Davies, M.R. Alexander, Analysis and prediction of defects in UV photo-initiated polymer microarrays, *J. Mater. Chem. B* 1 (7) (2013) 1035–1043.