



QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP STUDY ON THE INHIBITORY ACTIVITY OF COPPER SCHIFF BASE COMPLEXES AGAINST CANDIDA ALBICAN

U. MUHAMMAD^{1,*}, A. UZAIRU² and S.O. IDRIS²

¹Department of Pre-nd Sci. and Tech., School of General Studies, Kano State Polytechnic

²Department of Chemistry, Faculty of Science, Ahmadu Bello University, Zaria

* Correspondence Author's. ¹Department of Pre-nd Sci. and Tech., School of General Studies, Kano State Polytechnic, phone: +2348036755448 email address: umjidda58@gmail.com (U. Muhammad).

ARTICLE INFO

Article history:

Received 2018-08-07

Accepted 2018-08-31

Available online 2019-03-08

keywords

Candidaalbican,

Copper Schiff base complex,

Genetic Function Algorithm,

QSAR

ABSTRACT

Quantitative Structure Activity Relationship (QSAR) study was carried out on Copper Schiff base complexes. Multiple Linear Regression analysis and genetic function algorithms was employed to derive QSAR model for better activity. The derived QSAR model having Squared Correlation Coefficient $R^2 = 0.8345$, Cross Validation Squared Correlation Coefficient $Q^2 = 0.6681$ and predicted R squared (R^2_{pred}) = 0.5980. The robustness of the derived model was confirmed by internal and external cross validation techniques. The QSAR model indicate that the descriptors (MATS4p) Moron autocorrelation of lag 4 weighed by polarizability, (RCI) Ring Complexity index, (G2m) 2nd component symmetry directional WHIM index/weighted by mass, BIO [N-O] Presence/absence of N-O at topological distance 10 and (nF) Number of Fluorine atoms plays an important role in predicting the activities against anti-candida albican. The result obtained in this study was used in designing more potent Copper Complexes as anti-candida albican agents.

1. INTRODUCTION

The emergence and spread of antimicrobial resistance have become one of the most serious public health concerns across the world. Antimicrobial resistance refers to micro-organism that have developed the ability to in activate or block the inhibitory or lethal mechanism of the antimicrobial agent (Tenover,2006). The yeast *Candida albican* is considered to be the most opportunistic fungal pathogen as it presents itself in immunocompromised patients and over 75% of women suffer from vaginal candidiasis at some stage in their life time (Rathod, 2011). *Candida* infections are diverse in their manifestation , varying from superficial skin problems , chronic infection of the nails , mouth , throat or vagina to frequently fatal systemic diseases that involve the lungs , heart , gastrointestinal tract , central nervous system and other organs (Kim and Sudbery,2011).The occurrence of such disease is relatively constant. Despite the relative triviality of the infections, their eradication is often problematic and requires many weeks or months of therapy (Ajala *et al.*, 2016).

Metal based drugs are a research area of increasing interest for inorganic, pharmaceutical and medicinal chemistry and have concentrated much attention as an approach to new drug development (Kovala-Demertzi, 2006).The Schiff bases are widely used ligands due to their facile synthesis, significant versatility and good solubility in common solvents. Thus, they have played a vital role in the development of coordination chemistry as they readily form stable complexes with most metals in different oxidation states (Kumaran *et al.*, 2013)

Quantitative Structure Activity Relationship (QSAR) is one of the most important area in medicinal chemistry, which gives information that is useful for drug design. QSAR models are mathematical equations relating chemical structure to a wide variety of physical, chemical and biological

properties. The derived relationship between molecular descriptors and activity is used to estimate the property of other molecules and / or to find the parameters affecting the biological activity. The application of quantitative structure activity relationship (QSAR) methodologies has potential to decrease substantially the time and effort required to discover new medicines or to improve current ones in terms of their efficacy by avoiding the conventional trial and error approach employed in the discovery and development of novel medicines by avoiding leads unlikely to be successful. SAR studies are usually carried out by making minor changes to the structure of a lead to produce analogs and assessing the effect that these structural changes have on biological activity (Motta and Almeida, 2011).

The main aim of this work was to develop a mathematical model for predicting the inhibitory activity of Cu (II) complexes and use the information gathered to design more potent Cu (II) complexes for the treatment of infections caused by *Candida albican*.

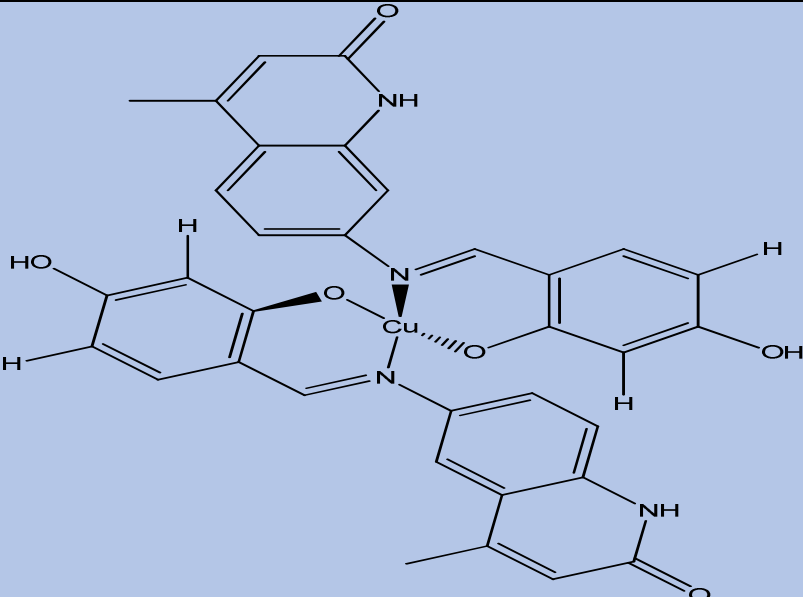
2. MATERIALS AND METHOD

2.1 QSAR studies

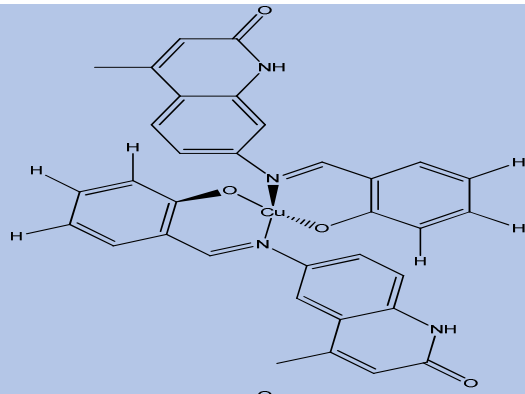
2.1.1 Data collection

The data set of twenty six Copper Complexes with their experimental minimum inhibitory concentration (MIC) values in $\mu\text{g/ml}$ were taken from the literature (Creaven *et al.*, 2010, 2009; Raman *et al.*,2006). The MIC values of the complexes were converted to $-\log\text{MIC}$ to decrease the variance and to approach normal distribution which is response variable for the QSAR analysis. The $-\log$ values of MIC along with the structure of complexes are presented in Table 1.

Table 1- Shows the structures and the activity (pIC_{50}) of the dataset.

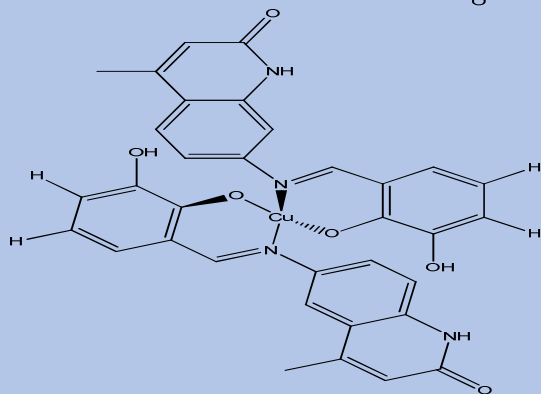
S/No	Structures	pMIC_{50}
1		5.273

2



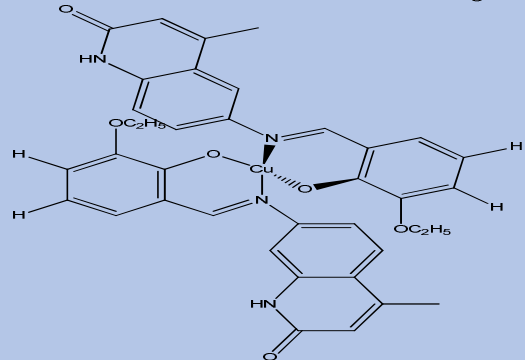
5.137

3



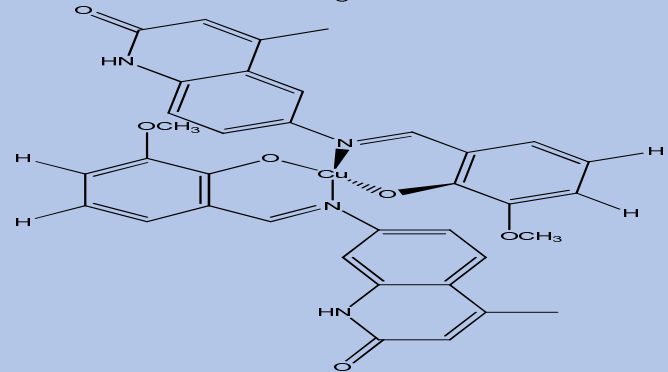
5.534

4



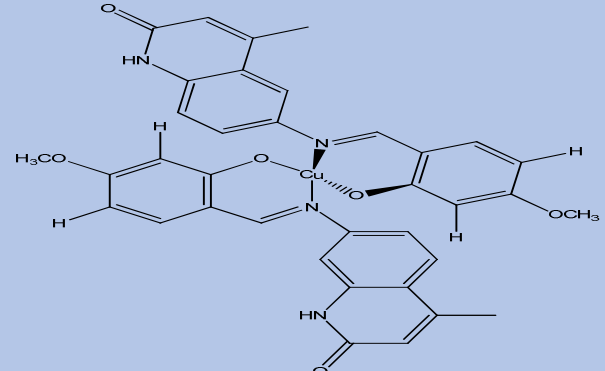
5.518

5



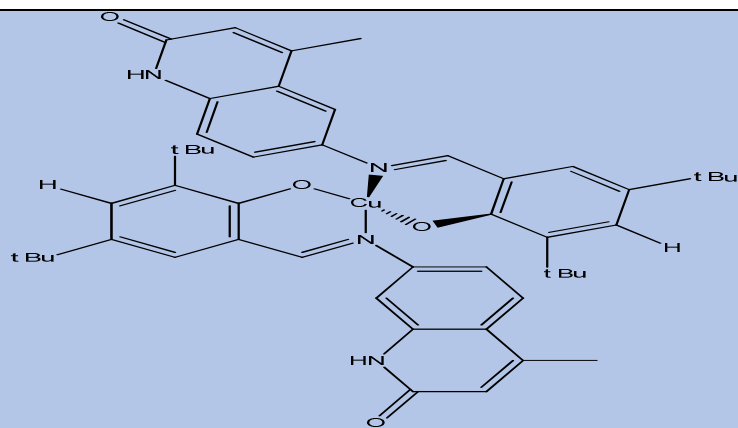
5.535

6



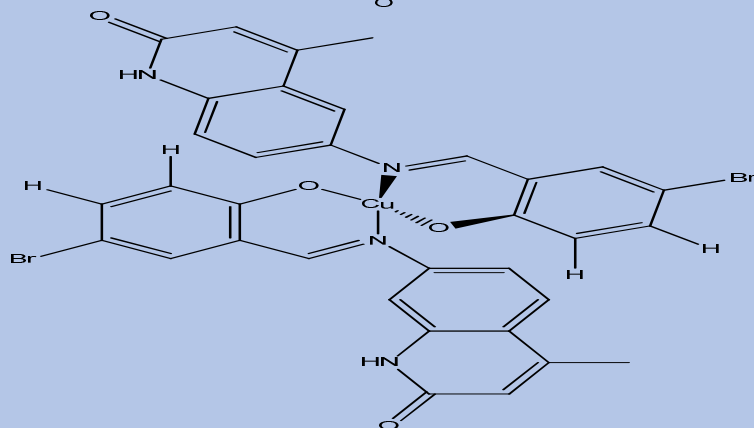
5.299

7



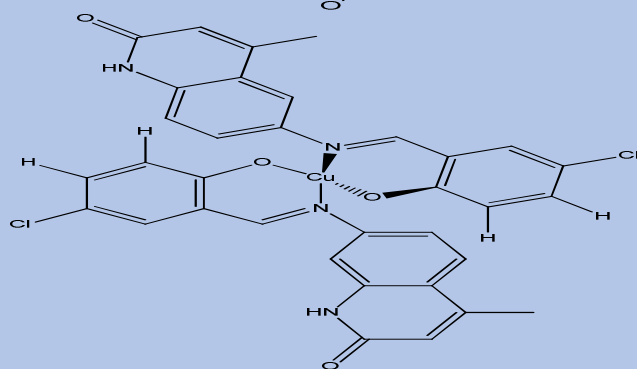
4.779

8



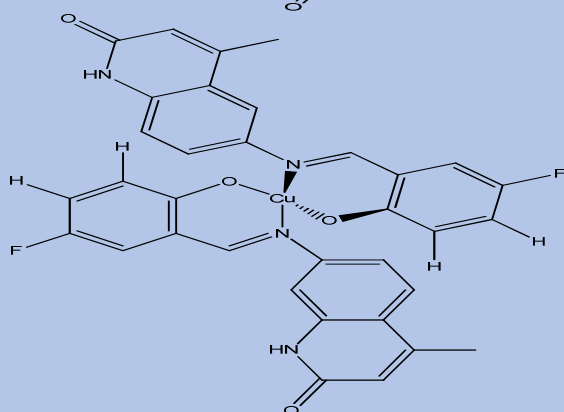
5.161

9



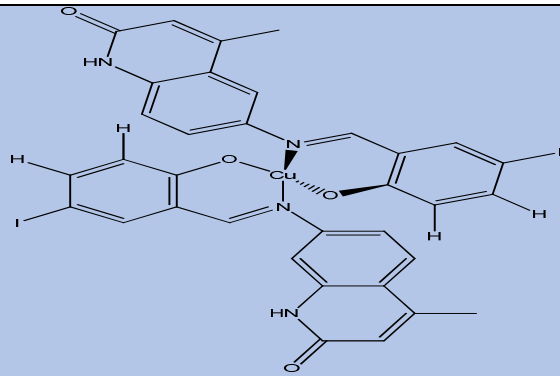
5.229

10



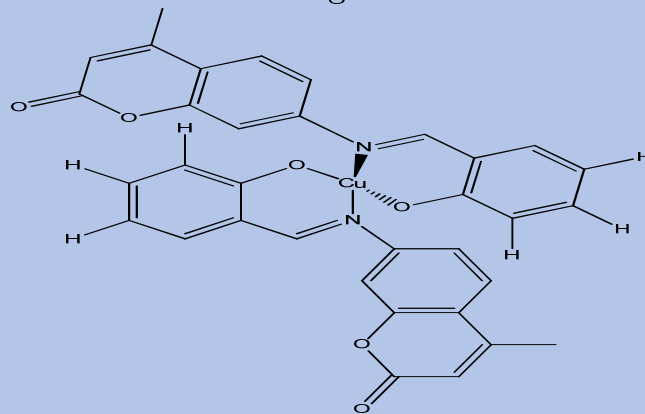
4.425

11



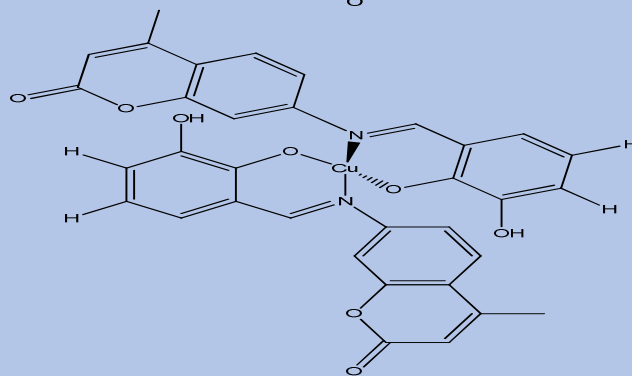
4.439

12



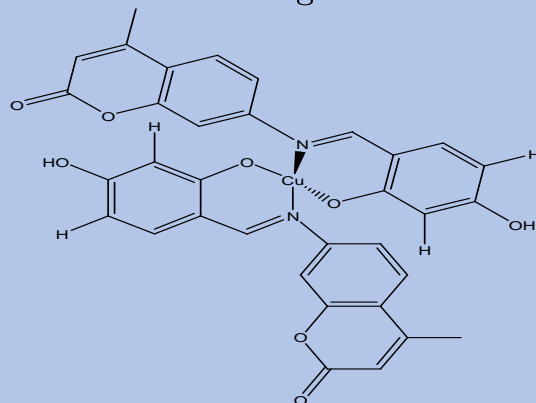
5.491

13



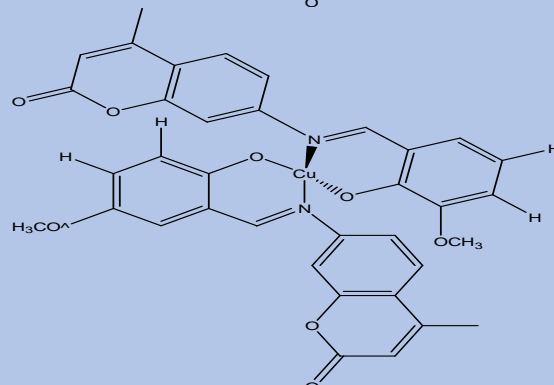
5.169

14



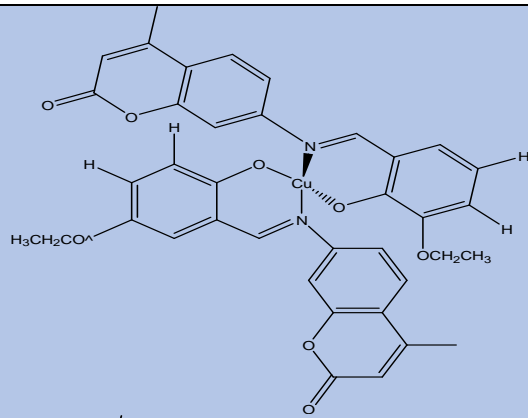
4.963

15



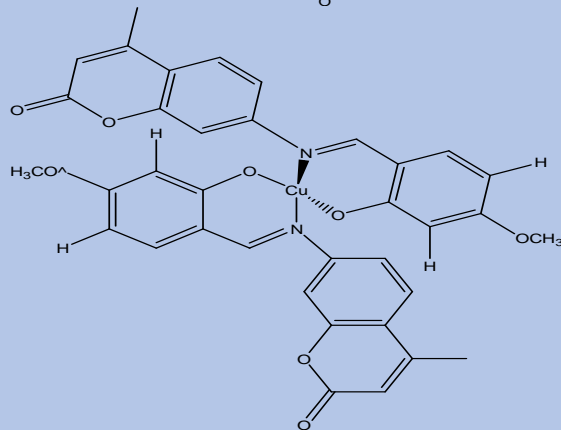
5.254

16



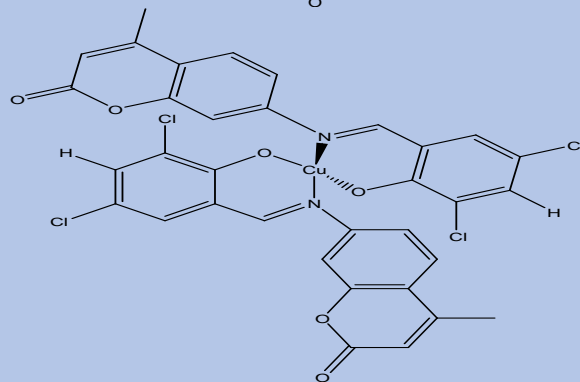
5.159

17



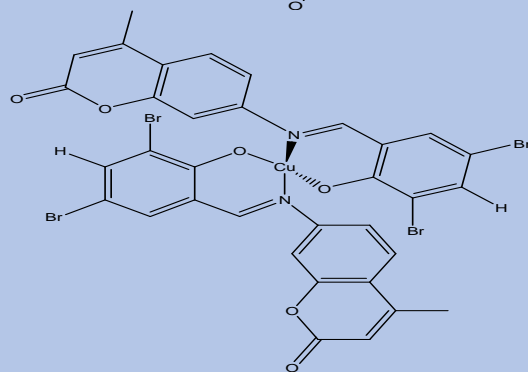
5.006

18



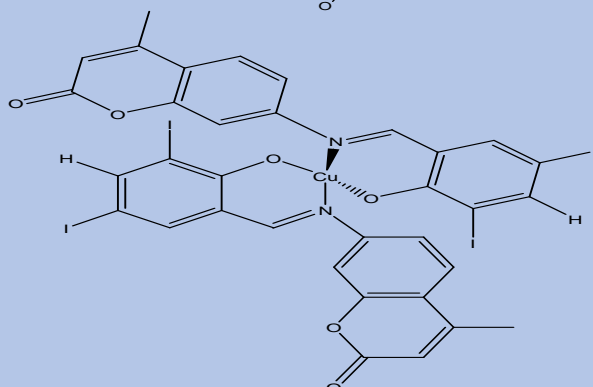
5.564

19

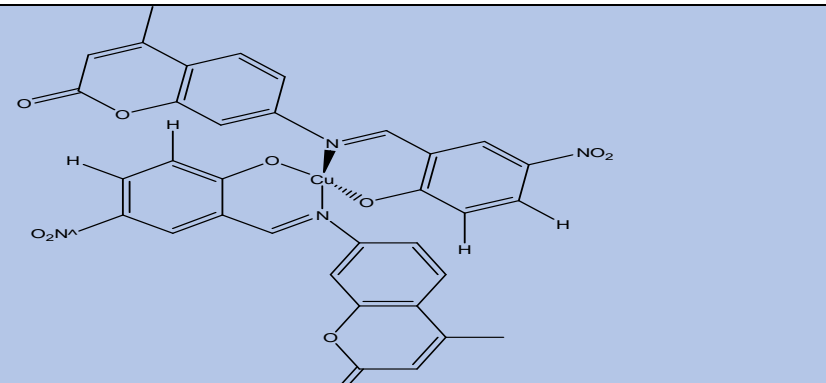
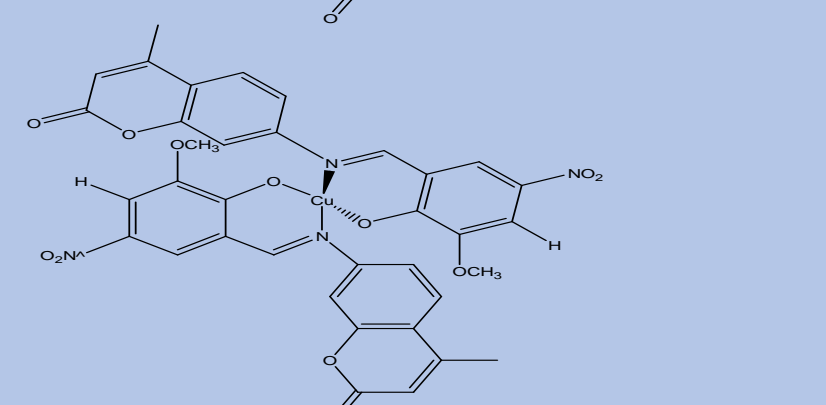
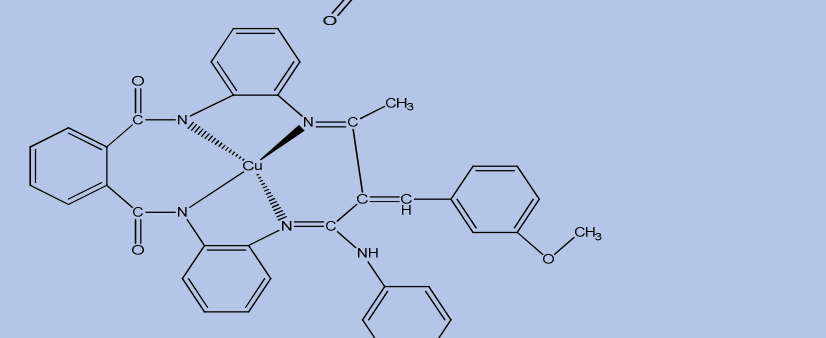
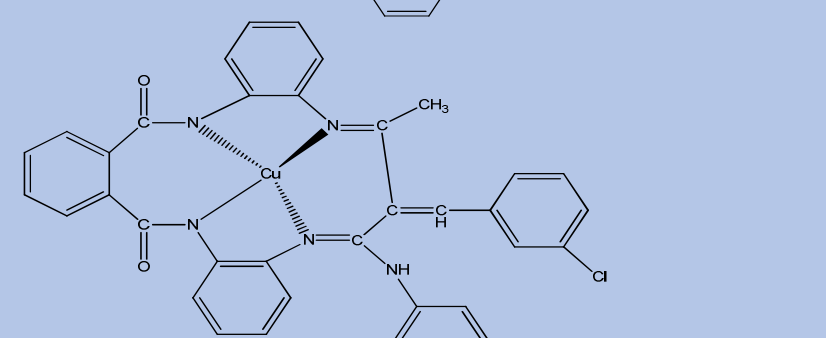
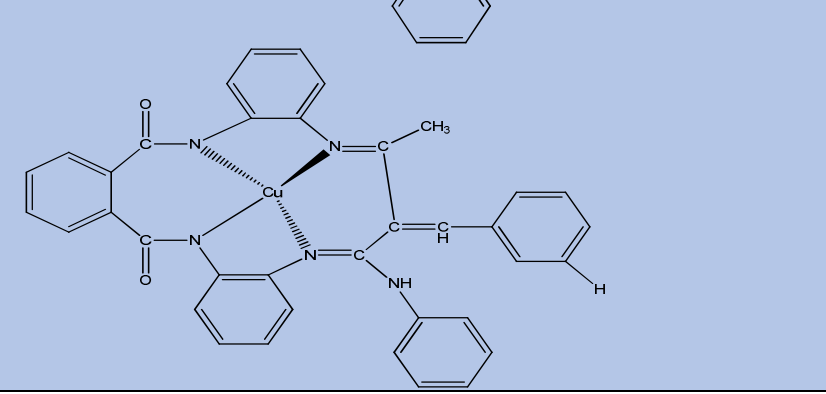


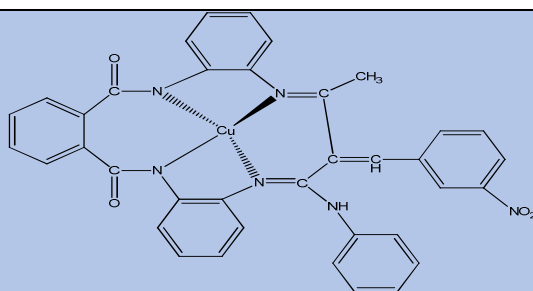
5.385

20



6.104

21		5.157
22		5.013
23		4.796
24		4.854
25		4.721



2.1.2 Molecular optimization

The 2D structure of each of the complexes was generated by using the sketch option on Spartan 14 graphic user interphase (GUI) and was then converted into 3D structure by using the view option on Spartan 14. From the build option on Spartan 14 the structures were clean by molecular mechanic force field (MM+) option in order to remove all strain from the molecular structure. (Viswanadhan *et al.*, 1989). The resulting geometries were further refined by means of Genetic functional theory (GFT) using the B3LYP and 6-31G basis set.

2.1.3 Generation of molecular descriptors

The quantum descriptors were computed from Spartan 14 V.1.1.0 and the 3D descriptors from paDEL descriptor version 2.18.

2.1.4 Generation of training and test sets

The data set consist of 26 complexes which were splitted into 18 training set and 7 test set. The training set which made up of about 70% of the entire data set was used in building the model. The test set constitute about 30% were used to confirmed the predictive ability of the model (Tropsha, 2010).

2.1.5 Model development

Multi-linear Regression (MLR) is a strategy utilized for displaying direct relationship between a dependent variable Y (Activities) and independent variable X (atomic descriptors). The model is fit such that sum-of-squares difference between the experimental and predicted activities of dada set is minimized. In regression analysis, contingent mean of dependent variable Y relies on (descriptors)X. MLR examination extends this thought to incorporate more than one autonomous variable, regression equation takes the form

$$Y = b_1x_1 + b_2x_2 + b_3x_3 + c \quad (1)$$

Where Y is dependent variable, 'b's are regression coefficients for corresponding 'x's (independent variable), 'c' is a regression constant or intercept.

2.1.6 Model validation

Validation is the process of evaluating the fitting ability, stability, reliability and predictive ability of the developed QSAR model (Ajala *et al.*, 2016). The model was validated and the validation parameters obtained were compared with the

minimum recommended value for a generally acceptable QSAR model shown in Table 2. The capability of the QSAR equation to predict bioactivity of new compounds was determined using the leave-one-out cross-validation method. The cross-validation regression coefficient (Q_{cv}^2) was calculated with the equation below;

$$Q_{CV}^2 = 1 - PRESS/TOTAL = 1 - \frac{\sum_{i=1}^n (y_{exp} - y_{pred})^2}{\sum_{i=1}^n (y_{exp} - \bar{y})^2} \quad (2)$$

y_{pred} , y_{exp} , and \bar{y} are the predicted, experimental, and mean values of experimental activity respectively.

2.1.7 Applicability domain.

Applicability domain is a theoretical region of a chemical space, defined by the model descriptors and modeled response and thus by the nature of the training set molecules (Atkinson *et al.*, 1985). A QSAR model is employed to determine outliers and influential compounds and to affirm the reliability and robustness of the model generated (Tropsha *et al.*, 2003). Leverage is one of the techniques used in evaluating the applicability domain of a QSAR model and is given for a chemical compound as h_i :

$$h_i = x_i (X^T X)^{-1} x_i^T \quad (i=K, \dots, P) \quad (3)$$

where x_i is the training compound matrix I , X is $n \times k$ descriptor matrix of the training set compounds and X^T is the transpose matrix X used to develop the model. As a prediction tool, the warning leverage (h^*) which is the limit for X values and it's defined

$$as: h^* = 3(p+1)/n \quad (4)$$

where n is the number of training compounds, and p is the number of descriptors in the model.

Table 2 - General minimum recommended value for the evaluation of QSAR model.

Symbol	Name	Value
R^2	Co-efficient of determination	≥ 0.6
$P_{(95\%)}$	Confidence interval at 95% confidence level	< 0.05
Q^2	Cross-Validation Co-efficient	≥ 0.5
$R^2 - Q^2$	Difference between R^2 and Q^2	≤ 0.3
$N_{(\text{ext. and test set})}$	Minimum number of external and test set	≥ 05
$R^2_{\text{ext.}}$	Co-efficient of determination of external and test set	≥ 0.5

(Veerasingh et al., 2011)

$$N=26, R^2=0.8345, R^2_{\text{adj}}=0.7655, Q^2=0.6681, (R^2_{\text{pred}}) = 0.5980 = 0.21$$

3. RESULT AND DISCUSSION

A QSAR study was performed to investigate the structure activity relationship of twenty (26) complexes as potent anti-candida *albican*. The QSAR model was built using GFA algorithm and the descriptors used in building the model are reported in Table 3.

The Genetic Function Algorithm (GFA) method employed in this study led to the selection of five descriptors which were used to build a linear model for calculating predictive activity of anti-candida *albican*. Ten QSAR models were built using Multi-linear Regression technique but due to the statistical significance, model 1 was selected as the best model as shown below;

Model

$$pMIC = 4.153 (\text{MATS4p}) + 17.494 (\text{RCI}) + 48.815 (\text{G2m}) + 0.5712 (\text{B10[N-O]}) - 0.389 (\text{nF}) - 27.763$$

Table 3 -The name and symbols of descriptors used in the models

S/no	Name	Description
1	MATS4P	Moran autocorrelation of lag 4 weighed by polarizability.
2	RCI	Ring complexity index
3	G2m	2 nd component symmetry directional WHIM index/weighted by mass
4	BIO[N-O]	Presence/absence of N-O at topological distance 10
5	Nf	Number of Fluorine atoms

The experimental and predicted activities of the complexes as a potent anti-candida *albican* and the residual values were presented in Table 4. The low residual value between Experimental and Predicted activity indicates that the model has high predictive power.

The p-value is a probability that measures the evidence against the null hypothesis. The null hypothesis implies that there is no association between the descriptors and the pMIC of the complexes. The P-values of all the descriptors in the model at 95% confidence level shown in Table 5 are less than 0.05. This implies that the alternative hypothesis is accepted. Thus, there is a significant relationship between the descriptors used in the model and the pMIC of the complexes.

Lead compound 20 with best activity serve as the reference complex for designing eleven (11) complexes as reported in Table 6. The QSAR model generated was used to predict the activities of the designed complex as reported in Table 7.

Table 4 - Comparison of Experimental (pIC₅₀), Predicted (pIC₅₀) and Residual of Model 1

Complex	ObservedpMIC	PredictedpMIC	Residuals
C2	5.137	5.064	0.073
C3	5.534	5.382	0.152
C4	5.518	5.651	-0.134
C5	5.535	5.341	0.194
C6	5.299	5.345	-0.046
C7	4.778	4.842	-0.064
C10	4.425	4.425	0.000
C11	4.349	4.443	-0.094
C12	5.491	5.044	0.447
C13	5.169	5.402	-0.233
C15	5.254	5.022	0.232
C17	5.006	5.226	-0.220
C19	5.385	5.422	-0.036
C20	6.104	6.027	0.078
C22	5.013	5.179	-0.166
C23	4.796	4.710	0.085
C24	4.854	4.917	-0.063
C25	4.721	4.926	-0.205

Table 5: Contribution of the individual descriptors in the model

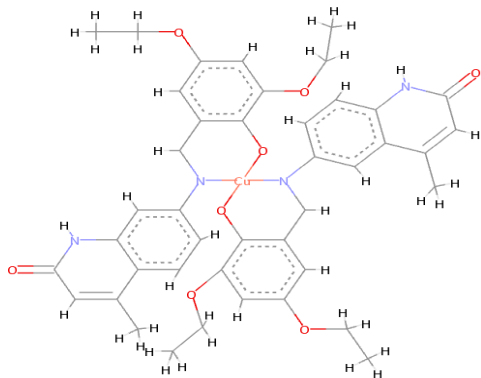
	Coefficients	Standard Error	P-value
MATS4p	4.15277	0.74145	0.00115
RCI	17.49361	4.03559	0.00097
G2m	48.81459	9.73386	0.00030
BIO[N-O]	0.57123	0.13696	0.00129
nF	-0.38917	0.11275	-0.14349

Table 6: The structure of the designed complexes

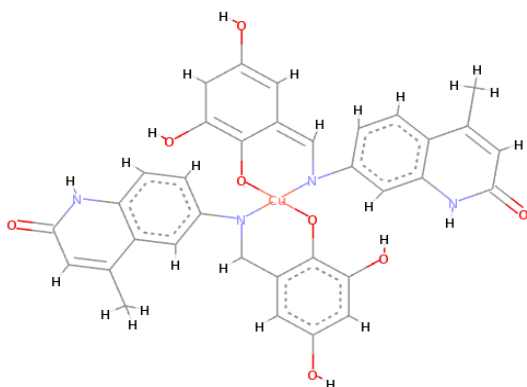
S/N	Complexes
20a	
20b	
20c	
20d	

20e	
20f	
20g	
20h	

20i



20j



20k

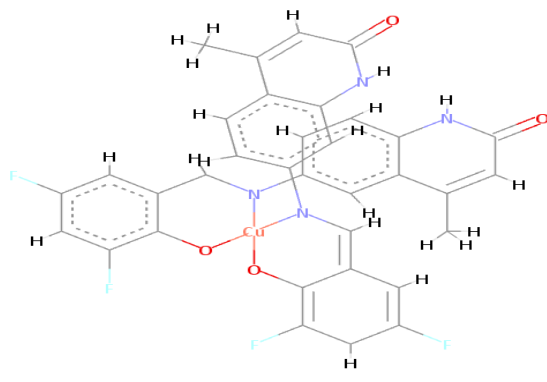


Table 7: Calculated descriptors for the designed complexes with the predicted activities

S/N	MATS4p	RCI	G2m	B10[N-O]	nF	pMIC
1b	0.089	1.457	0.185	0	0	7.126
2b	0.213	1.457	0.159	0	0	6.371
3b	0.321	1.447	0.147	0	0	6.059
4b	0.024	1.447	0.18	1	0	7.008
5b	-0.085	1.479	0.162	1	0	6.237
6b	0.005	1.454	0.157	1	0	5.929
7b	0.048	1.453	0.149	1	0	5.699
8b	-0.09	1.485	0.154	1	0	5.930
9b	-0.092	1.485	0.156	1	0	6.020
10b	-0.087	1.48	0.168	1	0	6.539
11b	-0.099	1.462	0.141	0	4	2.728

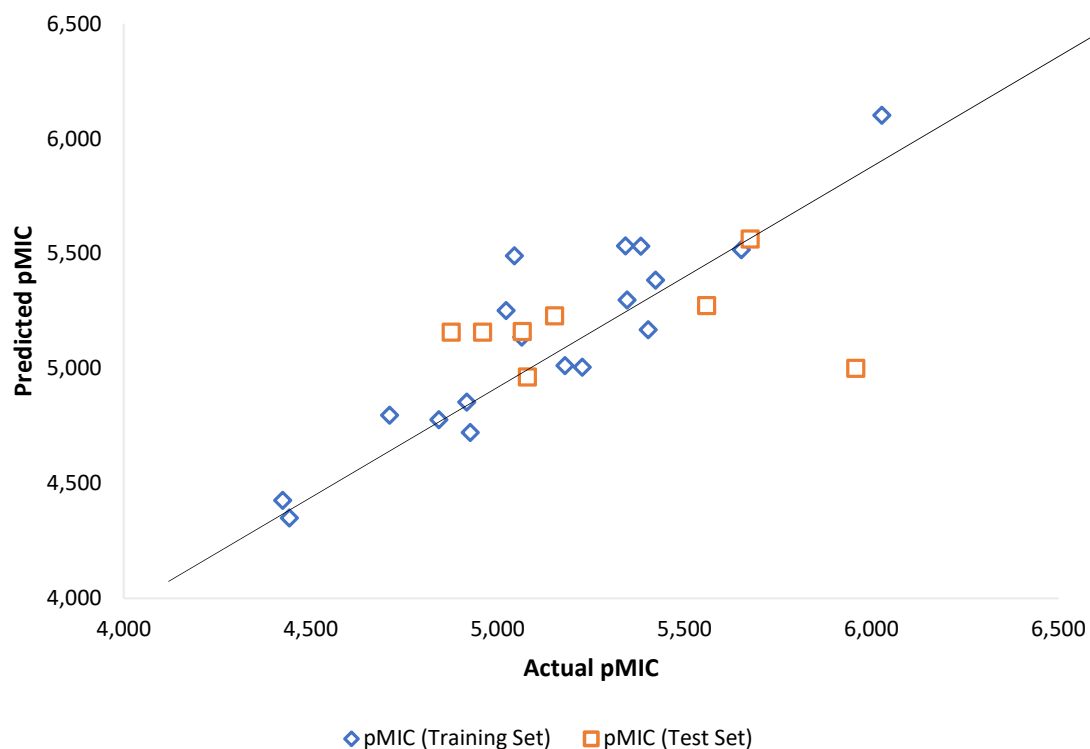


Figure 1 -The plot of the Experimental and Predicted activity of both the training and test sets of model 1.

The predicted values for pMIC of the complexes for the training and test sets were plotted against the experimental pMIC values as shown in Figure 1. The calculated values

for the pMIC are in good agreement with those of the experimental values.

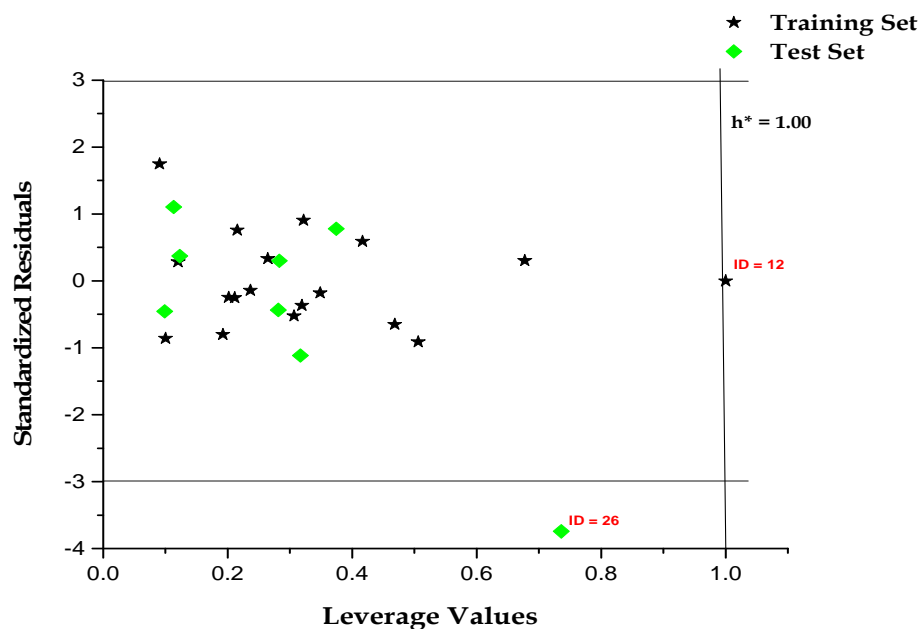


Figure 2: The Williams plot of the standardized residuals versus the leverage value

The leverages for every compound in the dataset were plotted against their standardized residuals, leading to discovery of outliers and influential chemicals in the models. The Williams plot of the standardized residuals versus the leverage value is shown in Figure 2. It is evident that all the compounds were within the square area

± 3 as shown in Figure 2. Therefore, one complex is said to be an outlier. Also, no complex is said to be an influential compound since its leverage value is greater than the warning leverage ($h^* = 1.0$)

4. CONCLUSION

QSAR study was carried out on twenty-six Copper (II) complexes, the best model obtained has $R^2 = 0.8345$, Cross Validation Squared Correlation Coefficient $Q^2 = 0.6681$ and predicted R squared (R^2_{pred}) = 0.5980. The QSAR model and the results obtained were used in designing more potent anti-candida agents. The reference complex C20 has pMIC of 6.104 and four out of the designed complexes ie 1b, 2b, 4b and 10b has pMIC of 7.126, 6.371, 7.008 and 6.539 respectively. This shows that the designed complexes are more potent than the reference complex.

Acknowledgments

The authors sincerely acknowledge TETFUND for the research grant and Prof Adamu Uzairu for his scholarly guidance.

REFERENCES

- Ajala, A., Uzairu, A., Suleiman, I.O., 2016. Chemometric study of some α , β -unsaturated ketone as potential antifungal agents using density function theory and GFA (ATCC 10231 and NCIM 3446 cell line). *Cogent Chem.* 2, 1175073.
- Atkinson, A.C. Plot, Transformation and Regression, Clarendon Press, Oxford UK (1985).
- Creaven, B.S., Devereux, M., Karcz, D., Kellett, A., McCann, M., Noble, A., Walsh, M., 2009. Copper (II) complexes of coumarin-derived Schiff bases and their anti-Candida activity. *J. Inorg. Biochem.* 103, 1196–1203.
- Creaven, B.S., Duff, B., Egan, D.A., Kavanagh, K., Rosair, G., Thangella, V.R., Walsh, M., 2010. Anticancer and antifungal activity of copper (II) complexes of quinolin-2 (1H)-one-derived Schiff bases. *Inorganica Chim. Acta* 363, 4048–4058.
- Kim, J., Sudbery, P., 2011. Candida albicans, a major human fungal pathogen. *J. Microbiol.* 49, 171.
- Kovala-Demertzi, D., 2006. Recent advances on non-steroidal anti-inflammatory drugs, NSAIDs: organotin complexes of NSAIDs. *J. Organomet. Chem.* 691, 1767–1774.
- Kumaran, J.S., Priya, S., Muthukumaran, J., Jayachandramani, N., Mahalakshmi, S., 2013. Journal of Chemical and Pharmaceutical Research, 2013, 5 (7): 56-69. *J. Chem. Pharm. Res.* 5, 56–69.
- Motta, L., Almeida, W., 2011. Quantitative structure-activity relationships (QSAR) of a series ketone derivatives as anti candida albicans. *Int. J. Drug Discov.* 3, 100–117.
- Raman, N., Joseph, J., Velan, A., Pothiraj, C., 2006. Antifungal activities of biorelevant complexes of copper (II) with biosensitive macrocyclic ligands. *Mycobiology* 34, 214–218.
- Rathod, A., 2011. Antifungal and Antibacterial activities of Imidazolylpyrimidines derivatives and their QSAR Studies under Conventional and Microwave-assisted. *Int J PharmTech Res* 3, 1942–1951.
- Tenover, F.C., 2006. Mechanisms of antimicrobial resistance in bacteria. *Am. J. Infect. Control* 34, S3–S10.
- Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, 29(6-7), 476-488.
- Tropsha, A., Gramatica, P., & Gombar, V. K. (2003). The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb Sci*, 22(1), 69-77.
- Veerasingh, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C.P., Agrawal, R.K., 2011. Validation of QSAR models-strategies and importance. *Int. J. Drug Des. Discov.* 3, 511–519.
- Viswanadhan, V. N., Ghose, A. K., Revankar, G. R., & Robins, R. K. (1989). Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *Journal of chemical information and computer sciences*, 29(3), 163-172