RESEARCH PAPER

A Quantitative Structure-Antioxidant Relationship (QSAR) model for 1,3,4-oxadiazole derivatives using PLS regression

Salim N.A Saber*1, Hikmat A. Mohamad1 and Madzlan Aziz2

1-Department of Chemistry department, College of Education, Salahaddin University-Erbil, Kurdistan region-Iraq. 2-Department of Chemistry, Faculty of Science, University Technology Malaysia, Johor, Malaysia

ABSTRACT:

Antioxidants can control the generation of free radical by terminating the reaction chain. 1,3,4-oxadizol derivatives are the most important heterocyclic compounds which exhibit antioxidant activity. This study aims to build a reliable quantitative structure antioxidant relationship (QSAR) model of 1,3,4-oxadiazol derivatives using iPLS as a variable selection method, and PLS as a regression method. The QSAR model was developed based on the correlation between the antioxidant activities in DPPH (2, 2-diphenyl-1-picryl hydrazyl) assay of 52 oxadiazol derivatives and their molecular descriptors. Only three descriptors that contribute to the antioxidant property are identified and selected to build QSAR model. The performance of QSAR model is reported as r2cal (0.92) and the model is validated by leave-one-out cross-validation technique r2cv (0.91), and predicted correlation coefficient r2pre (0.91). Based on the findings, the iPLS-PLS model can explain 91% variance of antioxidant activity. The result shows that the GATS2m and E3s descriptors are positively correlated with DPPH values, while the R7s+ has negative correlations with DPPH values. This finding proves that the increasing atomic mass and central symmetric atoms of 1,3,4-oxadiazol cause to increase the antioxidant activity. The final QSAR model can be used as a guide to predict free radical scavenger activities of new synthesized 1,3,4-oxadiazol compounds.

KEY WORDS: Antioxidant, 1,3,4-Oxadiazol, QSAR and PLS.
DOI: http://dx.doi.org/10.21271/ZJPAS.31.s4.17
ZJPAS (2019), 31(s4);109-115.

1. INTRODUCTION:

Free radicals are produced by various metabolic functions in human body which cause many diseases such as cell ageing, cardiovascular diseases, and cancerous tumor growth. Antioxidants can control the generation of free radical by terminating the reaction chain. They can be classified into two main classes such as natural antioxidants (glutathione reductase, vitamins E and C and carotenoids) and synthetic antioxidants (ebselen, trolox, and raxofelas).

In the last few decades, the variety of chemical compounds has been used as synthetic antioxidants (M. Rabie A., S. Tantawy A. and M. I. Badr S. 2016). Oxadiazole is a five membred hetrocyclic organic compound that has the one oxygen, two nitrogen and two carbon atoms in the ring. The general structure of 1,3,4-oxadiazole and it derivatives are illustrated in Figure (1). According to the position of nitrogen atoms, oxadiazoles have four isomer but 1,3,4-Oxadiazoles are the most important heterocyclic compounds because their derivatives have variety biological activity (P. C.a,b et al., 2015) such as anticancer (Sun J et al., 2013), antioxidant (Musad et al., 2011) and analgesic (Nazreen et al., 2014). According to reported literatures, the biological activities of 1,3,4-
oxadiazoles depends on their molecular structure (Zicane et al 2014 and Kotaiah et al., 2012).

There are variety methods have been reported to measured antioxidant activity of organic compounds, but DPPH (2, 2-diphenyl-1-picryl hydrazyl) assay is the most common assay to measure antioxidant activity. In DPPH assay, the antioxidant compounds are to react with DPPH. By electron donation, DPPH (purple color) converts to α,α-diphenyl-β-picryl hydrazine (yellow color). The DPPH reaction with antioxidant compound is shown in Figure (2) (Sauer et al., 2017)

Quantitative structure activity relationship (QSAR) is a method to identify the relationship between biological activity and molecular structure of compounds. QSAR model is a mathematical equation between biological activity and molecular structure. In QSAR modeling, the biological activity is used as a dependent variable when the molecular structure of the compounds as independent variables. The first step to build a QSAR model is to convert the molecular structure of compounds to numerical representation which called molecular descriptor. The molecular descriptors are classified into 0-Dimensional (0D), 1-Dimensional (1D), 2-Dimensional (2D) and 3-Dimensional (3D) (Meor et al 2017). The most significant step of QSAR modeling is to find the reliable variable selection method to select the suitable descriptors which related to the biological activity. Interval PLS (IPLS) is a variable selection method. IPLS selects a subset of variables which will give superior prediction compared to using all the variables in a data set. The concept of IPLS is to split the spectra into several equidistant subintervals, and construct PLS models on each subinterval (Chen, Jiang and Zhao, 2010). IPLS method performs exhaustive search for the most relevant variable or combination of variables. It begins by selecting a number of variables as a starting step. The next step is to add or remove new variables from the original variables (Abrahamsson et al., 2003). Partial least squares (PLS) approach is a regression method. PLS regression is an effective approach for finding the correlation between a molecule structure and its properties. Mathematically, PLS relates dependent variables matrix (Y) to molecular descriptors matrix (X). The objectives of PLS are to achieve several steps. PLS approximate the X and Y data matrices, and maximize the correlation between them. A regression equation relating each Y variable with the X matrix is created during the stepwise extraction of PLS components and the independently assessment of the importance of each component. PLS splits the X matrix into several latent variables with best correlation with the molecules activities (Sharma et al., 2013). Partial least squares regression (PLS regression) provides the equation which can explain the relation between one dependent variable (biological activity) and a group of independent variables (molecular descriptors) (Asadollahi et al., 2014). This study aims to build a reliable quantitative structure activity relationship (QSAR) model of 1,3,4-oxadiazole derivatives using interval partial least squares (IPLS) as a variable selection method, and PLS as a regression method.

2. METHODOLOGY
2.1. Data sets arrangement

1,3,4-Oxadiazole compounds with DPPH value as antioxidant activity were used as a data set. There are a number of oxadizole compounds with DPPH were obtained from literature (K. Mehta D. and Das R. 2011, Kotaiah Y., Harikrishna N., Nagaraju K., Venkata Rao C.2012, Sumangala et al., 2012, N. Mohana K. and B. Pradeep Kumar
C. 2013, S. Kareem et al., 2016, Aruna et al., 2016, R. R, Kumar Jat R. and Saravanan J. 2016 and Mustafa, 2014) (See Supplementary material). The compounds were rearranged in increased order of their DPPH values and then the data was spitted into training and test sets in an approximate ratio of 2:1. The number of training and test sets were 36 and 16 compounds, respectively. The training set was used for model development and test set for model validation.

2.2. Model development

Chem3D software (Cambridge Soft Corporation, Cambridge, MA) was utilized for drawing 2D-representation of 1,3,4-oxadazole compounds. After that 3D-representation of compounds was obtained by Chem3D Pro 7.0 software (Cambridge Soft Corporation, Cambridge, MA). Then the structures of compounds were optimized using the molecular mechanics (MM2) and molecular orbital package (MOPAC) module implemented in Chem3D software. DRAGON 6.0 Software was used to generate descriptors in different dimension blocks. The total number of descriptors was 4885 descriptors for each compound. The descriptors were reduced by the following methods:

1. DRAGON software was used to remove the descriptors with a relative standard deviation of less than 0.001 and removing those that were highly correlated (r ≥ 0.90)
2. The constant and near constant of descriptors were removed.
3. Then, the best descriptors were selected using IPLS variable selection method. Finally, three descriptors were selected to build the model.

The QSAR model was derived by performing the IPLS analysis with PLS in PLS tool-box in Solo software.

2.3. Model Validation

The performance model is reported as \( r^2_{(calc)} \). This parameter is squared correlation coefficient between experimental and prediction activity of training set compounds. The predicted activities were calculated by using developed QSAR model then compare with experimental values. The accepted value of \( r^2_{calc} \) is more than 0.6. Leave-one-out cross-validation technique was used to validate the developed QSAR model. In this technique, each compound from training set was removed once and then the model was built to remaining compounds. The new model was used to predict antioxidant activity for removed compound. These steps were repeated until all compounds were removed at least once. Then cross-validated squared correlation coefficient (\( r^2_{cv} \)) between predicted and experimental activities was calculated. This technique called internal validation technique. The predicted correlation coefficient (\( r^2_{pred} \)) was used as external validation technique. In this technique, the developed QSAR model was used to predict antioxidant activities of compounds that were not used to developed model. These compounds called test set compounds. The \( r^2_{pred} \) is squared correlation coefficient between experimental and predicted antioxidant activities of test set compounds (Meor et al 2017). Finally, to check co-linearity between descriptors, the variance inflation factor (VIF) value was calculated by the following equation:

\[
\text{VIF} = \frac{1}{1 - r^2}
\]

where \( r^2 \) is a correlation coefficient of multiple regressions between each variable and the other variables of the constructed QSAR model.

3. RESULTS AND DISCUSSION

To build the QSAR model for antioxidant activity of 1,3,4-oxadiazole derivatives, 52 compounds were used and all compounds contain 1,3,4-oxadizole ring with different other groups. DPPH values of all compounds are shown as IC\textsubscript{50} (µg/ml). All 1,3,4-oxadizole with DPPH value are shown in (Supplementary material).

QSAR Model

There are several commonly used regression methods such as multiple linear regression (MLR) and partial least squares (PLS). The MLR method works well as long as the X-variables are fairly few and fairly uncorrelated. Unlike MLR, PLS regression method can analyze data with strongly collinear (correlated) and numerous X-variables (Wold, Sjöström and Eriksson, 2001). In our
current work, a large number of the calculated molecular descriptor (X-variables) was obtained with high correlation between them; therefore, PLS regression method is of particular interest to model the obtained data. The number of the descriptors generated by DRAGON software should be reduced by using variable selection methods. After removing descriptors that have standard deviation less than 0.001 and highly correlated \((r_{ij} \geq 0.90)\), 4885 descriptors were reduced to 845 descriptors. By removing constant and near constant descriptors, the numbers of descriptors were reduced to 566. By iPLS, the 566 descriptors were reduced to 3 descriptors only. These 3 descriptors were used to build the QSAR model. The numbers of descriptors and the used variable selection methods were demonstrate in Table 1. This Table shows that, only 3 descriptors are remained after variable selection methods. The detail information of all 3 descriptors are shown in Table 2.

The developed QSAR model with 3 significant descriptors has high statistical parameter values, i.e., \(r_{cal}^2\), \(r_{cv}^2\), and \(r_{pre}^2\). The \(r_{cal}^2\) value of 0.92 is greater than 0.6 indicating that the predicted antioxidant activity of training set was closely fitted to the experimental antioxidant activity. The developed QSAR model was validated by internal validation and external validation techniques. Leave-one-out cross-validation technique was used as internal validation technique. The \(r_{cv}^2\) of 0.91 is less than the \(r_{cal}^2\) value, indicating that the developed QSAR model can avoid over fitting. The \(r_{pred}^2\) was used as external validation technique and its value of 0.91 shows that the predicted antioxidant activity of test set closely fitted to the experimental antioxidant activity indicating that the developed QSAR model can be applied as a guide to predict antioxidant activity of new synthesized 1,3,4-oxadiazol compounds as illustrate (3). All predicted antioxidant activity of training and test sets in DPPH assay of 1,3,4-oxadiazoles were calculated by equation 2 and then plotted against experimental antioxidant activity as illustrated in Figure 3. This Figure shows that all points were scattered around the line of fit.

\[
\begin{align*}
\text{IC}_{50} &= -146.46 + 198.62 \text{GATS2m} + 371.54 \text{E3s} - 156.30 \text{R7s}^+ \\
\text{R}_{\text{cal}}^2 &= 0.92, \quad \text{R}_{\text{cv}}^2 = 0.91, \quad \text{R}_{\text{pred}}^2 = 0.91 \quad \text{RMSEC: 17.83, RMSECV: 19.59, RMSEP: 20.91}
\end{align*}
\]

\[
\text{R}^2 = 0.928, \quad \text{RMSEC: 17.83, RMSECV: 19.59, RMSEP: 20.91}
\]

Table (2): Name and Descriptions with Block

| NO | Descriptor name | Description | Block |
|----|----------------|-------------|-------|
| 1  | GATS2m         | Geary autocorrelation of lag 2 weighted by mass 3rd component accessibility autocorrelation | 2D autocorrelations |
| 2  | E3s            | directional WHIM index / weighted by I-state | WHIM descriptors |
| 3  | R7s+           | R autocorrelation of lag 7 / weighted by I-state | GETAWAY descriptors |

The robust and reliable developed QSAR model was chosen according to \(r_{cal}\), \(r_{cv}\) and \(r_{pre}\), Root-Mean-Square-Error of Calibration (RMSEC), Root-Mean-Square-Error of Cross Validation (RMSECV) and Root-Mean-Square-Error of Prediction (RMSEP). The QSAR model with 3 descriptors is demonstrated by the equation 2.

\[
\text{Figure (3): The predicted against experimental antioxidant activity plot using QSAR model for training and test sets.}
\]
Table (3): Structure 1,3,4-oxadiazol compounds with Experimental and Predicted antioxidant activity

| Structure | Experimental (IC$_{50}$) | Predicted (IC$_{50}$) |
|-----------|--------------------------|-----------------------|
| ![Structure 1](image1) | 217.34 | 210.17 |
| ![Structure 2](image2) | 183.08 | 164.07 |
| ![Structure 3](image3) | 173.08 | 158.04 |
| ![Structure 4](image4) | 167.32 | 173.39 |
| ![Structure 5](image5) | 89.1 | 51.65 |
| ![Structure 6](image6) | 56.3 | 49.82 |
| ![Structure 7](image7) | 44 | 31.59 |
| ![Structure 8](image8) | 41.27 | 41.22 |
| ![Structure 9](image9) | 33.1 | 73.30 |
| ![Structure 10](image10) | 25.7 | 53.16 |
| ![Structure 11](image11) | 24.1 | 51.65 |
| ![Structure 12](image12) | 19.57 | 18.78 |
| ![Structure 13](image13) | 18.4 | 43.28 |
| ![Structure 14](image14) | 16.35 | 24.60 |
| ![Structure 15](image15) | 15.0 | 36.96 |

Experimental (IC$_{50}$) is antioxidant activity obtained from literatures. Predicted (IC$_{50}$) is antioxidant activity calculated from developed QSAR model.

According to the equation 2, GATS2m, E3s and R7s are the most significant descriptors that show significant contributions to antioxidant properties of 1,3,4-oxadiazoles. To check co-linearity problem between descriptors, VIF was calculated according to the equation 1 and the acceptable VIF value should be less than 5 for co-linearity problem. The obtained VIF values of all descriptors are less than 4, indicating that there is no co-linearity problem of descriptors in the developed QSAR model. The correlation matrixes between all descriptors are shown in Table (4). The correlation values between descriptors are
less than 0.3, indicating that the descriptors are not represent the same properties in the compounds. The equation 2 shows that the GATS2m and E3s descriptors are positively correlated with DPPH values, while the R7s+ descriptor has negative correlations with DPPH value. GATS2m is Geary autocorrelation of lag 2 weighted by mass and GATS2m is increased by increasing the atomic mass of compounds. It can be concluded that, the atomic mass of compounds will lead to increase the inhibitory activity. The second significant descriptor is E3s (3rd component accessibility directional WHIM index/weighted by I-state). By increasing the central symmetric atoms of compounds the value of this descriptor would increase, causing an increase in DPPH values. R7s+ (R autocorrelation of lag 7 / weighted by I-state) is R-GETAWAY descriptor. R7s+ is representation of molecular geometry in terms of an influence-distance matrix and obtained through the values of atomic Cartesian coordinates.

Table (4): Correlations between descriptors and VIF values.

|        | GATS2m | E3s | R7s | VIF  |
|--------|--------|-----|-----|------|
| GATS2m | 1      |     |     | 2.557|
| E3s    | 0.05   | 1   |     | 3.937|
| R7s+   | 0.00   | 0.05| 1   | 1.634|

4. CONCLUSIONS

The reliable QSAR model is successfully developed to explain the relationship between a series of 1,3,4-oxadiazole compounds and their antioxidant activities. The significance and robustness of the developed QSAR models have been confirmed by $r^2_{cal}$ and validated by cross-validated squared correlation coefficient ($r^2_{cv}$), predicted correlation coefficient ($r^2_{pre}$). Three significant descriptors were selected to build QSAR model. The result shows that the following descriptors: GATS2m and E3s are positively correlated with DPPH values, while the R7s+ has negative correlations with DPPH value. The increase in atomic mass and central symmetric atoms of 1,3,4-oxadiazole cause increase in the antioxidant activity values. The final QSAR model can be used as a guide to predict free radical scavenger activities of new synthesized 1,3,4-oxadiazol compounds.

Acknowledments

The authors acknowledge the University of Salahaddin-Erbil to express their appreciation for the support and sponsor of this project. The authors wish to thank the Department of Chemistry, Faculty of Science, University Technology Malaysia (UTM), for providing necessary facilities.

REFERENCES

Abrahamsson, C., Johansson, J., Sparén, A. and Lindgren, F. (2003). Comparison of different variable selection methods conducted on NIR transmission measurements on intact tablets. Chemometrics and Intelligent Laboratory Systems, 69(1-2), pp.3-12.

Aruna Sindhe M., D. Bodke Y., Kenchappa R., Telkar S. and Chandrashhekara A. (2016). Synthesis of a series of novel 2,5-disubstituted-1,3,4-oxadiazole derivatives as potential antioxidant and antibacterial agents. J Chem Biol, 9, pp 79–90.

Asadollahi T., Dafdarnia S., Mohammad Haji Shabani A. and B. Ghasemi J.(2014). Use of the Genetic Algorithm for Variable Selection of PLS Regression in a QSAR Study on [4,5-d] pyrimidinederivatives Antagonist of CXCR2. Communications in Mathematical and in Computer Chemistry, 71. pp 287-304.

Chen, Q., Jiang, P. and Zhao, J. (2010). Measurement of total flavone content in snow lotus (Saussurea involucrate) using near infrared spectroscopy combined with interval PLS and genetic algorithm. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 76(1), pp.50-55.

K. Mehta D. and DasR. (2011). Synthesis And In-Vitro Antioxidant Activity Of Some New 2, 5-Disubstituted-I, 3, 4-Oxadiazoles Containing Furan Moiety. International Journal of Pharmaceutical Sciences and Research, 2(11), pp 2959 – 2963.

Kotaiah Y., Harikrishna N., Nagaraju K. and Venkata Rao C. (2012). Synthesis and antioxidant activity of 1,3,4-oxadiazole tagged thieno[2,3-d] pyrimidine derivatives. European Journal of Medicinal Chemistry, 58, pp 340-345.

M. Rabie A., S. Tantawy A. and M. I. Badr S. (2016). Design, Synthesis, and Biological Evaluation of Novel 5-Substituted-2-(3,4,5-trihydroxyphenyl)-1,3,4-oxadiazoles as Potent Antioxidants. American Journal of Organic Chemistry, 6(2), pp.54-80.

Meor Ahmad M., Rafidah Wan Alwi S., Jamaludin R., Suan Chua L., and A. Azri Mustaffa A. (2017). Quantitative Structure-Activity Relationship Model for Antioxidant Activity of Flavonoid Compounds in Traditional Chinese Herbs. Chemical Engineering Transactions, 56, pp 1039 – 1044.

Musad, E., Mohamed, R., Ali Saeed, B., Vishwanath, B. and Lokanatha Rai, K. (2011). Synthesis and evaluation of antioxidant and antibacterial activities of new substituted bis(1,3,4-oxadiazoles), 3,5-bis(substituted) pyrazoles and...
isoxazoles. *Bioorganic & Medicinal Chemistry Letters*, 21(12), pp.3536-3540.

Mustafa Shakir R. (2014). Ph.D Dissertation, University Of Malaya.

N. Mohana K. and B. Pradeep Kumar C. (2013). Synthesis and Antioxidant Activity of 2-Amino-5-methylthiazol Derivatives Containing 1,3,4-Oxadiazole-thiol Moiety. *ISRN Organic Chemistry*, 2012, pp 1-8.

Nazreen, S., Alam, M., Hamid, H., Yar, M., Shafi, S., Dhulap, A., Alam, P., Pasha, M., Bano, S., Alam, M., Haider, S., Ali, Y., Kharbanda, C. and Pillai, K. (2014). Design, synthesis, in silico molecular docking and biological evaluation of novel oxadiazole based thiadizolidine-2,4-diones bis-heterocycles as PPAR-γ agonists. *European Journal of Medicinal Chemistry*, 87, pp.175-185.

P. C, R. J, Mustafa A, S., Kallurayaa B., S. K.c P. and A. M.d V. (2015). Synthesis, characterization, antidiabetic and antioxidant activity of 1,3,4-oxadiazole derivatives bearing 6-methyl pyridine moiety. *Der Pharma Chemica*, 7(12). pp 137-145.

R. R, Kumar Jat R. and Saravanan J. (2016). Synthesis, and in vitro antioxidant activity of novel 1, 3, 4-oxadiazole-2-thione. *Journal of Innovations in Pharmaceuticals and Biological Sciences*, 3 (3), pp 114-122, 2016

S. Kareem H., Nordin N., Heidelberg T., Abdul-Aziz A. and Ariffin A. (2016) Conjugated Oligo-Aromatic Compounds Bearing a 3,4,5-Trimethoxy Moiety: Investigation of Their Antioxidant Activity Correlated with a DFT Study. *Molecules*, 21, pp 1-19.

Sauer, A., Leal, J., Stefanello, S., Leite, M., Souza, M., Soares, F., Rodrigues, O. and Dornelles, L. (2017). Synthesis and antioxidant properties of organosulfur and organoselenium compounds derived from 5-substituted-1,3,4-oxadiazole/thiadiazole-2-thiols. *Tetrahedron Letters*, 58(1), pp.87-91.

Sharma, M., Sharma, S., Sahu, N. and Kohli, D. (2013). QSAR studies of some substituted imidazolinones angiotensin II receptor antagonists using Partial Least Squares Regression (PLSR) method based feature selection. *Journal of Saudi Chemical Society*, 17(2), pp.219-225.

Sumangala V., Poojarya B., Chidananda N., Arulmolib T. and ShenoyS. (2012). Synthesis, characterization, antimicrobial and antioxidant activity of some disubstituted [1,3,4]-oxadiazoles carrying 4-(methylsulfonyl/sulfinyl)benzyl moieties. *Journal of Chemical and Pharmaceutical Research*, 4(3), pp 1661-1669.

Sun J., Zhan H. Yang Z. and Zhu H. (2013). Synthesis, molecular modeling and biological evaluation of 2-aminomethyl-5-(quinolin-2-yl)-1,3,4-oxadiazole-2(3H)-thione quinolone derivatives as novel anticancer agent. *European Journal of Medicinal Chemistry*, 60, pp23-28.

Wold, S., Sjöström, M. and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), pp.109-130.

Zicane D., Tetere Z., Mierina I., Turks M. , Ravina I. and Leonciks A. (2014). Synthesis of quinazolinone-1,3,4-oxadiazole conjugates and studies of their antibacterial and antioxidant activity. *Journal of Chemical and Pharmaceutical Research*, 6(4), pp 1153-1158.