







## Development of Predictive Antioxidant Models for 1,3,4-Oxadiazoles by Quantitative Structure Activity Relationship

Ikechukwu Ogadimma Alisi , Adamu Uzairu , Stephen Eyije Abechi ,  
Sulaiman Ola Idris 

**Abstract:** The free radical scavenging properties of 1,3,4-oxadiazoles have been explored by the application of quantitative structure activity relationship (QSAR) studies. The entire data set of the oxadiazole derivatives were minimized and subsequently optimized at the density functional theory (DFT) level in combination with the Becke's three-parameter Lee-Yang-Parr (B3LYP) hybrid functional and 6-311G\* basis set. Kennard Stone algorithm was employed in data division into training and test sets. The training set was employed in QSAR model development by genetic function algorithm (GFA), while the test set was used to validate the developed models. The applicability domain of the developed model was accessed by the leverage approach. The variation inflation factor, degree of contribution and mean effect of each descriptor were calculated. Quantum chemical and molecular descriptors were generated for each molecule in the data set. Five predictive models that met all the requirements for acceptability with good validation results were developed. The best of the five models gave the following validation results:  $R = 0.944$ ,  $R^2 = 0.891$ ,  $Q^2(R^2_{cv}) = 0.831$ ,  $R^2_{pred} = 0.858$  and  $^cR^2_p = 0.810$ ,  $s = 0.114$ ,  $rmsep = 0.121$ . The QSAR analysis revealed that the sum of e-state descriptors of strength for potential hydrogen bonds of path length 9 (SHBint9) and topological radius (topoRadius) are the most crucial descriptors that influence the free radical scavenging activities of 1,3,4-oxadiazole derivatives.

**Keywords:** Antioxidants, 1,3,4-Oxadiazoles, Descriptors, model validation, QSAR.

**Submitted:** March 15, 2018 . **Accepted:** November 27, 2018.

**Cite this:** Alisi I, Uzairu A, Abechi S, Idris S. Development of Predictive Antioxidant Models for 1,3,4-Oxadiazoles by Quantitative Structure Activity Relationship. JOTCSA. 2019;6(2):103-14.

**DOI:** <https://dx.doi.org/10.18596/jotcsa.460207>.

**\*Corresponding author. E-mail:** [ikeogadialisi@gmail.com](mailto:ikeogadialisi@gmail.com).

### INTRODUCTION

Antioxidants are molecular systems that inhibit or delay the oxidation of other molecules to generate free radicals by employing various mechanisms in their mode of action (1). In biological systems to be precise, antioxidants are involved in the control of oxidative chemical processes. Natural sources of antioxidants for the human system include: ascorbic acid, thyme, pepper, nutmeg and tea. Free radicals are molecules or their fragments that contain one or more unpaired electrons in their atomic or molecular orbitals (2). The considerable level of reactivity associated with free radicals is due to the presence of these unpaired electrons.

Various methods have been developed to measure the efficacy of antioxidants either as pure compounds or in extracts. The methods

employed to screen compounds for their free radical scavenging activities include: nitric oxide radical ( $\text{NO}^*$ ) scavenging assay (3), superoxide anion radical ( $\text{O}_2^{\cdot-}$ ) scavenging assay (4), hydroxyl radical ( $\text{OH}^*$ ) scavenging assay (3,4) hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) scavenging assay (3), 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonate) cationic radical ( $\text{ABTS}^{\cdot+}$ ) scavenging assay (4, 5, 6), and ferric reducing antioxidant power (FRAP) methods (6). The most accepted method employed in screening the free radical scavenging activities of antioxidants is the DPPH (2,2'-diphenyl-1-picrylhydrazyl) radical scavenging assay (3-7). This method involves a change in color from deep-violet to light-yellow due to the reduction of the stable free radical DPPH to the reduced form DPPH-H (2,2'-diphenyl-1-picrylhydrazine) (8). The radical scavenge examined in this research is based on the DPPH method.

Normal cellular metabolism in living systems results in the production of reactive oxygen species (ROS) (superoxide anion radical, the hydroxyl radical, etc.) and reactive nitrogen species (RNS) (nitric oxide radical, peroxyxynitrite, etc.). External factors that induce free radical generation in the body include the penetration of the skin by ultraviolet light, and polluted air with high levels of smog (9). Free radicals at moderate levels are essential in maintaining the physiological processes in the body. High levels of ROS and RNS in the body are harmful and may cause damage to proteins and other cell structures which results to oxidative stress and nitrosative stress respectively (2). The body can overcome these adverse effects by an efficient antioxidant protection system.

The 1,3,4-oxadiazole derivatives have been studied extensively and found to possess antioxidant activities (6, 10,11). They are a class of heterocyclic compounds that contain one oxygen and two nitrogen atoms in a five membered ring system.

A method widely employed to correlate the relationship between molecular structure and biological activities of compounds is quantitative structure activity relationship (QSAR) (12-16). In this research, QSAR through genetic function algorithm was employed to investigation the antioxidant properties of 1,3,4 oxadiazole derivatives. The data set of the oxadiazoles and their IC<sub>50</sub> antioxidant activities were generated from the literature (6, 10, 11). The molecular structures were initially subjected to energy minimization and subsequently optimized at the density functional theory (DFT) level. The entire data set were divided into training and test sets by Kennard Stone algorithm (KSA). The training set was used to develop the QSAR models by genetic function algorithm (GFA), while the test set was employed in model validation. The variation inflation factor, mean effect and degree of contribution of each descriptor in the developed model were all calculated. Also, the applicability domain of the developed model was accessed by the leverage approach.

## COMPUTATIONAL METHODS

### Data set generation

The data set of 50 derivatives of 1,3,4-oxadiazole and their IC<sub>50</sub> antioxidant activities based on DPPH free radical scavenging assay was sourced from literature (6, 10, 11). The antioxidant activities were converted to uniform unit values in  $\mu\text{g/mL}$  and subsequently to their  $pIC_{50}$  values.

### Geometry optimization / Descriptors calculation

ChemDraw software (17) was utilized in drawing the chemical structures of the compounds. The minimization of the molecular geometries and their subsequent optimization were executed with

the Spartan 14 program (18) at the density functional theory (DFT) level using Becke's three-parameter Lee-Yang-Parr hybrid functional (B3LYP) in combination with the 6-311G\* basis set without symmetry constraints (19). This operation also resulted in the generation of quantum chemical descriptors. The PaDEL program package version 2.20 (20), was employed in the generation of molecular descriptors for the optimized molecular structures.

### Data pre-treatment, normalization and division

The program "Data Pre-Treatment GUI 1.2" (21, 22), was employed in data pre-treatment in which constant value descriptors and pairs of descriptors whose correlation coefficients are greater than 0.9 were removed. Also, in order to ensure that no particular descriptor dominates the model due to very high or low values, data normalization was conducted by scaling between the interval  $N(0,1)$  (23, 24). A rational selection of training and test sets from the data set was carried out using the program "Dataset Division GUI 1.2" tool (25), by employing the Kennard Stone algorithm technique. This technique is a classical selection method that selects a representative subset from a pool of  $N$  samples that employs the Euclidean distances  $d_x(y, z)$  between the  $\mathbf{x}$ -vectors of each pair  $(y, z)$  of samples.

### Development and validation of the QSAR models

The training set was used to develop the QSAR model in which the independent variables (descriptors) and the dependent variables ( $pIC_{50}$ ) were subjected to multivariate analysis by Genetic Function Approximation (GFA) using the material studio software. During the model development, 50,000 crossovers, a smoothness value of 1.00 and other default settings were used for each combination.

The internal stability and predictive ability of the developed models were tested by subjecting them to internal validation by leave- one- out (LOO) cross- validation. Various internal validation metrics such as the correlation coefficient,  $R$ ; the cross-validated squared correlation coefficient,  $R_{cv}^2(Q^2)$ ; the adjusted  $R^2$  ( $R_a^2$ ); the variance ratio,  $F$  and the standard error of estimate,  $s$  were calculated (12,16,26-28).

The external validation metrics calculated include the following:

The predictive  $R^2$  ( $R^2_{pred}$ ) which is the predicted correlation coefficient calculated from the predicted activity of all the test set compounds was calculated as presented in Equation 1.

$$R^2_{pred} = 1 - \frac{\sum(Y_{pred(Test)} - Y_{(Test)})^2}{\sum(Y_{(Test)} - \bar{Y}_{(Training)})^2} \quad (1)$$

Where  $Y_{pred(Test)}$  and  $Y_{(Test)}$  are the predicted and observed activity values of the test set compounds.  $\bar{Y}_{(Training)}$  is the mean activity value of the training set.

Due to the dependence of  $R^2_{pred}$  on  $\sum(Y_{(Test)} - \bar{Y}_{(Training)})^2$  its value may not reflect the predictive potential of the developed QSAR model (29) For this reason, a modified  $R^2$  called  $r^2_m$  is thus introduced (30), as presented in equation 2.

$$r^2_m = r^2 \left(1 - \sqrt{r^2 - r_0^2}\right) \quad (2)$$

A plot of observed activities against the corresponding predicted values has slope equal to  $k$  and squared correlation coefficient  $r_0^2$  with intercept set to zero. When the intercept is not set to zero, the squared correlation coefficient  $r^2$  is obtained. The interchange of the axis gives the parameter  $r_m'^2$  (Equation 3) (24, 31-33).

$$r_m'^2 = r^2 \times \left(1 - \sqrt{r^2 - r_0'^2}\right) \quad (3)$$

The Y-randomization test was employed in checking the robustness of the models by permuting the activity values against the matrix of the descriptors (34). The robustness of the model is reflected in the value of  ${}^cR_p^2$  parameter computed as presented in equation 4.

$${}^cR_p^2 = R \times \sqrt{R^2 - R_r^2} \quad (4)$$

#### Calculation of the variation inflation factor parameter

The multi-collinearity, among the descriptors in the developed model was estimated using equation (5).

$$VIF = \frac{1}{1 - r^2} \quad (5)$$

Where  $VIF$  is the variation inflation factor, while  $r$  is the correlation coefficient of multiple regressions of a given descriptor with the other descriptors.

#### Mean effect and degree of contribution estimates

The mean effect (MF) which signifies the relative significance of each descriptor in the model, was estimated using Equation 6.

$$MF_j = \frac{\beta_j \sum_{i=1}^n d_{ij}}{\sum_j \beta_j \sum_1^n d_{ij}} \quad (6)$$

Where  $MF_j$  = mean effect for the considered descriptor  $j$ .

$\beta_j$  = coefficient of the descriptor  $j$

$d_{ij}$  = value of the target descriptors for each molecule,

$m$  = number of descriptors in the model.

The Degree of Contribution (DC), which is the standardized regression coefficient was also calculated for each descriptor.

#### Applicability domain investigation

The applicability domain of a model is the range of compounds within which the model makes predictions with a given reliability (35).

The applicability domain of the developed 1,3,4-oxadiazoles antioxidants' QSAR model was accessed by the leverage approach (36). The leverage value for all compounds in the dataset  $X$  were calculated from the hat matrix (H) (equation 7).

$$H = X(X^T X)^{-1} X^T \quad (7)$$

Where  $X$  is the two-dimensional  $n \times k$  descriptor matrix of the training set compounds that contains  $n$  compounds and  $k$  descriptors, The transpose of  $X$  is reoresented as  $X^T$ . The leverage value of the  $i$ th compound ( $h_i$ ) was estimated using equation 8.

$$h_i = x_i(X^T X)^{-1} x_i^T \quad (i = 1, \dots, m) \quad (8)$$

The cut-off leverage,  $h^*$ , which is the limit of normal values for  $X$  outliers, was also calculated as presented in equation 9 (37).

$$h^* = \frac{3(k+1)}{n} \quad (9)$$

For the estimation of the standard residuals, equation 10 was employed.

$$\text{Standard Residual} = \frac{\text{Residual}}{\text{RMSE}} \quad (10)$$

Where  $RMSE$  is the Root Mean Square Error.

The Williams plot was obtained by plotting the standard residuals against the leverage values. (38).

## RESULTS AND DISCUSSION

#### Data set generation and pre-treatment

The chemical name of the entire data set together with their corresponding  $IC_{50}$  ( $\mu\text{g}/\text{mL}$ ) antioxidant activities are presented in Table 1. A total of 32 quantum chemical descriptors were obtained upon geometry optimization of the molecular structures. Also 1875 molecular descriptors were generated after descriptors calculation. These computations gave a total of 1907 descriptors. The resulting descriptors can broadly be categorized into electronic, spatial, structural, thermodynamic, topological, constitutional, geometrical, RDF and 3D-Morse. The entire data set upon feature processing produced 1181 descriptors. Recall that data pre-treatment ensures that the predictivity and generalization of the resulting model does not fail under the prevailing conditions. The resulting data set was then normalized and upon data division, gave 40 training set compounds and 10 test set compounds.

**Analysis of the results of developed models and their validation**

A total of five QSAR models were developed by genetic function approximation. These models are represented in Table 2. These descriptors are

**Table 2:** Developed models for 1,3,4-oxadiazole antioxidant derivatives by genetic function approximation

Model No	Equation
1	$pIC_{50} = 0.514 * ATSC7v - 2.505 * SpAbs\_Dzv + 0.644 * SHBint9 + 1.800 * topoRadius - 0.334 * TDB9m + 0.448 * RDF85e + 4.024$
2	$pIC_{50} = 0.635 * ATSC7v - 2.119 * SpAbs\_Dzv + 0.341 * VE3\_Dzv + 0.338 * SHBint9 + 1.324 * topoRadius + 0.468 * RDF85e + 3.813$
3	$pIC_{50} = 0.531 * ATSC7v - 2.494 * SpAbs\_Dzv + 0.564 * SHBint9 + 1.755 * topoRadius - 0.368 * TDB9m + 0.401 * RDF75m + 0.390 * RDF85e + 3.977$
4	$pIC_{50} = 0.534 * ATSC7v - 2.368 * SpAbs\_Dzv + 0.485 * nHBint9 + 1.464 * topoRadius - 0.291 * TDB9m + 0.556 * RDF85e + 4.017$
5	$pIC_{50} = 0.661 * ATSC7v - 0.236 * GATS5i - 2.038 * SpAbs\_Dzp + 0.442 * nHBint9 + 1.079 * topoRadius + 0.633 * RDF85e + 4.058$

A summary of the internal validation results of the five models is presented in Table 3. All the five models met the requirements for model acceptability based on the results of their internal validation. For instance, for models 1, 2, 3, 4 and 5 their ( $R^2 - R_a^2$ ) values are 0.019885, 0.019963, 0.020424, 0.020266 and 0.020297, respectively. These results are by far lower than the threshold value of 0.3. Thus the number of descriptors in the developed models are within the acceptable range since the value of ( $R^2 - R_a^2$ ) for each model

observed to fall within the categories of autocorrelation, barysz matrix, electrotopological state atom type, topologica and radial distribution function descriptors.

is less than 0.3. Also their  $R^2$  values are well above the threshold limit of 0.6. Model 3 has the highest  $R^2$  and  $R_a^2$  values of 0.907 and 0.886 respectively. This model also has the lowest standard error value of 0.107. While model 5 has the lowest  $R^2$  and  $R_a^2$  values of 0.888 and 0.868 respectively. Also, model 2 has the highest  $R_{cv}^2(Q^2)$  value of 0.840 while model 3 has the lowest value of 0.807.

**Table 3:** Summary of internal validation results for 1,3,4-oxadiazole antioxidant derivatives.

Parameters	Equation 1	Equation 2	Equation 3	Equation 4	Equation 5
Friedman LOF	0.023	0.024	0.024	0.024	0.024
R-squared	0.891	0.890	0.907	0.889	0.888
Adjusted R-squared	0.871	0.870	0.886	0.868	0.868
Cross validated R-squared	0.831	0.840	0.807	0.815	0.834
Significant Regression	Yes	Yes	Yes	Yes	Yes
Significance-of-regression F-value	44.790	44.590	44.390	43.840	43.770
Critical SOR F-value (95%)	2.404	2.404	2.321	2.404	2.404
Replicate points	0.000	0.000	0.000	0.000	0.000
Computed experimental error	0.000	0.000	0.000	0.000	0.000
Lack-of-fit points	33.000	33.000	32.000	33.000	33.000
Min expt. error for non-significant LOF (95%)	0.095	0.096	0.089	0.096	0.096
Standard Error of Estimate	0.114	0.115	0.107	0.115	0.116

\*The criteria for model acceptability is:  $R^2 \geq 0.6$

From the results of Y- randomization tests as presented in Table 4, all the models gave encouraging results. The five models passed the tests for model acceptability with model parameter results well above the threshold values :  $R \geq 0.8$  ,  $R^2 \geq 0.6$  ,  $Q^2 > 0.5$  ,  $cR_p^2 \geq 0.5$  (39). Model 2 recorded the highest  $cR_p^2$  value of 0.822,

while model 3 has the lowest value of 0.789. Thus model 2 is the most robust of the five models. Recall that based on the Y-randomization test, the predictive power of a model which is reflected on the value of  $cR_p^2$  must be greater than or equal to 0.5 (40).

**Table 4:** Results of Y-randomization for 1,3,4-oxadiazole antioxidant derivatives.

Parameters	Model 1	Model 2	Model 3	Model 4	Model 5
$R$	0.944	0.944	0.952	0.943	0.943
$R^2$	0.891	0.890	0.907	0.889	0.888
$Q^2$	0.831	0.840	0.807	0.815	0.834
Random Model Parameters					
Average $r$	0.392	0.363	0.469	0.405	0.374
Average $r^2$	0.167	0.141	0.230	0.175	0.145
Average $Q^2$	-0.590	-0.560	-0.420	-0.280	-0.400
$cR_p^2$	0.810	0.822	0.789	0.802	0.815

\*Model acceptability criteria:  $R \geq 0.8$ ,  $R^2 \geq 0.6$ ,  $Q^2 > 0.5$ ,  $cR_p^2 \geq 0.5$  (39).

**Table 5:** External validation results for 1,3,4-oxadiazole antioxidant derivatives.

Validation Parameters	Model 1	Model 2	Model 3	Model 4	Model 5
$r^2$	0.829	0.774	0.760	0.723	0.682
$r_0^2$	0.821	0.745	0.695	0.686	0.606
Reverse $r_0^2$	0.822	0.769	0.760	0.710	0.676
$r_m^2$	0.752	0.641	0.566	0.585	0.494
Reverse $r_m^2$	0.759	0.719	0.749	0.643	0.627
Average $r_m^2$	0.755	0.68	0.658	0.614	0.561
Delta $r_m^2$	0.007	0.078	0.182	0.058	0.133
$r^2 - r_0^2/r^2$	0.010	0.038	0.085	0.050	0.112
$r^2 - r_0'^2/r^2$	0.009	0.007	3E-04	0.017	0.010
$k$	0.994	0.996	0.987	0.988	0.993
$k'$	1.005	1.003	1.012	1.011	1.006
$ r_0^2 - r_0'^2 $	0.001	0.024	0.065	0.024	0.070
rmsep	0.121	0.142	0.166	0.165	0.178
$R_{pred}^2$	0.858	0.805	0.735	0.737	0.694

The acceptable threshold values for the given parameters are as follows:

$R_{pred}^2 > 0.5$ ,  $r^2 > 0.6$ ,  $r_m^2 \geq 0.5$ , Delta  $r_m^2 < 0.2$ ,  $|r_0^2 - r_0'^2| < 0.3$ ,  $(r^2 - r_0^2)/r^2 < 0.1$  and  $0.85 \leq k \leq 1.15$ , or  $(r^2 - r_0'^2)/r^2 < 0.1$  and  $0.85 \leq k' \leq 1.15$  (24).

The external validation results for the developed models are presented in Table 5. These developed models passed all the Golbraikh and Tropsha criteria for acceptability, as the results of their external validation are within the recommended threshold values as presented in Table 5.

The computed predicted activities and residual values for the training set and test set compounds are also presented in Table 1. These results indicate very small residual values for each compound. Also the plot of predicted activities against experimental activities for the training set are presented in Figure 1. While, that of the test set are presented in Figure 2.

The acceptable threshold values for the given parameters are as follows:

$R_{pred}^2 > 0.5$ ,  $r^2 > 0.6$ ,  $r_m^2 \geq 0.5$ , Delta  $r_m^2 < 0.2$ ,  $|r_0^2 - r_0'^2| < 0.3$ ,  $(r^2 - r_0^2)/r^2 < 0.1$  and  $0.85 \leq k \leq 1.15$ , or  $(r^2 - r_0'^2)/r^2 < 0.1$  and  $0.85 \leq k' \leq 1.15$  (24).

From the results of Y-randomization tests as presented in Table 4, all the models gave encouraging results. The five models passed the tests for model acceptability with model parameter results well above the threshold values:  $R \geq 0.8$ ,  $R^2 \geq 0.6$ ,  $Q^2 > 0.5$ ,  $cR_p^2 \geq 0.5$  (39). Model 2 recorded the highest  $cR_p^2$  value of 0.822,

while model 3 has the lowest value of 0.789. Thus model 2 is the most robust of the five models. Recall that based on the Y-randomization test, the predictive power of a model which is reflected on the value of  $cR_p^2$  must be greater than or equal to 0.5 (40).

The external validation results for the developed models are presented in Table 5. These developed models passed all the Golbraikh and Tropsha criteria for acceptability, as the results of their external validation are within the recommended threshold values as presented in Table 5.

The computed predicted activities and residual values for the training set and test set compounds are also presented in Table 1. These results indicate very small residual values for each compound. Also the plot of predicted activities against experimental activities for the training set are presented in Figure 1. While, that of the test set are presented in Figure 2.

Model 1 has the highest  $R_{pred}^2$  value of 0.858, the highest  $r^2$  value of 0.829, the lowest Delta  $r_m^2$  value of 0.007 and the lowest rmsep value of 0.121. Recall that the higher the  $R_{pred}^2$  and  $r^2$  values in combination with very low Delta  $r_m^2$  and rmsep values, the better the model. Since model 1 possess the best results based on these

standards, in combination with the good internal validation results, it is hereby recognized as the best of the five developed models. Thus, the predicted activities and residual values for the entire data set as presented in Tables 1 and the plots of predicted activities against the experimental activities as presented in Figures 1 and 2 are generated from model 1 results.

This model is given by:

$$pIC_{50} = 0.514 * ATSC7v - 2.505 * SpAbs\_Dzv + 0.644 * SHBint9 + 1.800 * topoRadius - 0.334 * TDB9m + 0.448 * RDF85e + 4.024$$

$$R = 0.944, R^2 = 0.891, Q^2(R^2_{cv}) = 0.831, R^2_{pred} = 0.858 \text{ and } cR^2_p = 0.810 \quad s = 0.114, \text{ rmsep} = 0.121$$

### Analysis of applicability domain results

The results of the applicability domain of the developed model for the training and test set compounds is presented in the William's plot (Figure 3). This is a plot of the standard residual for each molecule against its corresponding leverage value. In this plot, the applicability domain for the developed model was established inside a squared area within  $\pm 2.5$  bound for residuals and a warning leverage,  $h^*$  value of

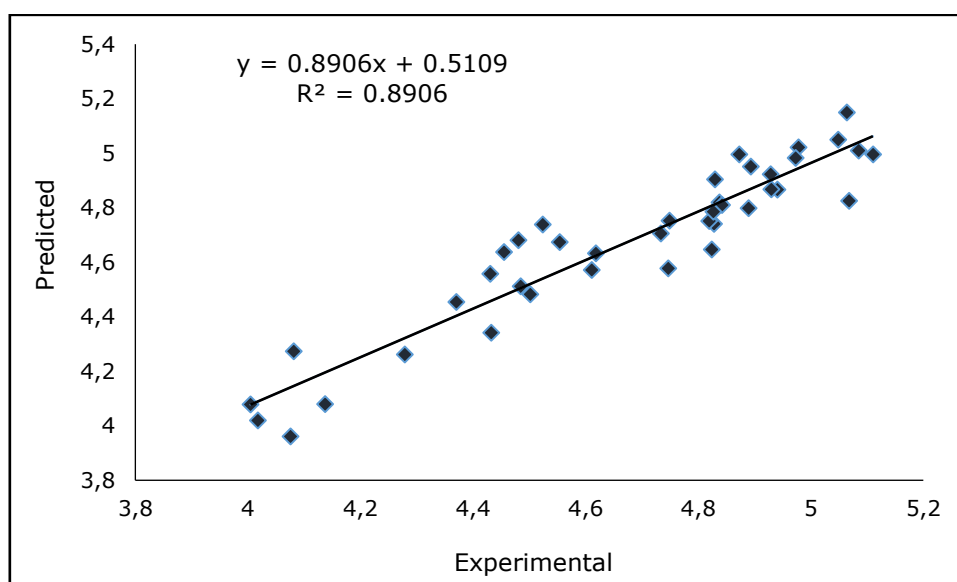
0.525. From this plot, no response outliers were identified for both the training set and test set compounds as all the compounds were found within the limit of  $\pm 2.5$  standard deviation units. In terms of structural outliers, only one training set compound (M30) and two test set compounds (M37 and M42) were identified. A close observation on the molecular structure of these compounds reveal that these structural outliers have peculiar molecular structure. For M30 a long alkyl group is attached to the terminal phenyl ring. While M37 and M42 are the only molecules with two methyl groups linked to the amino group attached to this terminal phenyl ring. These molecules are recognized as extrapolations from the structural domain of the model since their  $h$  values are greater than  $h^*$  (0.525) (41). Thus prediction for these compounds are considered unreliable. Based on these results, the applicability domain of the 1,3,4-oxadiazole antioxidant derivatives is thus established.

### Molecular descriptors and their significance

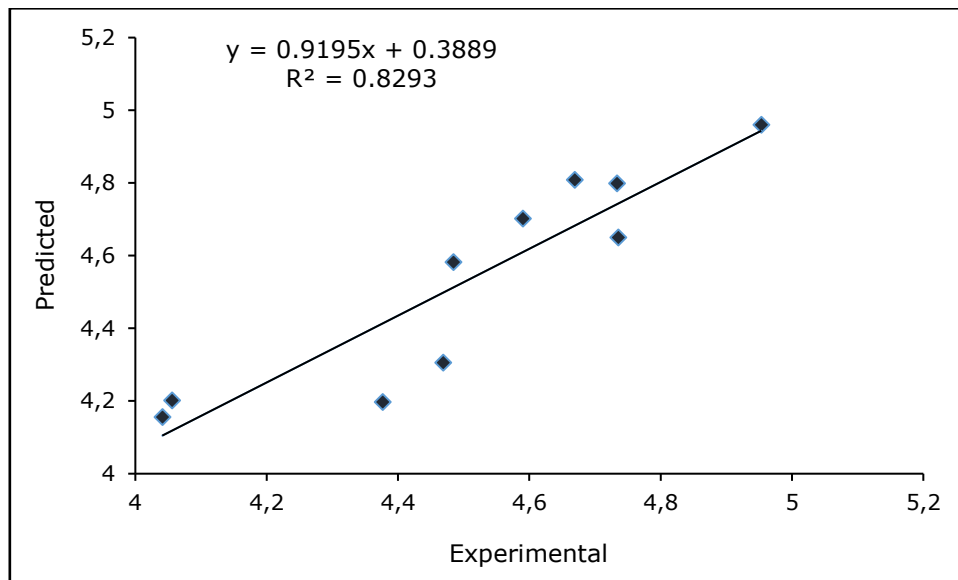
The computed results of Coefficient, Standard Error, Mean Effect, Variation Inflation Factor and Degree of Contribution of the Descriptors for the oxadiazole antioxidants are presented in Table 6.

**Table 6:** Specifications of coefficient, standard error, mean effect, variation inflation factor and degree of contribution of the descriptors for 1,3,4-oxadiazole antioxidants.

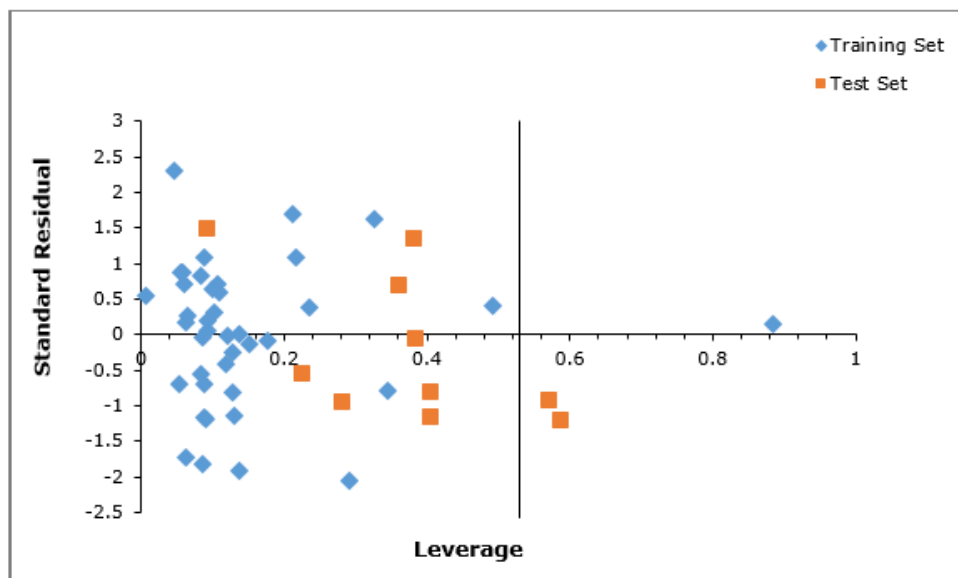
Descriptor	Coefficient	Standard Error	P-Value	DC	MF	VIF
ATSC7v	0.514	0.106	2.96E-05	4.839	0.906	1.512
SpAbs_Dzv	-2.505	0.279	2.14E-10	-8.995	-4.417	6.359
SHBint9	0.645	0.094	7.52E-08	6.876	1.136	3.908
topoRadius	1.800	0.272	1.65E-07	6.605	3.173	5.774
TDB9m	-0.334	0.113	0.006	-2.967	-0.589	1.915
RDF85e	0.448	0.119	0.001	3.751	0.790	2.264



**Figure 1:** Graphical plot of experimental activities against predicted activities for training set of 1,3,4-oxadiazole antioxidants.



**Figure 2:** Graphical plot of experimental activities against predicted activities for test set of oxadiazole antioxidants.



**Figure 3:** William's Plot for Oxadiazole Antioxidant Derivatives

**Table 1:** Data set chemical name and their  $IC_{50}$  and  $pIC_{50}$  Values

Comp No	Compounds	$IC_{50}$	$pIC_{50}$		
			Observed	Predicted	Residual
M01*	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-((phenylamino)methyl)-1,3,4-oxadiazole-2(3H)-thione	11.140	4.953	4.960	-0.007
M02	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((2-fluorophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	13.401	4.873	4.996	-0.123
M03	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((3-fluorophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	12.780	4.893	4.951	-0.058
M04	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((4-fluorophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	10.519	4.978	5.021	-0.043
M05	3-(((2,6-difluorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	8.2267	5.085	5.009	0.075
M06	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((3,4,5-trifluorophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	7.757	5.11	4.996	0.115
M07	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((2-(trifluoromethyl)phenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	27.945	4.554	4.673	-0.120
M08	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((3-(trifluoromethyl)phenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	14.541	4.837	4.819	0.018
M09	3-(((3,5-bis(trifluoromethyl)phenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	42.630	4.370	4.453	-0.083
M10	3-(((2-chlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	11.463	4.941	4.866	0.074
M11	3-(((3-chlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	8.5578	5.068	4.825	0.242
M12	3-(((4-chlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	11.764	4.929	4.868	0.062
M13	3-(((2,4-dichlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	17.827	4.749	4.752	-0.003
M14	3-(((2,5-dichlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	14.384	4.842	4.810	0.032
M15	3-(((3,4-dichlorophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	14.856	4.828	4.741	0.088
M16	3-(((2-bromophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	15.156	4.819	4.751	0.068
M17*	3-(((3-bromophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	21.418	4.669	4.809	-0.14
M18	3-(((4-bromophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	29.962	4.523	4.738	-0.215
M19	3-(((2,4-dibromophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	17.931	4.746	4.577	0.170
M20*	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((2-nitrophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	18.477	4.733	4.800	-0.066
M21	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((3-nitrophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	14.803	4.830	4.903	-0.074
M22	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((4-nitrophenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	18.477	4.733	4.705	0.028
M23	3-(((2-chloro-4-nitrophenyl)amino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	35.076	4.455	4.637	-0.182
M24	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((o-tolylamino)methyl)-1,3,4-oxadiazole-2(3H)-thione	12.902	4.889	4.798	0.091
M25	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((p-tolylamino)methyl)-1,3,4-oxadiazole-2(3H)-thione	8.9459	5.048	5.049	-4.17E-04
M26	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((2-methoxyphenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	10.637	4.973	4.982	-0.009
M27	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((4-methoxyphenyl)amino)methyl)-1,3,4-oxadiazole-2(3H)-thione	11.792	4.928	4.922	0.006



M28	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-(((2-ethoxyphenyl)amino) methyl)-1,3,4-oxadiazole-2(3H)-thione	8.6339	5.064	5.149	-0.086
M29*	3-((cyclohexylamino)methyl)-5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-1,3,4-oxadiazole-2(3H)-thione	32.798	4.484	4.583	-0.099
M30	5-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-3-((hexadecylamino)methyl)-1,3,4-oxadiazole-2(3H)-thione	52.680	4.278	4.262	0.017
M31	(E)-5-((2-((2,4-dimethoxybenzylidene)amino)-5-methylthiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	14.900	4.827	4.785	0.042
M32	(E)-5-((2-((benzo[d][1,3]dioxol-5-ylmethylene)amino)-5-methylthiazol-4-yl)methyl)-1,3,4-oxadiazole-2-thiol	32.800	4.484	4.511	-0.027
M33*	(E)-5-((2-((4-methoxybenzylidene)amino)-5-methylthiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	18.400	4.735	4.651	0.084
M34	(E)-5-((2-((4-isopropylbenzylidene)amino)-5-methylthiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	24.100	4.618	4.632	-0.014
M35	(E)-4-(((4-((5-mercapto-1,3,4-oxadiazol-2-yl)methyl)-5-methylthiazol-2-yl)imino)methyl)-2-methoxyphenol	15.000	4.824	4.647	0.177
M36	(E)-2-(4-(((4-((5-mercapto-1,3,4-oxadiazol-2-yl)methyl)-5-methylthiazol-2-yl)imino)methyl)phenyl)ethanol	24.500	4.611	4.571	0.039
M37*	(E)-5-((2-((4-(dimethylamino)benzylidene)amino)-5-methylthiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	25.700	4.59	4.703	-0.112
M38	(E)-5-((2-((furan-3-ylmethylene)amino)-5-methylthiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	37.100	4.431	4.557	-0.126
M39	(E)-5-((5-methyl-2-((2-nitrobenzylidene)amino)thiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	31.500	4.502	4.482	0.019
M40	(E)-5-((5-methyl-2-((pyridine-2-ylmethylene)amino)thiazol-4-yl) methyl)-1,3,4-oxadiazole-2-thiol	33.100	4.480	4.680	-0.200
M41	1-(2-phenyl-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	73.000	4.137	4.079	0.058
M42*	1-(2-(4-(dimethylamino)phenyl-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	88.000	4.056	4.202	-0.147
M43	1-(2-(2-chlorophenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	99.000	4.004	4.078	-0.074
M44	1-(2-(2-nitrophenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	84.000	4.076	3.961	0.115
M45*	1-(2-(4-methoxyphenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl) ethanone	42.000	4.377	4.198	0.179
M46	1-(2-(3-chloro-4-nitrophenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl) ethanone	96.000	4.018	4.019	-0.002
M47*	1-(2-(3-nitrophenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	91.000	4.041	4.156	-0.115
M48	1-(2-(4-hydroxy-3-methoxyphenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	83.000	4.081	4.273	-0.192
M49	1-(5-(pyridine-3-yl)-2-(p-tolyl)-1,3,4-oxadiazol-3(2H)-yl)ethanone	37.000	4.432	4.341	0.091
M50*	1-(2-(3,4-dimethoxyphenyl)-5-(pyridine-3-yl)-1,3,4-oxadiazol-3(2H)-yl) ethanone	34.000	4.469	4.306	0.162

\*Test Set

From Table 6, we observe that the minimum and maximum VIF values are 1.512 and 6.359 for the descriptors ATSC7v and SpAbs\_Dzv respectively. These results are within the acceptable limits, hence the developed model should be accepted. Recall that there is no inter-correlation among the descriptors when  $VIF = 1$ . The model is acceptable when the  $VIF$  result lies in the range 1 – 5. Also, if  $VIF > 10$ , the developed model is unstable (42).

**SHBint9** (Sum of E-State descriptors of strength for potential hydrogen bonds of path length 9). This descriptor possess the highest DC value of 6.876, the second largest MF value of 1.136 in addition to the smallest standard error value of 0.094. For a given descriptor, the higher the MF and DC values, the higher the relative significance and contribution of that descriptor in the developed model. To substantiate these results, this descriptor is positively correlated with the antioxidant activities of the oxadiazole derivatives. These results imply that an increase in the strength for potential hydrogen bonds of path length 9 on the oxadiazole moiety improve their activity. This descriptor is hereby recognized as the most influential descriptor for the 1,3,4-oxadiazoles. Thus in the design of new set of oxadiazole antioxidants, emphasis should be laid on this descriptor.

**topoRadius:** Topological radius. This descriptor for a given molecule, implies the minimum atom eccentricity of that molecule. This descriptor has the highest MF and coefficient values of 3.173 and 1.800 with a very encouraging value for DC of 6.605. Also it is positively correlated with the antioxidant activities of the oxadiazoles. Thus the larger the value of the minimum atom eccentricity among the oxadiazoles, the higher their activities.

**TDB9m** (3D topological distance based autocorrelation - lag 9 / weighted by mass). This is a 3D descriptor in which the topological distance between pairs of atoms is 9 with a weighting component of relative atomic mass. This descriptor is negatively correlated with the antioxidant activities of the 1,3,4-oxadiazoles with very low values of -2.967 and -0.589 for DC and MF respectively.

**SpAbs\_Dzv:** Graph energy from Barysz matrix / weighted by van der Waals volumes. As can be seen in Table 6, these two descriptors possess the lowest results for coefficient, DC and MF. In addition, they are negatively correlated with the antioxidant activities of the oxadiazole antioxidants. These results portray the low contribution and influence of these descriptors on the antioxidant activities of the oxadiazoles.

**RDF85e** (Radial distribution function - 085 / weighted by relative Sanderson electronegativities). This is a 3D descriptor that describes the probability distribution of locating an atom in a spherical volume of radius  $r$  in which

the associated weighing scheme is the Sanderson electronegativity. In our developed model, this descriptor is positively correlated with the antioxidant activities of the oxadiazoles with a coefficient of 0.448. As can be observed from Table 6, the DC value for this descriptor is 3.751, while the MF value is 0.790. These results suggest that the addition of electronegative entities improve the free radical scavenging activities of the oxadiazoles.

**ATSC7v** (Centered Broto-Moreau autocorrelation - lag 7 / weighted by van der Waals volumes) This descriptor describes the distribution of van der Waals volume with a lag of 7 along the topological structure of the 1,3,4-oxadiazole molecule. Recall that the lag is the topological distance between pairs of atoms. In the developed model, the ATSC7v descriptor has moderate values for DC and MF, in addition of being positively correlated with the antioxidant activity of the 1,3,4-oxadiazoles.

## CONCLUSION

Quantitative structure activity relationship (QSAR) by genetic function algorithm (GFA) was successfully employed to investigate the antioxidant properties of 1,3,4-oxadiazoles. Five QSAR models were developed and subjected to various validation tests. All the developed models passed the various validation tests with highly encouraging results. The best of the five developed models which is represented as:  $pIC_{50} = 0.514 * ATSC7v - 2.505 * SpAbs\_Dzv + 0.644 * SHBint9 + 1.800 * topoRadius - 0.334 * TDB9m + 0.448 * RDF85e + 4.024$  is a reference basis in the design of new antioxidants with 1,3,4-oxadiazole moiety.

The results of mean effect, degree of contribution and variation inflation factor show that the most important descriptors that influence the free radical scavenge of 1,3,4-oxadiazole derivatives are: Sum of E-State descriptors of strength for potential hydrogen bonds of path length 9 (SHBint9) and topological radius (minimum atom eccentricity) (topoRadius) descriptors. In the design of new set of compounds with potent antioxidant activities, these highly influential descriptors should be considered.

## REFERENCES

1. Brewer MS. Natural Antioxidants: Sources, Compounds, Mechanisms of Action, and Potential Applications. *Compr. Rev. Food Sci. Food Saf.*, 2011 June;10: 221-247. DOI: 10.1111/j.1541-4337.2011.00156.x
2. Valko M, Leibfritz D, Moncol J, Cronin MTD, Mazur M, Telser J. Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* 2007; 39(1):44-84.

3. Kotaiah Y, Harikrishna N, Nagaraju K, Venkata RC. Synthesis and antioxidant activity of 1,3,4-oxadiazole tagged thieno[2,3-d]pyrimidine derivatives. *Eur J Med Chem.* 2012 Dec; 58:340-5. Doi: 10.1016/j.ejmech.2012.10.007
4. Zhang Y, Zou B, Chen Z, Pan Y, Wang H, Liang H, Yi X. Synthesis and antioxidant activities of novel 4-Schiff base-7-benzyloxy-coumarin derivatives. *Bioorg Med Chem Lett.* 2011 Nov;21(22):6811-5. Doi: 10.1016/j.bmcl.2011.09.029
5. Kang TS, Jo HO, Park WK, Kim JP, Konishi Y, Kong JY, Park NS, Jung YS. Synthesis and antioxidant activities of 3,5-dialkoxy-4-hydroxycinnamamides. *Bioorg Med Chem Lett.* 2008 Mar; 18(5):1663-7. doi: 10.1016/j.bmcl.2008.01.061
6. Ma L, Xiao Y, Li C, Xie ZL, Li DD, Wang YT, Ma HT, Zhu HL, Wang MH, Ye YH. Synthesis and antioxidant activity of novel Mannich base of 1,3,4-oxadiazole derivatives possessing 1,4-benzodioxan. *Bioorg Med Chem.* 2013 Nov; 21(21):6763-70. Doi: 10.1016/j.bmc.2013.08.002
7. Lu J, Li C, Chai YF, Yang DY, Sun CR. The antioxidant effect of imine resveratrol analogues. *Bioorg Med Chem Lett.* 2012 Sep; 22(17):5744-7. doi: 10.1016/j.bmcl.2012.06.026
8. Molyneux, P. The use of the stable free radical diphenylpicrylhydrazyl (DPPH) for estimating antioxidant activity. *Songklanakarin J. Sci. Technol.*, 2004, 26(2): 211-219.
9. Hamid AA, Aiyelaagbe OO, Usman LA, Ameen OM, Lawal A. Antioxidants: Its medicinal and pharmacological applications. *Afr. J. Pure Appl. Chem.* 2010 Aug; 4(8), 142-151.
10. Mohana KN, Kumar CBP. Synthesis and Antioxidant Activity of 2-Amino-5-methylthiazol Derivatives Containing 1,3,4-Oxadiazole-2-thiol Moiety. *ISRN Org Chem.* 2013 Aug; 2013: 620718. doi: 10.1155/2013/620718
11. Aanandhi MV, Mansoori MH, Shanmugapriya S, George S, Shanmugasundaram P. Synthesis and In-vitro antioxidant activity of substituted Pyridinyl 1, 3, 4 oxadiazole derivatives. *Res. J. Pharm., Biol. Chem. Sci.*, 2010 Oct-Dec;1(4), 1083-1090.
12. Ogadimma AI, Adamu U. Quantitative Structure Activity Relationship Analysis of Selected Chalcone Derivatives as Mycobacterium tuberculosis Inhibitors. *OALib.* 2016 Mar; 3: e2432 doi: <http://dx.doi.org/10.4236/oalib.1102432>.
13. Sanmati KJ, Achal M. 3D QSAR Analysis on Isatin Derivatives as Carboxyl Esterase Inhibitors Using K-Nearest Neighbor Molecular Field Analysis. *J Theor Comput Sci.* 2015 May; 2:124. Doi:10.4172/2376-130X.1000124
14. Aicha K, Belaidi S, Lanez T. Computational Study of Structure-Property Relationships for 1,2,4-Oxadiazole-5-Amine Derivatives. *Quantum Matter.* 2016 Feb; 5(1): pp. 45-52(8). DOI:10.1166/qm.2016.1253
15. Yokoi T, Nakagawa Y, Miyagawa H. Quantitative structure-activity relationship of substituted imidazothiadiazoles for their binding against the ecdysone receptor of Sf-9 cells. *Bioorg. Med. Chem. Lett.*, 2017 Oct; 27(23):5305-5309. DOI: 10.1016/j.bmcl.2017.10.013
16. Alisi IO, Uzairu A, Abechi SE, Idris SO, Quantitative structure activity relationship analysis of coumarins as free radical scavengers by genetic function algorithm. *Phys. Chem. Res.* 2018 Jan; 6(1):208-222. DOI:10.22036/pcr.2017.95755.1409
17. Li Z, Wan H, Shi Y. Ouyang ,P., Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J Chem Inf Comput Sci.* 2004 Sep-Oct; 44(5): 1886–1890. DOI:10.1021/ci049794h
18. Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert ATB, Slipchenko LV, Levchenko SV, O'Neill DP, DiStasio RA, Lochan RC, Wang T, Beran GJO, Besley NA, Herbert JM, Lin CY, Van Voorhis T, Chien SH, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EFC, Dachsel H, Doerksen RH, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu CP, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang WZ, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock III HL, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF, Kong J, Krylov AI, Gill PMW, Head-Gordon M. Advances in methods and algorithms in modern quantum chemistry program package. *Phys. Chem. Chem. Phys.*, 2006 June; 8(27): 2006, pages 3172-3191. Doi.org/10.1039/B517914A
19. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens. Matter*, 1988 Jan; 37(2): 785-789. DOI:<https://doi.org/10.1103/PhysRevB.37.785>
20. Yap C. W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* 2011 May; 32(7): 1466–1474. Doi: 10.1002/jcc.21707
21. Ballabio D, Consonni V, Mauri A, Claeys-Bruno M, Sergeant M, Todeschini R. Comb. Chem. High Throughput Screening. 2014 June; 136: 147-154. DOI: 10.1016/j.chemolab.2014.05.010
22. Ambure P, Aher RB, Gajewicz A, Puzyn T. NanoBRIDGES" software: Open access tools to perform QSAR and nano-QSAR modelling. *Chemom. Intell. Lab. Syst.* 2015 Oct; 147: 1-13. doi.org/10.1016/j.chemolab.2015.07.007
23. Brignole M, Auricchio A, Baron-Esquivias G, Bordachar P, Boriani G, Breithardt OA, Cleland J, Deharo JC, Delgado V, Elliott PM, Gorenek B, Israel CW, et al., 2013 ESC Guidelines on cardiac pacing and cardiac resynchronization therapy: the Task Force on cardiac pacing and resynchronization therapy of the European Society of Cardiology (ESC). Developed in collaboration with the European Heart Rhythm Association (EHRA). *Eur Heart J.* 2013 Aug;34(29):2281-329. doi: 10.1093/eurheartj/ehf150.
24. Golbraikh A1, Tropsha A. Beware of q2! *J Mol Graph Model.* 2002 Jan;20(4):269-76. Doi.org/10.1016/S1093-3263(01)00123-1
25. Todd MM, Harten P, Douglas MY, Muratov EN, Golbraikh A, Zhu H, Tropsha A. Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR

- Modeling? *J. Chem. Inf. Model.* 2012 Oct; 52 (10): 2570–2578. DOI: 10.1021/ci300338w
26. Mitra I, Saha A, Roy K. Chemometric QSAR modeling and in silico design of antioxidant NO donor phenols. *Sci Pharm.* 2011 Jan-Mar; 79(1):31-57. Doi: 10.3797/scipharm.1011-02.
27. Das ND, Roy K. Development of classification and regression models for *Vibrio fischeri* toxicity of ionic liquids: Green solvents for the future. *Toxicol. Res.* 2012 July; 1: 186-195. DOI:10.1039/C2TX20020A
28. Tropsha A, Gramatica P, Gombar, VK. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* 2003 April; 22(1), 69-77. DOI: 10.1002/qsar.200390007
29. Kar S, Roy K. Development and validation of a robust QSAR model for prediction of carcinogenicity of drugs. *Indian J Biochem Biophys.* 2011 Apr;48(2):111-22.
30. Roy PP, Roy K. On some aspects of variable selection for partial least squares regression models *QSAR Comb. Sci.* 2008 March; 27(3): 302-313. Doi:10.1002/qsar.200710043
31. Roy K, Mitra I. On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. *Comb Chem High Throughput Screen.* 2011 Jul;14(6):450-74. DOI:10.2174/138620711795767893
32. Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN. Some case studies on application of " $r(m)^2$ " metrics for judging quality of quantitative structure-activity relationship predictions: emphasis on scaling of response data. *J Comput Chem.* 2013 May 5;34(12):1071-82. doi: 10.1002/jcc.23231
33. Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inf.* 2010 July; 29(6-7): 476-488. DOI: 10.1002/minf.201000061
34. Roy K, Kar S, Das RN. Statistical Methods in QSAR/QSPR. *Springer Briefs in Molecular Science.* 2015. p. 37-59.
35. Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MT, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts D, Schultz T, Stanton DW, van de Sandt JJ, Tong W, Veith G, Yang C. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM Workshop 52. *Altern Lab Anim.* 2005 Apr;33(2):155-73.
36. Gramatica P, Giani E, Papa E. Statistical external validation and consensus modeling: a QSPR case study for Koc prediction. *J. Mol. Graphics Modell.* 2006 Aug; 25(6):755-766 DOI: 10.1016/j.jmgm.2006.06.005
37. P. Gramatica, Chemiometric methods and theoretical molecular descriptors in predictive QSAR modeling of the environmental behavior of organic pollutants. In: T. Puzyn et al. (eds.), *Recent Advances in QSAR Studies*, (Dordrecht, Heidelberg, London, 2010), pp. 327-366.
38. Sharma BK, Singh P. Chemometric Descriptor Based QSAR Rationales for the MMP-13 Inhibition Activity of Non-Zinc-Chelating Compounds. *Med chem.* 2013 April; 3:168-178. Doi:10.4172/2161-0444.1000134
39. Mitra I, Saha A, Roy K. Chemometric modeling of free radical scavenging activity of flavone derivatives. *Eur J Med Chem.* 2010 Nov;45(11):5071-9. Doi: 10.1016/j.ejmech.2010.08.016
40. Veerasamy R, Rajak H, Abhishek J, Sivadasan S, Varghese CP, Agrawal RK. Validation of QSAR Models - Strategies and Importance. *Int. J. Drug Des. Discovery* 2011 July - September; 2(3): 511-519.
41. Saaidpour S. Quantitative Modeling for Prediction of Critical Temperature of Refrigerant Compounds. *Phys. Chem. Res.* 2016 March; 4(1): 61-71. DOI: 10.22036/pcr.2016.11759
42. Baumann K. Chance Correlation in Variable Subset Regression: Influence of the Objective Function, the Selection Mechanism, and Ensemble Averaging. *QSAR Comb. Sci.* 2005 November; 24(9):1033-1046. DOI: 10.1002/qsar.200530134.