Research Article

# Isoxazoline-containing Podophyllotoxin/2′(2′,6′)-(DI) Halogenopodo - phyllotoxin Derivatives as Acaricidal Activities against *Tetranychus cinnabarinus*. 2D-QSAR Study by using Molecular Operating Environment (MOE)

**Fatima En-nahli[1], Hanane Zaki[1], Abdellah EL AISSOUQ[3] , Abdelkrim Ouammou[3], Tahar Lakhlifi[1] and Mohammed Bouachrine[1, 2*]**

[1]MCNS Laboratory, Faculty of Science, Moulay Ismail University, Meknes, Morocco
[2]EST Khenifra, Sultan Moulay Slimane University, Morocco
[3]LIMOME Faculty of Science, Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco
*Corresponding Author: E-mail: m.bouachrine@umi.ac.ma

**Abstract** The carmine spider mite, *Tetranychus cinnabarinus*, is one of the most serious mite pests on crops. It is capable to rapidly develop resistance to pesticides and the control methods still difficult. Thus, make the development of new natural-product-based insecticidal agents a very desirable method for sustainable pest control. The present study was carried out to explain the relationships between chemical structure and experimental observations. In which we have employed the Molecular Operating Environment (MOE), powerful molecular visualization software that can be implemented on a variety of operating platforms. A total of 29 molecules of isoxazoline-containing podophyllotoxin/2′(2′,6′)-(di)halogenopodophyllotoxin derivatives were subjected to quantitative structure-activity relationship (2D-QSAR) analysis, based on their acaricidal activity, to find two models at 48h and 72h. The two 2D-QSAR models demonstrated that two descriptors lip-don and PEOE-VSA-FNEG are likely to influence the acaricidal activity of these compounds.
Statistically robust 2D-QSAR models were developed for all studied compounds with ($R^2$= 0.83, $Q^2$ = 0.79, $R^2_{test}$ = 0.64 and RMSE =0.06) for the first model at 48h, moreover, the second model at 72h showed better performance with the($R^2$= 0.92, $Q^2$ = 0.83, $R^2_{test}$ = 0.67 and RMSE = 0.16).

## Introduction

The carmine spider mite, *Tetranychus cinnabarinus* (Boisduval) is classed as a type of arachnid, is considered as one of the most crop-threatening insect pests of field and greenhouse [1], These mites attack several crops such as cotton, melons, beans, eggplants, tomatoes, peppers, cucurbits and strawberries and cause crops damage in term of productivity and plants performances. As a result of the economic loss, it became a serious problem for agriculture in different countries around the globe [2]. Spider mites usually feed on the cellular contents of the plant and cause a negative effect in transpiration and photosynthesis [3]. Plants lightly infested with the mite showed leaf discoloration, defoliation, and reduced fruit quality. On the other hand, severe infestations cause the death of the plants and complete crop loss [4]. The mite life cycle is very specialized to the parasitism. First, the mite species

survive the cold climate of the winter as eggs on the leaves and bark of the host plant, waiting for the spring where the temperature increased and the conditions became more adequate for the development of tiny six-legged larvae, then for few days they start feeding and hatching until the first nymphal stage, where the young nymphs developed eight legs and two more molts before the final maturity stage. The females can produce about 300 eggs in a few weeks. Finally, in hot, dry weather the eggs pass rapidly to pests (only in 5 days) and the mite can have many overlapping generations per year [5][6]. Spider mites have a very small size, colored and difficult to be seen. Therefore, it's difficult to prevent and control this arachnid because of its high fecundity, high rate of inbreeding, and high to develop resistance of pesticide [7].

Several acaricide active ingredients are used in pesticides, but the current problem is that pesticides used for pest control also kill beneficial insects and toxic to the environment and negative impacts on human health. Also, the control of spider mites has become difficult due to their resistance to many synthetic pesticides [8]. Nowadays, the discovery and development of natural products-based pesticidal agents have received much research attention, lots of synthetic agrochemicals have been introduced to protect crops from pests and enhance crop yields [9].

Quantitative structure-activity relationship (QSAR) studies are very useful tools to correlate the chemical structure of a compound with the bioactivity [10]. QSAR plays an important role in the design and discovery of new insecticidal and acaricidal agents, these substances are used for increasing production yields and quality of agricultural products [11].

The objective of this study was to develop a 2D-QSAR model by using Molecular Operating Environment (MOE) software to a series of isoxazoline-containing podophyllotoxin/2′(2′,6′)-(di)halogenopodophyllotoxin derivatives, and to extract the most relevant descriptors with the final mortality rate of the compounds. The performance of the proposed model was evaluated by several statistic parameters, such as squared correlation coefficient ($R^2$), leave-one-out cross-validation correlation coefficient ($Q^2$) for internal predictive ability and root mean square error (RMSE). The plot of the standardized residuals versus the leverage value (Williams plot), was employed to describe the applicability domain (AD) of the QSAR model. To develop a stable QSAR model with high predictive power by calculating the $R^2_{test}$ for external validation [12].

**Material and Methods**

**Acaricidal activity**

As an initial step before working on the QSAR model fetching is to collect a list of compounds, a group of 29 molecules with a variety in mortality rate were used as the model dataset. In our work, 2D-QSAR models were devised to quantitatively correlate the 2D molecular feature descriptors of this series with their acaricidal activities against *Tetranychus cinnabarinus*. A series of isoxazoline-containing podophyllotoxin/2′(2′,6′)-(di)halogenopodophyllotoxin derivatives, were synthesized by *Yang et al* [13] by the introduction of the isoxazoline fragment into esters of podophyllotoxin/2′(2′,6′)-(di)halogenopodophyllotoxins. Their acaricidal activities were evaluated against *Tetranychus cinnabarinus* at 48h and 72h.

The structures of molecules were designed using Chem Draw Professional software, and then, were optimized by the molecular force field (MM2) method included in ChemBioOffice software, these molecules were used for calculation of the descriptors 2D.

For the QSAR study, the activity values were transformed as follows [14]:

$$Activity= -\log c + \log it$$

Where c is the molar concentration= concentration ($\mu$g/ml) $\times$ 0.001/molecular weight

$$Log\ it= \log[\%\ inhibition/(100-\%inhibition)].$$

The Structures of compounds with their biological activities are displayed in Table 1.

**Table 1:** Structure of data set and their experimental activities against female adults of *T. Cinnabarinus*

| | Compd | R | MR (%) ( 48h) | pIC$_{50}$ (48h) | MR (%) (72h) | pIC$_{50}$ (72h) |
|---|---|---|---|---|---|---|
| | 1 | C$_6$H$_5$ | 10.7 | 8.14 | 19.9 | 8.46 |
| | 2 | (4-Me)C$_6$H$_4$ | 13.3 | 8.26 | 22.9 | 8.55 |
| | 3 | (4-Cl)C$_6$H$_4$ | 14 | 8.30 | 23.4 | 8.57 |
| | 4 | (4-F)C$_6$H$_4$ | 10.6 | 8.15 | 20.2 | 8.48 |
| | 5 | (4-NO$_2$)C$_6$H$_4$ | 7.7 | 8.02 | 15.2 | 8.35 |
| | 6 | C$_6$H$_5$ | 12.9 | 8.26 | 12.2 | 8.23 |
| | 7 | (4-Me)C$_6$H$_4$ | 9 | 8.09 | 21.8 | 8.54 |
| | 8 | (4-Cl)C$_6$H$_4$ | 3.8 | 7.71 | 29.4 | 8.73 |
| | 9 | (4-F)C$_6$H$_4$ | 9.6 | 8.13 | 22.5 | 8.57 |
| | 10 | (4-NO$_2$)C$_6$H$_4$ | 8.4 | 8.08 | 15.3 | 8.38 |
| | 11 | C$_6$H$_5$ | 9.9 | 8.15 | 20.7 | 8.53 |
| | 12 | (4-Me)C$_6$H$_4$ | 14.2 | 8.34 | 24.2 | 8.63 |
| | 13 | (4-Cl)C$_6$H$_4$ | 11.5 | 8.25 | 34 | 8.85 |
| | 14 | (4-F)C$_6$H$_4$ | 5.4 | 7.88 | 25 | 8.65 |
| | 15 | (4-NO$_2$)C$_6$H$_4$ | 7.2 | 8.03 | 18.4 | 8.50 |
| | 16 | C$_6$H$_5$ | 10.1 | 8.17 | 19 | 8.49 |
| | 17 | (4-Me)C$_6$H$_4$ | 9.2 | 8.13 | 13.3 | 8.31 |
| | 18 | (4-Cl)C$_6$H$_4$ | 9.8 | 8.18 | 22.4 | 8.60 |
| | 19 | (4-F)C$_6$H$_4$ | 12.2 | 8.27 | 20.1 | 8.53 |
| | 20 | (4-NO$_2$)C$_6$H$_4$ | 10.1 | 8.20 | 15.5 | 8.41 |
| | 21 | | 29.8 | 8.54 | 65.7 | 9.19 |

| | | | | | |
|---|---|---|---|---|---|
| | 22 | | 1.3 | 7.03 | 1.9 | 7.20 |
| | 23 | X= Cl, Y= H | 6.8 | 7.81 | 6 | 7.75 |
| | 24 | X= Y= Cl | 7.1 | 7.83 | 6.6 | 7.80 |
| | 25 | X= Br, Y= H | 5.6 | 7.76 | 5.6 | 7.76 |
| | 26 | X= Y= H | 6.9 | 7.84 | 13.3 | 8.15 |
| | 27 | X= Cl, Y= H | 9.9 | 8.04 | 11.4 | 8.11 |
| | 28 | X= Y= Cl | 14 | 8.24 | 17.8 | 8.36 |
| | 29 | X= Br, Y= H | 8.6 | 8.01 | 11.7 | 8.16 |

**2D-QSAR analysis**
**Molecular descriptors with MOE**
Before establishing 2D-QSAR models we need first to calculate molecular descriptors. This analysis was carried using Molecular Operating Environment (MOE) software (https://www.chemcomp.com.) for a total of 360 molecular descriptors.

MOE is a molecular modeling program, which is specifically designed to use several different force fields and semi-empirical and ab initio quantum mechanics calculations and calculated 2D descriptors for a set of molecules.

The 360 of 2D descriptors can be divided into three classes: 2D descriptors which are based only on atoms and physicochemical properties (molecular weight, density, some of the atomic polarisabilities), atom counts and bonds, Kier & Hall and Krappa indices, pharmacophore descriptors and finally partial charges; internal 3D descriptors (i3D) based on 3D coordinate information but not affected by orientation, this type of descriptors interested by calculation of energy, surface area, and volume; the last type is external 3D descriptors (x3D), based on 3D coordinate information but sensitive to absolute orientation) such as Van Der Waals interaction energy and principal moment of inertia. Additionally, news descriptors can be introduced in MOE using the SVL function [15].

The second step is to filter the most influential 2D-descriptors by using MOE analysis, for linking the best descriptors selected to biological activities.

**Statistical Analysis**
**Partial Least Squares (PLS)**
Partial Least Squares (PLS) is a technique that reduces descriptors molecular and to establish a linear 2D-QSAR model. Several parameters were taken to choose the best model: high value of determination coefficient ($R^2$), high Fischer's value (F-test) and low Root Mean Square Error (RMSE) [16]. The XLSTAT 2014 software packages were used for detailed statistical analysis of the 2D-QSAR model. Hence, we used the internal and external validation for the purpose to evaluate the predictive ability of the developed QSAR models.

**Internal validation**

The internal validation uses the dataset for checking the internal stability and determining the power of the model. This technique refers to the LOO (leave-one-out) method implemented in the MOE package [17].

This process consists to extract a certain number k of compounds from the initial data set (N) and to build a new model with the (N-k) using the selected descriptors [18]. This method allowed us to calculate the value of the Cross-validation coefficient ($Q^2$) by using the following equation [19]:

$$Q^2 = 1 - \left[ \frac{\Sigma_i(y_{iobs} - y_{ipred})^2}{\Sigma_i(y_{iobs} - \bar{y}_{moy})^2} \right] \quad \text{Eq. (1)}$$

Where $y_{iobs}$ and $y_{ipred}$ are the actual and predicted activities in the training set respectively while $\bar{y}_{moy}$ is the average activity of all of the molecules in the training set.

**External validation**

For the prediction of the property/activity related to several molecules, which are absent in the initial set and called test series. Several parameters of validation were evaluated such us $R^2$ (test) using the mathematical equation [19]:

$$R^2_{Pred} = 1 - \left[ \frac{\Sigma_i(y_{iobs} - y_{ipred})^2}{\Sigma_i(y_{iobs} - \bar{y}_{moy})^2} \right] \quad \text{Eq. (2)}$$

Where $y_{iobs}$ and $y_{ipred}$ are the actual and predicted activities of the ith molecule in the test set, respectively, and $\bar{y}_{moy}$ is the average activity of all of the molecules in the training set.

**Applicability Domain**

The Applicability Domain (AD) defines the area in which a compound can be predicted with confidence. The AD corresponds to the region of the chemical space including the compounds in the learning set and similar compounds that are close in this same space. There are several methods for the determination of the applicability domain of a QSAR model [20]. The most method is the leverage approach which based on the determination of the leverage value (hi) of each compound i.

$$h_i = x_i^T(X^T X)^{-1}x_i \quad (i = 1, …, n), \qquad \text{Eq. (3)}$$

Where x is the vector of descriptors of a query compound, and X is the matrix formed with rows corresponding to the descriptors of molecules from the training set [21].

The Applicability Domain (DA) defines the area in which a compound can be predicted with confidence. If a compound had the leverage value higher than the critical value (h*) (h* = 3(p+1)/n, where p is the number of descriptors and n is several compounds in the training set) this compound is considered outside of the model developed. If hi<h*, the probability of agreement between the measured and predicted values of compound "i" is as high as that of the compounds in the database. Compounds with hi>h* strengthen the model when they belong to the training set, but will otherwise have dubious predicted values without necessarily being outliers, as the residuals may be low [22].

**Results and Discussion**

**Model 1: 2D-QSAR Model at 48h**

**Partial least squares (PLS)**

The PLS analysis was applied in this study to select the most significant descriptors on the activities and also used to establish a linear QSAR model. The training set contains 23 molecules and 6 molecules taken as test molecules (22, 28, 26, 8, 14, and 12).

The best PLS model obtained form 2D-QSAR study presented in Eq.4:

$$pIC_{50} = 7.53945 - 0.22617*\text{lip-don} + 1.67343*\text{PEOE-VSA-FNEG} \qquad \text{Eq. (4)}$$

$$N = 23, RMSE = 0.068, r^2 = 0.83, r^2_{adjusted} = 0.81, F_{model} = 49.53, p < 0.0001, \alpha = 1\%$$

$$q^2_{LOO} = 0.79, RMSE_{LOO} = 0.07, r^2_{test} = 0.64$$

N represents the number of compounds in training set.

The PLS model was established using two descriptors significant which shows less correlation between them (Table 2). The select descriptors had an explained variance of 83% ($r^2 = 0.83$) with a lowest error RMSE= 0.068. Furthermore, the F value of the model is 49.53 ($F_{model} > F_{tabulate}$), this result suggests that the model is highly significant in the levels $\alpha = 1\%$.

The observed pIC50 and predicted pIC50 activities of the training and the test sets are displayed in Table 3 and graphically represented in figure 1a, Figure 1bgives the contribution of descriptors, as shown in this figure and the eq.4the lip-don decreases the values of biological activity with the negative sign (**-0.22**). Moreover the PEOE-VSA-FNEG with a positive sign (+**1.67**)is favorable to increase biological activity.

The multi-colinearity between the descriptors used in the 2D-QSAR model was verified by calculating the Variance Inflation Factor VIF. The VIF is defined as $1/(1-r^2)$, where $r^2$ is the coefficient of inter-correlation between a descriptor and all other descriptors of the proposed model[23]. A VIF value greater than 5 indicates that the model is unstable and a value between 1 and 4 indicates that the model is acceptable. These values are collected in Table 4 and their analyses show that the descriptors used in the proposed models have a very low correlation which means the model has good stability.

The significance of each descriptor was examined by the Student test, which showed us that the descriptors selected are very significant at the level $\alpha = 1\%$ (Table 4).

To estimate the stability of the obtained model, the correlation coefficients of the cross-validation ($q^2_{LOO}$) and the external validation ($R^2_{test}$) were calculated. The high value of $q^2_{LOO}$($q^2_{LOO} = 0.79$) and the low associated error $RMSE_{LOO}$($RMSE_{LOO} = 0.07$) means that the 2D-QSAR model is stable. Moreover, the high value of the $r^2_{test}$($r^2_{test} = 0.64$) indicates great predictability of the obtained 2D-QSAR Model.

## Interpretation of the Selected Descriptors

To understand and explain the variation of acaricidal activities for all compounds of data set, it's necessary to define each descriptor selected in the final model.

## Lip-don

It represents the number of OH and NH atoms. The negative sign of the lip-don in Eq.4 indicates that the increase in the number of H-Bond in compounds decreases the values of $pIC_{50}$. As seen from Table 3 and the structure from Table 1, we could observe that the compound (23, 24 and25)which are characterized by group OH with the highest values of lip-don (Lip-don=1) have the lowest pIC50 (7.79, 7.79 and 7.83) successively; however, the other compounds have the high pIC50 values, they are greater than 8 with a group of OH.

## PEOE-VSA-FNEG

Partial Equalization of Orbital Electronegativities – Vander Waals surface Area- Fractional Negative or fractional negative VDW surface area. In the model 1, the coefficient of this descriptor has a positive weight, which demonstrates the significance of the molecular surface area bearing a fractional negative charge for molecules inhibitor against the female adults of *T. cinnabarinus*. This involves an electrostatic interaction between electronegative substituents having a fractional negative charge in the aromatic ring and a group carrying a fractional positive charge present in the enzyme as TcPMCA1 and Ca2+-ATPase which are regulated when the spider is exposed to molecules known by their acaricidal activity[3].

As shown in Table 3, the compounds 13 and 21 have good values of PEOE-VSA-FNEG (0.40 and 0.60) respectively this is the reason why the two compounds have high values of pIC50 (8.25 and 8.54) successively. The molecule 21 was used as a positive control which means it has a high mortality rate and the best acaricidal activity. Also, the molecule 13 showed the most promising acaricidal activity.

**Applicability Domain (AD)**

The applicability domain of the established 2D-QSAR model was checking by William's plot (Figure2) by using Minitab software, which is based on the calculation of the leverage value (hi) of each compound i and a leverage threshold h*, in this study h*= 0.39. According to this figure, we can note that the compounds 23, 24, 25 have a lever value greater than h* for this reason these compounds are considered as X outlier. These three molecules which are outside of the applicability domain are characterized by the lowest acaricidal activity with a group of OH.

**Table 2:** Correlation Matrix between the two descriptors obtained by model 1 using the PLS analysis.

| Variables | lip-don | PEOE-VSA-FNEG | $pIC_{50}$ |
|---|---|---|---|
| lip-don | 1.0000 | | |
| PEOE-VSA-FNEG | -0.4735 | 1.0000 | |
| $pIC_{50}$ | -0.7401 | 0.8201 | 1.0000 |

]**Table 3**: Observed, predicted activities, and residuals for model 1. And descriptors values for training and test set

| | Compounds | $pIC_{50}$ | préd($pIC_{50}$) | Residuals | Lip-don | PEOE-VSA-FNEG |
|---|---|---|---|---|---|---|
| **Training set** | **23** | 7.8162 | 7.7916 | 0.0246 | 1 | 0.28 |
| | **24** | 7.8364 | 7.7916 | 0.0448 | 1 | 0.28 |
| | **25** | 7.7674 | 7.8369 | -0.0695 | 1 | 0.31 |
| | **27** | 8.0434 | 8.0686 | -0.0252 | 0 | 0.31 |
| | **29** | 8.0129 | 8.1066 | -0.0937 | 0 | 0.33 |
| | **1** | 8.1486 | 8.1293 | 0.0193 | 0 | 0.35 |
| | **2** | 8.2662 | 8.1375 | 0.1287 | 0 | 0.35 |
| | **3** | 8.3065 | 8.1645 | 0.1420 | 0 | 0.37 |
| | **4** | 8.1572 | 8.1231 | 0.0341 | 0 | 0.34 |
| | **5** | 8.0234 | 8.0898 | -0.0664 | 0 | 0.32 |
| | **6** | 8.2654 | 8.1645 | 0.1009 | 0 | 0.37 |
| | **7** | 8.097 | 8.1713 | -0.0743 | 0 | 0.37 |
| | **9** | 8.1333 | 8.1581 | -0.0248 | 0 | 0.36 |
| | **10** | 8.0876 | 8.1245 | -0.0369 | 0 | 0.34 |
| | **11** | 8.1592 | 8.1906 | -0.0314 | 0 | 0.38 |
| | **13** | 8.2542 | 8.2215 | 0.0327 | 0 | 0.40 |
| | **15** | 8.0368 | 8.1504 | -0.1136 | 0 | 0.36 |
| | **16** | 8.1754 | 8.1943 | -0.0189 | 0 | 0.39 |
| | **17** | 8.1396 | 8.2001 | -0.0605 | 0 | 0.39 |
| | **18** | 8.1827 | 8.2254 | -0.0427 | 0 | 0.40 |
| | **19** | 8.2792 | 8.1879 | 0.0913 | 0 | 0.38 |
| | **20** | 8.2038 | 8.1536 | 0.0502 | 0 | 0.36 |
| | **21** | 8.5431 | 8.5537 | -0.0106 | 0 | 0.60 |
| **Test set** | **22** | 7.038 | 7.735 | -0.6970 | **1** | 0.25 |
| | **8** | 7.714 | 8.197 | -0.4830 | **0** | 0.39 |
| | **14** | 7.886 | 8.184 | -0.2980 | **0** | 0.38 |
| | **26** | 7.8416 | 8.0223 | -0.1807 | **0** | 0.28 |
| | **12** | 8.3462 | 8.1965 | 0.1497 | **0** | 0.39 |
| | **28** | 8.2429 | 8.1034 | 0.1395 | **0** | 0.33 |
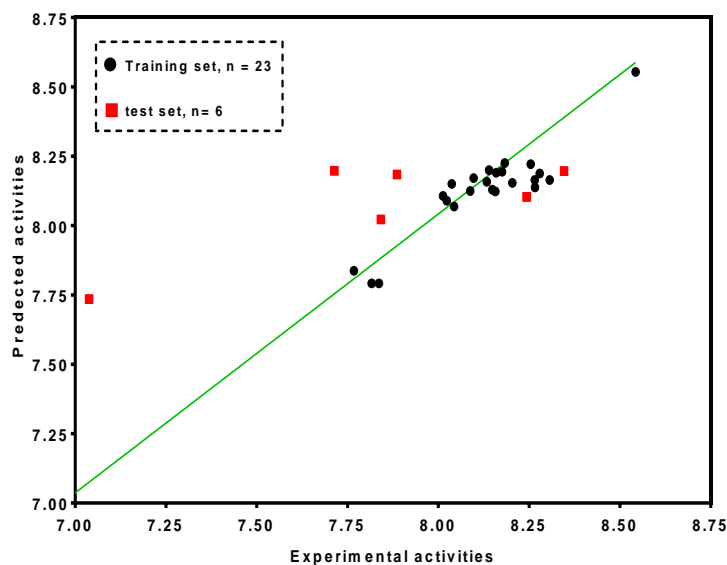
1a
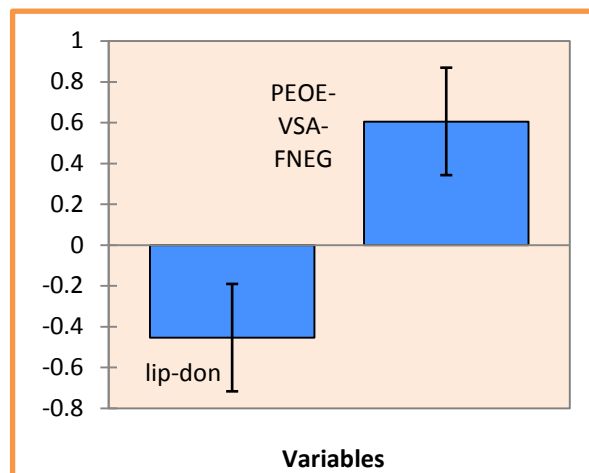


1b

*Figure1: 1a: Correlations of observed and predicted activities (training set in black and test set in red). 1b: Contribution of descriptors obtained by model 1.*

**Table 4:** Student test and VIF values of the selected descriptors in model 1.

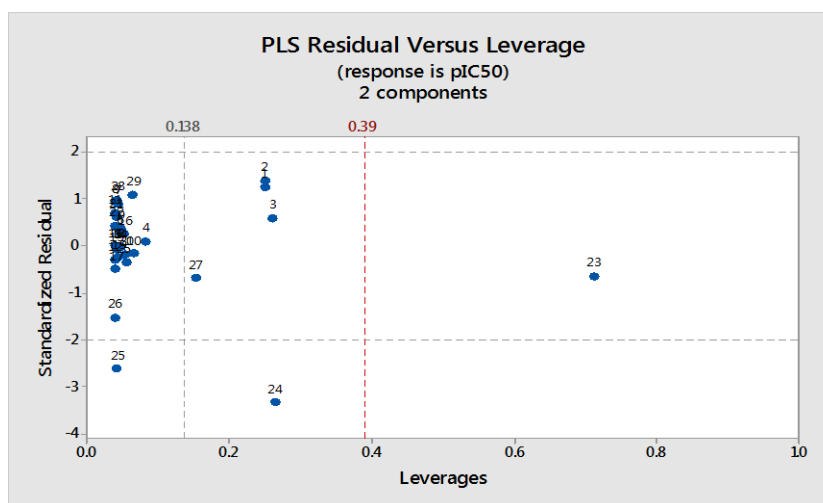| Descriptors | coefficient | Erreur standard | t | p-value | VIF |
|---|---|---|---|---|---|
| **lip-don** | -0.2262 | 0.0519 | -4.3587 | 0.0003 | 1.28 |
| **PEOE-VSA-FNEG** | 1.6734 | 0.2876 | 5.8182 | < 0.0001 | 1.28 |



*Figure 2: William's plot of model 1, training, and test set.*

**Model 2: 2D-QSAR Model at 72h**

**Partial least squares (PLS)**

Also for model 2, the PLS analysis was used to link the biological activities with descriptors. The data set was split into a training set contains 24 compounds and the test set contains 5 compounds (24, 13, 17, 8, and 6). The second model obtained from 2D-QSAR study presented in Eq. 5:

$$\text{pIC}_{50} = 7.17318 - 0.60155 \cdot \text{lip-don} + 3.52035 \cdot \text{PEOE-VSA-FNEG} \qquad \text{Eq. (5)}$$

$$N = 24, \text{RMSE} = 0.11, r^2 = 0.92, r^2_{adjusted} = 0.91, F_{model} = 125.50 \ p < 0.0001, \alpha = 1\%$$

$$q^2_{LOO} = 0.83, \text{RMSE}_{LOO} = 0.16, r^2_{test} = 0.67$$

N represents the number of compounds in the training set.

Model 2 explains 92% of the variance in the observed activity using two descriptors significant (lip-don and PEOE-VSA-FNEG)with inter-correlation $|r| < 0.6$(Table 5) and with the lowest error RMSE= 0.11. Moreover, the F value of the model is 125.50 ($F_{model} > F_{tabulate}$), this result indicates that the model is highly significant in the levels $\alpha = 1\%$. The predictive ability of model 2 is also good as measured by its cross-validated $q^2_{LOO} = 0.83$ with low error $\text{RMSE}_{LOO} = 0.16$. The value of the $r^2_{test}$($r^2_{test} = 0.67$) demonstrated the robustness and external high prediction of the 2D-QSAR model.

The observed predicted activities, residual, and values of descriptors used in the QSAR models are depicted in table 6. The correlation between the predictions and the observed values for all molecules is shown in Figure 2a.

As in the previous model, the two descriptors lip-don and PEOE-VSA-FNEG are the most influential on acaricidal activity. According to the eq.5 and Figure 2b, the first descriptor lip-don influences negatively on the activity with a value of (-0.60) which means decreasing activity. The second descriptor PEOE-VSA-FNEG influences positively on the acaricidal activity with a value of (+3.52) which indicates increasing activity.

The Student t-test indicates that the two descriptors in model 2 are highly significant at the level $\alpha = 1\%$ as shown in Table 7. The presence of multicollinearity is detected by calculating VIF values. In model 2, no multicollinearity seems to be present. As the models satisfy the diagnostics of the molecular descriptors, they can be considered as robust and acceptable for further consideration (Table 7).

**Applicability Domain (AD)**

Applicability domain was used in this model to analyze the similarity of the molecules, first of all, we calculated critical value h* (h*= 0.37), From the Williams plot (Figure 3), it can be seen that 2 compounds (24 and 1) are out of applicability domain with h > h*. Hence, the molecule 24 was only a mortality rate of less than 7% at 72h

**Table 5:** The correlation matrix of two descriptors by model 2.

| Variables | lip-don | PEOE-VSA-FNEG | pIC$_{50}$ |
|---|---|---|---|
| **lip-don** | 1.0000 | | |
| **PEOE-VSA-FNEG** | -0.4660 | 1.0000 | |
| **pIC$_{50}$** | -0.8047 | 0.8392 | 1.0000 |

**Table 6:** Values of significant molecular descriptors, experimental, predicted acaricidal activities, and residuals from model 2.

| | Compounds | pIC$_{50}$ | préd(pIC$_{50}$) | Residuals | lip-don | PEOE-VSA-FNEG |
|---|---|---|---|---|---|---|
| **Training set** | 1 | 7.2 | 7.4591 | -0.2591 | 1 | 0.2521 |
| | 11 | 7.75 | 7.5777 | 0.1723 | 1 | 0.2858 |
| | 13 | 7.76 | 7.6731 | 0.0869 | 1 | 0.3129 |
| | 14 | 8.15 | 8.1888 | -0.0388 | 0 | 0.2885 |
| | 15 | 8.11 | 8.2863 | -0.1763 | 0 | 0.3162 |
| | 16 | 8.36 | 8.3595 | 0.0005 | 0 | 0.337 |
| | 17 | 8.16 | 8.3662 | -0.2062 | 0 | 0.3389 |
| | Ia | 8.46 | 8.4141 | 0.0459 | 0 | 0.3525 |
| | Ib | 8.55 | 8.4314 | 0.1186 | 0 | 0.3574 |
| | Ic | 8.57 | 8.4880 | 0.0820 | 0 | 0.3735 |

*The Pharmaceutical and Chemical Journal*

|          |              |      |        |         |   |        |
|----------|--------------|------|--------|---------|---|--------|
|          | **Id**       | 8.48 | 8.4011 | 0.0789  | 0 | 0.3488 |
|          | **Ie**       | 8.35 | 8.3310 | 0.0190  | 0 | 0.3289 |
|          | **IIb**      | 8.54 | 8.5025 | 0.0375  | 0 | 0.3776 |
|          | **IId**      | 8.57 | 8.4747 | 0.0953  | 0 | 0.3697 |
|          | **IIe**      | 8.38 | 8.4039 | -0.0239 | 0 | 0.3496 |
|          | **IIIa**     | 8.53 | 8.5430 | -0.0130 | 0 | 0.3891 |
|          | **IIIb**     | 8.63 | 8.5553 | 0.0747  | 0 | 0.3926 |
|          | **IIId**     | 8.65 | 8.5296 | 0.1204  | 0 | 0.3853 |
|          | **IIIe**     | 8.5  | 8.4585 | 0.0415  | 0 | 0.3651 |
|          | **IVa**      | 8.49 | 8.5507 | -0.0607 | 0 | 0.3913 |
|          | **IVc**      | 8.6  | 8.6162 | -0.0162 | 0 | 0.4099 |
|          | **IVd**      | 8.53 | 8.5373 | -0.0073 | 0 | 0.3875 |
|          | **IVe**      | 8.41 | 8.4652 | -0.0552 | 0 | 0.367  |
|          | **spirodiclofen** | 9.19 | 9.3069 | -0.1169 | 0 | 0.6061 |
| **Test set** | **12**   | 7.8  | 7.57   | 0.23    | 1 | 0.2858 |
|          | **Iia**      | 8.23 | 8.48   | -0.25   | 0 | 0.3735 |
|          | **Iic**      | 8.73 | 8.55   | 0.18    | 0 | 0.3931 |
|          | **IIIc**     | 8.85 | 8.6    | 0.25    | 0 | 0.4076 |
|          | **Ivb**      | 8.31 | 8.56   | -0.25   | 0 | 0.3948 |
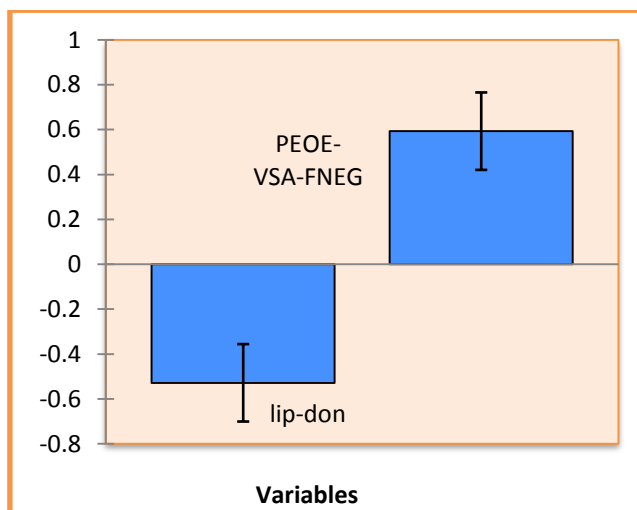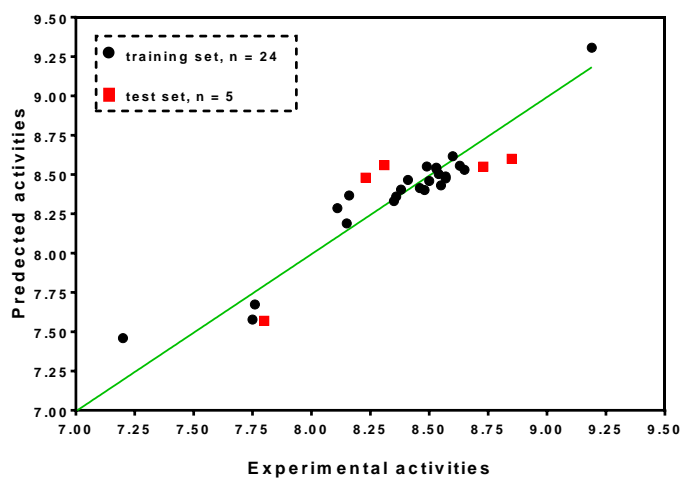


*Figure 2:2a: Plot of experimental versus predicted pIC$_{50}$ values of training sets and test sets against female T. cinnabarinus (training set in black and test set in red). 2b: Contribution of descriptors obtained by model 2.*

**Table 7:** Student test and VIF values of the selected descriptors in model 2.

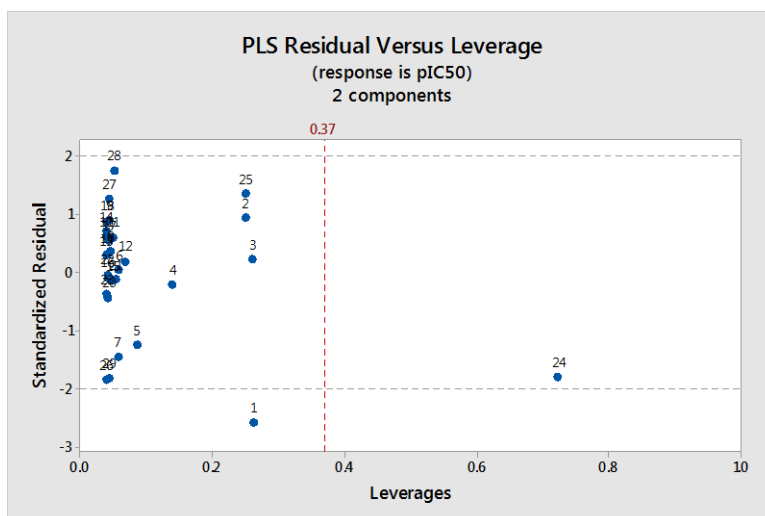| Descriptors    | Coefficient | Erreur standard | t       | p-value   | VIF  |
|----------------|-------------|-----------------|---------|-----------|------|
| lip-don        | -0.6016     | 0.0780          | -7.7098 | < 0.0001  | 1.27 |
| PEOE-VSA-FNEG  | 3.5204      | 0.4068          | 8.6534  | < 0.0001  | 1.27 |

*Figure 3: William's plot of the PLS model N°2, training, and test set.*

**Conclusion**

The present work aimed to develop a 2D-QSAR study of 29 molecules of isoxazoline-containing podophyllotoxin/2′(2′,6′)-(di)halogenopodophyllotoxin derivative against the female adults of *T. cinnabarinus*. The 2D-QSAR models were established by using MOE descriptors which show the importance of lip-don and PEOE-VSA-FNEG to describe the acaricidal activity.

These results allowed us to better understand the nature and the structures of molecules used in pesticides to protect crops and a better estimation of the adverse effects of the molecules in the environment.

**References**

[1]. Sertkaya, E., Kaya, K. & Soylu, S. (2010). Acaricidal Activities of the Essential Oils from Several Medicinal Plants against the Carmine Spider Mite (Tetranychus Cinnabarinus Boisd.) (Acarina: Tetranychidae). *Industrial Crops and Products*, *31*(1), 107–112.

[2]. Shi, G.L., Zhao, L.L., Liu, S.Q., Cao, H., Clarke, S.R. & Sun, J.H. (2006). Acaricidal Activities of Extracts of Kochia Scoparia against Tetranychus Urticae, Tetranychus Cinnabarinus, and Tetranychus Viennensis (Acari: Tetranychidae). *Journal of Economic Entomology*, *99*(3), 858–863.

[3]. Luo, J., Lai, T., Guo, T., Chen, F., Zhang, L., Ding, W. & Zhang, Y. (2018). Synthesis and Acaricidal Activities of Scopoletin Phenolic Ether Derivatives: Qsar, Molecular Docking Study and in Silico Adme Predictions. *Molecules*, *23*(5).

[4]. Lemaire, É. (2012). LA LUTTE BIOLOGIQUE CONTRE LES TÉTRANYQUES DANS LES FRAMBOISIÈRES SOUS TUNNELS.

[5]. Mo, T.L. & Liu, T.X. (2006). Biology, Life Table and Predation of Feltiella Acarisuga (Diptera: Cecidomyiidae) Feeding on Tetranychus Cinnabarinus Eggs (Acari: Tetranychidae). *Biological Control*, *39*(3), 418–426.

[6]. Ma, X., Zhang, Y., Zhou, H., Liu, J., Guo, F., Luo, J., Ding, W. & Zhang, Y. (2020).Silencing T-Type Voltage-Gated Calcium Channel Gene Reduces the Sensitivity of Tetranychus Cinnabarinus (Boisduval) to Scopoletin. *Comparative Biochemistry and Physiology Part - C: Toxicology and Pharmacology*, *Elsevier*, *227*, 108644.

[7]. Lin, H., Chuan-hua, X., Jin-jun, W., Ming, L., Wen-cai, L. & Zhi-mo, Z. (2009). Resistance Selection and Biochemical Mechanism of Resistance to Two Acaricides in Tetranychus Cinnabarinus (Boiduval). *Pesticide Biochemistry and Physiology*, *Elsevier Inc.*, *93*(1), 47–52.

[8]. Wang, Y.N., Shi, G.L., Zhao, L.L., Liu, S.Q., Yu, T.Q. & Clarke, S.R. (2007). Acaricidal Activity of

Juglans Regia Leaf Extracts on Tetranychus Viennensis and Tetranychus Cinnabarinus ( Acari : Tetranychidae ). 1298–1303.

[9]. Lv, M., Liu, G., Jia, M. & Xu, H. (2018). Synthesis of Matrinic Amide Derivatives Containing 1,3,4-Thiadiazole Scaffold as Insecticidal/Acaricidal Agents. *Bioorganic Chemistry*.

[10]. Chang, H., Kim, H.J. & Chun, H.S. (2007). Quantitative Structure − Activity Relationship ( QSAR ) for Neuroprotective Activity of Terpenoids. *80*, 835–841.

[11]. Hamadache, M., Benkortbi, O., Hanini, S. & Amrane, A. (2016). A Quantitative Structure Activity Relationship for Acute Oral Toxicity of Pesticides on Rats : Validation , Domain of Application and Prediction. *Journal of Hazardous Materials*, *Elsevier B.V.*, *303*, 28–40.

[12]. Jing, G., Zhou, Z. & Zhuo, J. (2012). Chemosphere Quantitative Structure – Activity Relationship (QSAR) Study of Toxicity of Quaternary Ammonium Compounds on Chlorella Pyrenoidosa and Scenedesmus Quadricauda. *Chemosphere*, *Elsevier Ltd*, *86*(1), 76–82.

[13]. Yang, R., Zhang, Y. & Xu, H. (2018). Synthesis of Novel Isoxazoline-Containing Podophyllotoxin/2′(2′,6′)-(Di)Halogenopodophyllotoxin Derivatives and Their Insecticidal/Acaricidal Activities. *Bioorganic and Medicinal Chemistry Letters*, *28(8)*, 1410–1416.

[14]. Manvar, A., Malde, A., Verma, J., Virsodia, V., Mishra, A., Upadhyay, K., Acharya, H., Coutinho, E. & Shah, A. (2008). Synthesis, Anti-Tubercular Activity and 3D-QSAR Study of Coumarin-4-Acetic Acid Benzylidene Hydrazides. *European Journal of Medicinal Chemistry*, *Elsevier Masson SAS*, *43*(11), 2395–2403.

[15]. Vilar, S., Cozza, G. & Moro, S. (2008). Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Current Topics in Medicinal Chemistry*, *8*(18), 1555–1572.

[16]. Baviskar, B.A., Deore, S.L. & Rathi, P. (2019). QSAR Studies of Triterpenoid Saponin Analogues for Nematicidal Activity. *10*(1), 15–20.

[17]. Olasupo, S.B., Uzairu, A., Shallangwa, G. & Uba, S. (2020). QSAR Modeling, Molecular Docking and ADMET/Pharmacokinetic Studies: A Chemometrics Approach to Search for Novel Inhibitors of Norepinephrine Transporter as Potent Antipsychotic Drugs. *Journal of the Iranian Chemical Society*, Springer Berlin Heidelberg.

[18]. El Aissouq, A., Toufik, H., Stitou, M., Ouammou, A. & Lamchouri, F. (2019). In Silico Design of Novel Tetra-Substituted Pyridinylimidazoles Derivatives as c-Jun N-Terminal Kinase-3 Inhibitors, Using 2D/3D-QSAR Studies, Molecular Docking and ADMET Prediction. *International Journal of Peptide Research and Therapeutics*, *Springer Netherlands*.

[19]. Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C.P. & Agrawal, R.K. (2011). Validation of QSAR Models - Strategies and Importance. *International Journal of Drug Design and Disocovery*, *2*(3), 511–519.

[20]. Sahigara, F., Mansouri, K., Ballabio, D., Mauri, A., Consonni, V. &Todeschini, R. (2012).Comparison of Different Approaches to Define the Applicability Domain of QSAR Models. *Molecules*, *17(5)*, 4791–4810.

[21]. Rybka, M., Mercader, A.G. & Castro, E.A. (2014). Chemometrics and Intelligent Laboratory Systems Predictive QSAR Study of Chalcone Derivatives Cytotoxicity Activity against HT-29 Human Colon Adenocarcinoma Cell Lines. *Chemometrics and Intelligent Laboratory Systems*, Elsevier B.V., *132*, 18–29.

[22]. Chtita, S., Aouidate, A., Belhassan, A., Ousaa, A., Taourati, A.I., Elidrissi, B., Ghamali, M., Bouachrine, M. & Lakhlifi, T. (2020). QSAR Study of N -Substituted Oseltamivir Derivatives as Potent Avian Influenza Virus H5N1 Inhibitors Using Quantum Chemical Descriptors and Statistical Methods. *New Journal of Chemistry*, *44*(5), 1747–1760.

[23]. Jaiswal, M., Khadikar, P. V., Scozzafava, A. & Supuran, C.T. (2004). Carbonic Anhydrase Inhibitors: The First QSAR Study on Inhibition of Tumor-Associated Isoenzyme IX with Aromatic and Heterocyclic Sulfonamides. *Bioorganic and Medicinal Chemistry Letters*, *14(12)*, 3283–3290.