Journal of Computational Methods in Molecular Design, 2015, 5 (2):83-91

Scholars Research Library
(http://scholarsresearchlibrary.com/archive.html)

ISSN : 2231- 3176
CODEN (USA): JCMMDA

The biological activity of pyrazinecarboxamides derivatives as an herbicidal agent: combining DFT and QSAR studies

B. Elidrissi[a*], A. Ousaa[a], M. Ghamali[a], S. Chtita[a], M. A. Ajana[a], M. Bouachrine[b] and T. Lakhlifi[a]

[a]*Molecular Chemistry and Natural Substances Laboratory, Faculty of Science, University Moulay Ismail, Meknes, Morocco*
[b]*ESTM, University Moulay Ismail, Meknes, Morocco*

ABSTRACT

A set of nineteen Pyrazinecarboxamides derivatives with herbicidal activity was subjected to the two dimensional quantitative structure activity relationships studies. This work was conducted using the principal component analysis (PCA) method, the multiple linear regression method (MLR), the multiple non-linear regressions (MNLR) and the artificial neural network (ANN). The predicted results of various study compounds afford reliable prediction of IC_{50} with respect to experimental data. Density functional theory (DFT) calculations have been carried out in order to get insights into the structure, chemical reactivity and property information for the series of study compounds. This study shows that the PCA, MLR and MNLR have served also to predict activities, but when compared with the results given by the ANN (R^2= 0.994) , we realized that the predictions fulfilled by this latter were more effective as indicated by the value of cross validated squared correlation coefficient (R^2_{CV} = 0.998). Thus, this validated model brings important structural insight to aid the design of novel herbicidal agents.

Keywords: DFT study, QSAR, pyrazinecorboxamides, herbicidal.

INTRODUCTION

Herbicides are inhibitors of individual metabolic processes in plants, used in agriculture as a selective means of defense against weeds. Weeds compete with crops for sunshine, water, nutrients, and physical space and are thus capable of greatly influencing the growth of crops and undermining both crop quality and yield. Also, many weeds are the harbor or nest of pathogens, viruses, and pests, which may result in the occurrence and spread of plant diseases and insect pests in crops. Herbicides, as the main weed control tool, play a very important role in modern agriculture. Crop protection continually needs the discovery of novel herbicides.

Since the discovery of the herbicide 2, 4-dichlorophenoxyacetic acid (2, 4-D) and 2, 4, 5-trichlorophenoxyacetic acid (2, 4, 5-T) Fig. 1, the agrochemical industry has successfully developed a wide array of herbicides with various chemical structures and modes of action [1]. However, an inevitable problem associated with the use of herbicides is the occurrence of herbicide resistant weeds [2]. Therefore, it is necessary to develop efficient herbicides with novel structures or modes of action to overcome the resistance of weeds.

Analogues of pyrazinecarboxamides belong to the group of herbicides inhibiting the photosynthetic electron transport in spinch chloroplast. On the other hand, the pyrazinamide ring system has received much attention in biologically active molecules, such as potent inhibitors of mycobacterium and fungal. In this study, we have modeled the inhibition in spinach chloroplast (IC_{50}) of a series of pyrazine-2-corboxylic acid amides derivative

(Table 1), using several statistical tools, principal components analysis (PCA), multiple linear regression (MLR), multiple non-linear regression (MNLR) and artificial neural network (ANN) calculations. The objectives of this work are to develop predictive QSAR models for the toxicity of our studied molecules. On the other hand, several quantum chemical methods and Quantum-chemistry calculations have been performed in order to study the molecular structure and electronic properties [3,4]. The geometry as well as the nature of their molecular orbital, HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) is involved in the properties of biological activity of organic compounds. The more relevant molecular properties were calculated, these properties are the highest occupied molecular orbital energy $\mathbf{E_{HOMO}}$, the lowest unoccupied molecular orbital energy $\mathbf{E_{LUMO}}$, energy gap $\mathbf{\Delta E}$, dipole moment $\boldsymbol{\mu}$, the total energy $\mathbf{E_T}$, the activation energy $\mathbf{E_a}$ and the absorption maximum $\boldsymbol{\lambda_{max}}$and factor of oscillation $\mathbf{f_{(SO)}}$.

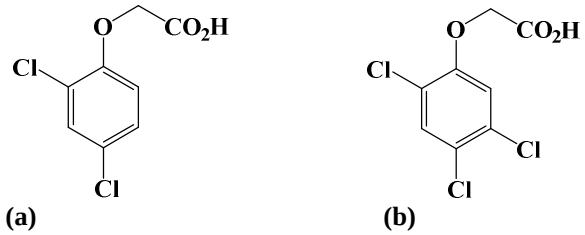


Figure 1. Structure of (a) 2, 4-D and (b) 2, 4, 5-T

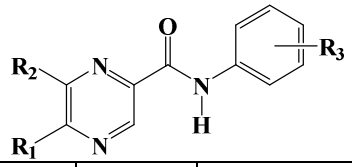
MATERIALS AND METHODS

EXPERIMENTAL DATA

A dataset of 19 compounds was taken from the published pyrazinecarboxamides derivatives as an herbicidal agent [5, 6]. The activity under investigation is the inhibition in spinach chloroplast (IC_{50}) of a series of pyrazine-2-corboxylic acid amides derivatives by 50% (IC_{50}).

The structures and their herbicidal activities are listed in Table1. The inhibitory activity IC_{50} (µmol/L) values were converted to logarithmic form (pIC_{50}) and used as dependent variables in the 3D-QSAR analyses.

Table1. Observed pIC_{50} of the pyrazinecarboxamide derivatives as herbicidal agents



Mol. N°	R_1	R_2	R_3	pIC_{50}(obs.)
1*	H	H	2-Cl-5-OH	2.86
2*	H	H	4-F	2.68
3*	H	H	2-CF_3	2.57
4	H	H	3-CF_3	2.11
5	H	H	4-CH_3	3.17
6	Cl	H	2-Cl-5-OH	2.79
7	Cl	H	4-F	2.58
8	Cl	H	2-CF_3	2.74
9	Cl	H	3-CF_3	2.36
10	Cl	H	4-CH_3	3.18
11	H	$(CH_3)_3C$	4-F	2.72
12	H	$(CH_3)_3C$	2-CF_3	1.74
13	H	$(CH_3)_3C$	3-CF_3	2.45
14	H	$(CH_3)_3C$	4-CH_3	2.21
15	Cl	$(CH_3)_3C$	2-Cl-5-OH	2.79
16	Cl	$(CH_3)_3C$	4-F	2.01
17	Cl	$(CH_3)_3C$	2-CF_3	2.31
18	Cl	$(CH_3)_3C$	3-CF_3	2.24
19	Cl	$(CH_3)_3C$	4-CH_3	1.86

*test set

CALCULATION OF MOLECULAR DESCRIPTORS

CALCULATION OF DESCRIPTORS USING GAUSSIAN 03W

DFT (density functional theory) methods were used in this study. These methods have become very popular in recent years because they can reach similar precision to other methods in less time and less cost from the computational point of view. In agreement with the DFT results, energy of the fundamental state of a polyelectronic system can be expressed through the total electronic density, and in fact, the use of electronic density instead of wave function for calculating the energy constitutes the fundamental base of DFT [7,8] using the B3LYP functional [9] and a 6-31G(d) basis set. The B3LYP, a version of DFT method, uses Becke's three-parameter functional (B3) and includes a mixture of HF with DFT exchange terms associated with the gradient corrected correlation functional of Lee, Yang and Parr (LYP). The geometry of all species under investigation was determined by optimizing all geometrical variables without any symmetry constraints.

The 3D structures of the molecules were generated using the Gauss View 3.0, and then, all calculations were performed using Gaussian 03W program series, Geometry optimization of nineteen compounds was carried out by B3LYP method employing 6–31G (d) basis set.

CALCULATION OF DESCRIPTORS USING ACD/CHEMSKETCH

ChemSketch program (Demo version 10.0) [10] was employed to calculate the others molecular descriptors, Molar Volume (MV (cm^3)), Molecular Weight (MW), Molar Refractivity (MR (cm^3)), Parachor (Pc (cm^3)), Density (D (g/cm^3)), Refractive Index (n) [11].

STATISTICAL ANALYSIS

PRINCIPAL COMPONENTS ANALYSIS (PCA)

The molecules of pyrazinecarboxamide derivatives (1 to 19) were studied by statistical methods based on the principal component analysis (PCA) [11] using the software XLSTAT 2009.

This is an essentially a descriptive statistical method which aims to present, in graphic form, the maximum informations contained in the data table 1.

PCA is a statistical technique useful for summarizing all the informations encoded in the structures of compounds. It is also very helpful for understanding the distribution of the compounds.

MULTIPLE LINEAR REGRESSIONS (MLR)

The multiple linear regression statistic technique is used to study the relation between one dependent variable and several independent variables. It is a mathematic technique that minimizes differences between actual and predicted values. The multiple linear regression model (MLR) [9] was generated using the software XLSTAT 2009, to predict IC_{50}. It has served also to select the descriptors used as the input parameters for a back propagation network (ANN).

ARTIFICIAL NEURAL NETWORKS (ANNS)

The ANNs analysis was performed with the use of Mathlab software v 2008a Neural Fitting tool (nftool) toolbox on a data set of pyrazinecarboxamide derivatives IC_{50} activity [12].

A number of individual models of ANN were designed built up and trained. Generally the network was built for three layers; one input layer, one hidden layer and one output layer were considered [13]. The input layer consisted of fifteen artificial neurons of linear activation function. The number of artificial neural in the hidden layer was adjusted experimentally. The hidden layer consisted of 8 artificial neural. One neuron formed the output layer of sigmoid function activation.

The data subjected to ANN analysis were randomly divided into three sets: a learning set, a validation set and a testing set. Prior to that, the whole data set was scaled within the 0–1 range.

The set of pyrazinecarboxamide derivatives of IC_{50} activity [14] were subjected to the ANN analysis. First, forth learning set of compounds, 13 inhibitors were used. ANN models were designed, built and trained. The learning set of data is used in ANNs to recognize the relationship between the input and output data. Then for the revision of the ANN model designed and selected, the validation set of 3 compounds was used. Testing set with 3 compounds was provided to be an independent evaluation of the ANN model performance for the finally applied network. In this study, we selected the Sigmoid as a basis function [15]. The operation of the output layer is linear, which is given as below:

$$y_k(X) = \sum_{j=1}^{n_k} w_{kj} h_j(X) + b_k \qquad (1)$$

Where y_k is the k^{th} output layer unit for the input vector X, w_{kj} is the weight connection between the k^{th} output unit and the j^{th} hidden layer unit and b_k is the bias that allows a transfer function "non-zero" given by the following equation:

$$\text{Bias} = \sum(\bar{y} - y) \qquad (2)$$

where y is the measured value and $\bar{y}$ is the value predicted by the model.
The accuracy of the model was mainly evaluated by the root mean square error (RMSE). Formula is given as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}.\sum_{i=1}^{n}(p_{exp} - p_{pred})^2} \qquad (3)$$

where n = number of compounds, p_{exp} = experimental value, p_{pred} = predicted value and summation is of overall patterns in the analyzed data set [16,17]. The scripts were run on a personal PC.

RESULTS AND DISCUSSION

This study was carried for a series of 19 of pyrazinecarboxamide derivatives, in order to determine a quantitative relationship between structure and herbicidal activity IC_{50}. Table 2 shows the values of the calculated parameters obtained by DFT/B3LYP 6-31G* optimization and ACD/ChemSketch program of the studied compounds.

Table 2. Values of the calculated parameters obtained by DFT/B3LYP 6-31G* optimization and ACD/ChemSketch program of the studied compounds

N°	pIC_{50}	MW	MR (cm^3)	MV (cm^3)	Pc (cm^3)	D (g/cm^3)	α	E_T (eV)	E_{HOMO} (eV)	E_{LUMO} (ev)	ΔE (eV)	μ (Debay)	E_a (ev)	λ_{max} (nm)	f(so)
1*	2.86	249.653	68.780	165.100	486.000	1.511	25.280	- 32 645.8	-5.994	-2.128	3.866	2.627	120.125	281.050	0.0410
2*	2.68	217.199	57.000	158.900	442.300	1.366	22.590	-20 785.0	-6.003	-2.159	3.844	4.096	118.619	284.610	0.0495
3*	2.57	267.207	61.990	188.200	492.300	1.419	24.570	-27 260.4	-6.383	-2.189	4.194	1.378	111.044	269.080	0.0141
4	2.11	267.207	61.990	188.200	492.300	1.419	24.570	-27 260.4	-6.406	-2.275	4.131	6.119	119.755	281.920	0.0452
5	3.17	213.235	61.830	171.000	472.800	1.246	24.510	-19 153.5	-5.825	-2.053	3.772	3.513	116.389	290.070	0.0474
6	2.79	284.098	68.680	177.000	521.900	1.604	27.220	-45 160.6	-6.085	-2.424	3.661	2.005	129.503	260.700	0.0197
7	2.58	251.644	61.890	170.900	478.100	1.472	24.530	-33 299.8	-6.126	-2.466	3.661	2.439	128.637	262.450	0.0157
8	2.74	301.652	66.880	200.200	528.200	1.506	26.510	-39 775.2	-6.510	-2.476	4.034	0.796	134.236	251.510	0.0216
9	2.36	301.652	66.880	200.200	528.200	1.506	26.510	-39 394.5	-6.441	-4.112	2.329	1.982	103.354	326.650	0.0199
10	3.18	247.680	66.730	182.900	508.700	1.353	26.450	-31 668.3	-5.947	-2.362	3.585	2.034	127.140	265.540	0.0174
11	2.72	273.305	75.390	225.600	590.300	1.211	29.880	-25 067.2	-5.909	-1.927	3.982	4.806	124.577	271.000	00212
12	1.74	323.313	80.370	254.900	640.400	1.268	31.860	-31 542.6	-6.286	-1.961	4.325	1.853	126.924	265.990	0.0261
13	2.45	323.313	80.370	254.900	640.400	1.268	31.860	-31 542.6	-6.316	-2.043	4.274	3.829	126.437	267.020	0.0254
14	2.21	269.342	80.220	237.700	620.800	1.133	31.800	-23 435.7	-5.735	-1.827	3.908	3.778	122.404	275.820	0.0153
15	2.79	340.205	87.060	243.700	669.900	1.395	34.510	-49 442.6	-6.008	-2.183	3.824	1.177	130.704	258.300	0.0377
16	2.01	307.751	80.280	237.500	626.200	1.295	31.820	-37 581.8	-6.034	-2.218	3.817	3.297	128.580	262.570	0.0279
17	2.31	537.758	85.270	266.900	676.900	1.340	33.800	-44 057.2	-6.414	-2.236	4.178	0.584	133.849	252.230	0.0640
18	2.24	357.758	85.270	266.900	676.300	1.340	33.800	-44 057.3	-6.444	-2.325	4.119	2.967	133.718	252.480	0.0466
19	1.86	303.787	85.110	249.600	656.700	1.216	33.740	-35 950.4	-5.859	-2.119	3.740	2.116	127.608	264.570	0.0351

The set of sixteen descriptors encoding the 19 of pyrazinecarboxamide compounds, electronic, energetic and topologic parameters are submitted to PCA analysis [18]. The first three principal axes are sufficient to describe the information provided by the data matrix. Indeed, the percentages of variance are 42.52%; 20.36% and 13.64% for the axes F1, F2 and F3, respectively. The total information is estimated to a percentage of 76.52%. The principal component analysis (PCA) [19, 20] was conducted to identify the link between the different variables. Bold values are different from 0 at a significance level of p= 0.05. Correlations between the fifteen descriptors are shown in table 3 as a correlation matrix and in fig. 2 these descriptors are represented in a correlation circle.

The Pearson correlation coefficients are summarized in the following table 3. The obtained matrix provides information on the negative or positive correlation between variables.

* The Molecular Refractivity **MR** is strongly correlated with: the Polarizability **α** (r= 0.996), the Parachor **Pc** (r=0.984), and is well correlated with the Molecular Volume **MV**(r=0.933).
* The Molar Weight **MW** is well correlated with the Parachor **Pc** (r =0.979) and with the Polarizability **α** (r= 0.952).
* The Parachor **Pc** is strongly correlated with the Polarizability **α** (r =0.994).
* The Energy of activation $\mathbf{E_a}$ is negatively correlated with the maximum of absorption $\boldsymbol{\lambda_{max}}$ (r= -0.902).

Table 3. Correlation matrix (Pearson (n)) between different obtained descriptors

	pIC_{50}	MW	MR	MV	Pc	D	α	E_T	E_{HOMO}	E_{LUMO}	ΔE	μ	E_a	λ_{max}
MW	-0.383	**1**												
MR	-0.496	0.679	**1**											
MV	-0.627	0.729	**0.933**	**1**										
Pc	-0.558	0.720	**0.984**	**0.979**	**1**									
D	0.334	-0.060	-0.471	-0.548	-0.496	**1**								
α	-0.512	0.686	**0.994**	**0.952**	**0.994**	-0.496	**1**							
$\mathbf{E_T}$	0.134	-0.638	-0.531	-0.416	-0.512	-0.454	-0.525	**1**						
$\mathbf{E_{HOMO}}$	0.262	-0.465	0.042	-0.178	-0.068	-0.468	0.024	0.401	**1**					
$\mathbf{E_{LUMO}}$	-0.024	-0.043	0.252	0.202	0.218	-0.546	0.239	0.344	0.441	**1**				
ΔE	-0.176	0.217	0.256	0.326	0.281	-0.338	0.251	0.153	-0.082	0.858	**1**			
μ	-0.094	-0.461	-0.237	-0.163	-0.219	-0.349	-0.232	0.623	0.262	0.251	0.128	**1**		
$\mathbf{E_a}$	-0.102	0.470	0.550	0.471	0.536	-0.092	0.555	-0.504	-0.018	0.469	0.530	-0.246	**1**	
$\boldsymbol{\lambda_{max}}$	0.080	-0.416	-0.459	-0.415	-0.459	0.085	-0.467	0.402	0.070	-0.604	-0.711	0.317	**-0.902**	**1**
$\mathbf{f_{(SO)}}$	-0.060	0.424	0.172	0.160	0.159	-0.091	0.150	-0.084	-0.096	0.179	0.254	0.098	0.175	-0.062

CORRELATION CIRCLE

The principal component analysis (PCA) was also performed to detect the connection between the different variables. The principal component analysis revealed the correlation circle (Fig. 2) shows that the F1 axis (42.59% of the variance) appears to represent the Molar Refractivity (**MR**), and the F2 axis (20.36% of the variance) seems to represent the lowest unoccupied molecular orbital energy ($\mathbf{E_{LUMO}}$).

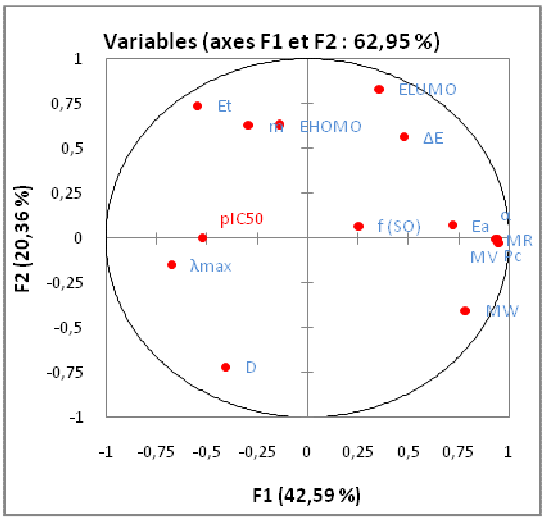


Figure 2. Correlation circles

Analysis of projections according to the planes F1–F2 and F1-F3 (62.95% and 56.23% of the total variance respectively) of the studied molecules (Fig. 3) shows that the molecules are dispersed in two Regions: Region 1 contains compounds having a values of $\mathbf{E_T}$ between -19 153.5 eV and -37 581.8 eV, Region 2 contains compounds having a values of $\mathbf{E_T}$ between -39 394.5 eV and - 49 442.6 eV.

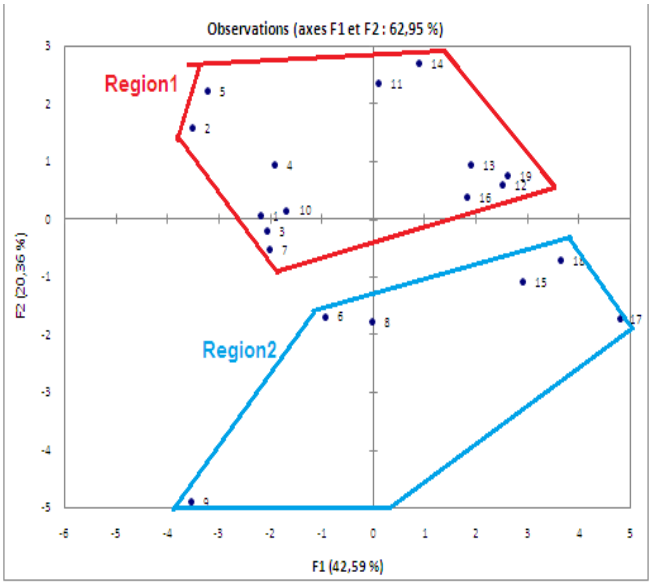


Figure 3. Cartesian diagram according to F1-F2

MULTIPLE LINEAR REGRESSIONS (MLR)

To establish quantitative relationships between herbicidal activity **pIC$_{50}$** and selected descriptors, our array data were subjected to a multiple linear and nonlinear regression. Only variables whose coefficients are significant were retained.

MULTIPLE LINEAR REGRESSION OF THE VARIABLE PIC$_{50}$ (MLR)

Modeling herbicidal activity pI50 value of all training compounds (compounds 16 pyrazinecarboxamides derivatives) led to the best value corresponding to the linear combination of the following descriptors: Molar Volume **MV**, Molecular Weight **MW**, Molar Refractivity **MR**, Parachor **Pc**, Density **D**, the total energy **E$_T$**, the highest occupied molecular orbital energy **E$_{HOMO}$**, the lowest unoccupied molecular orbital energy **E$_{LUMO}$**, the total dipole moment of the molecule **μ**, the activation energy **E$_a$**, absorption maximum **λ$_{max}$** and factor of oscillation **f$_{(SO)}$.**

The resulting equation is:

$$\mathbf{pIC_{50}}=-128.88+2.76\times10^{-03}\times\mathbf{MW}+1.89\times10^{-02}\times\mathbf{MR}-0.59\times\mathbf{MV}+0.27\times\mathbf{Pc}-16.03\times\mathbf{D}+4.82.10^{-04}\times\mathbf{E_t}-16.19\times\mathbf{E_{HOMO}}-3.52\times\mathbf{E_{LUMO}}-6.36.10^{-02}\times\boldsymbol{\mu}+0.22\times\mathbf{E_a}+2.25\times10^{-02}\times\boldsymbol{\lambda_{max}}+8.48.10^{-02}\times\mathbf{f_{(SO)}} \quad (4)$$

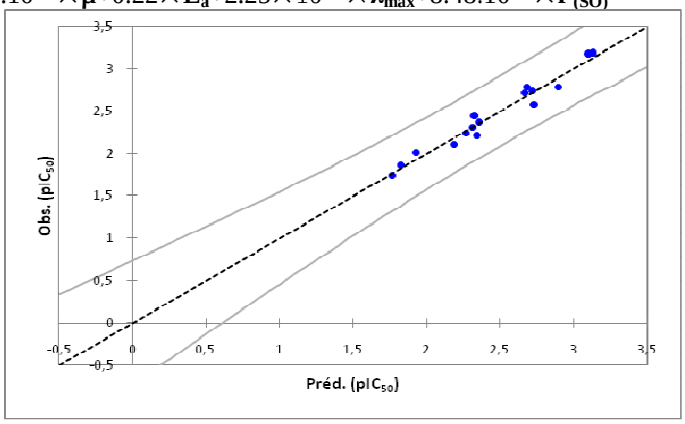


Figure 4. Graphical representation of calculated and observed pIC$_{50}$ by MLR

For our 16 compounds, the correlation between experimental and calculated pIC$_{50}$one based on this model are quite significant (Figure 4) as indicated by statistical values:

N = 16 R = 0.98 R^2 =0.96 RMSE = 0.19

The figure 4 shows a very regular distribution of pIC$_{50}$ values depending on the experimental values.

Available online at www.scholarsresearchlibrary.com

MULTIPLE NONLINEAR REGRESSION OF THE VARIABLE PIC_{50} (MNLR)

We have used also the technique of nonlinear regression model to improve the predicted pIC_{50} in a quantitative way. It takes into account several parameters. This is the most common tool for the study of multidimensional data.

The resulting equation is:

$$\mathbf{pIC_{50}} = -118.584+2.082\times10^{-03}\times\mathbf{MW}+12.975\times\mathbf{MR}-0.594\times\mathbf{MV}+0.254\times\mathbf{Pc}-15.626\times\mathbf{D}-32.323\times\boldsymbol{\alpha}+5.071\times10^{-04}\times\mathbf{E_t}-16.913\times\mathbf{E_{HOMO}}-3.633\times\mathbf{E_{LUMO}}-0.110\times\boldsymbol{\mu}+0.153\times\mathbf{E_a}-5.802\times10^{-03}\times\boldsymbol{\lambda_{max}}+2.126\times\mathbf{f_{(SO)}} \quad (5)$$

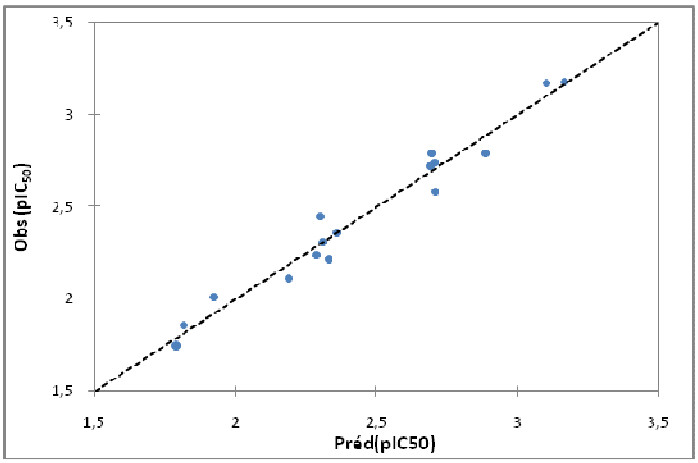


Figure 5. Graphical representation of calculated and observed pIC_{50} by MNLR

The obtained parameters describing the topologic and the electronic aspects of the studied molecules are:

N = 16 R = 0.98 R^2 = 0.96 RMSE= 0.31

The **pIC_{50}** value predicted by this model is somewhat similar to that observed. The fig. 5 shows a very regular distribution of IC_{50} values based on the observed values.

As part of this conclusion, we can say that the herbicidal activity values obtained from MLR and RNLR are highly correlated to that of the observed herbicidal activity.

True predictive power of a QSAR model is to test their ability to predict accurately the activities of compounds from an external test set (compounds which were not used for the model development), the activities of the remained set of 3 compounds (1-3) are deduced from the quantitative model proposed with the 16 molecules (training set) by MLR and MNLR.

The comparison of the values of **pIC_{50}** (test) to **pIC_{50}** (obs.) shows that a good prediction has been obtained for the 3 compounds:

MLR	**N = 3**	**R_{test}= 0.85**	**R^2_{test} = 0.73**
MNLR	**N = 3**	**R_{test}= 0.94**	**R^2_{test} = 0.88**

ARTIFICIAL NEURAL NETWORKS (ANN)

Artificial Neural Networks (ANN) can be used to generate predictive models of quantitative structure-activity relationships (QSAR) between a set of molecular descriptors obtained from the MLR and observed activity.

The correlations coefficients and Standard Error of Estimate, obtained with the Neural Network, show that the selected descriptors by MLR are pertinent and that the model proposed to predict activity is relevant.

The values of predicted activities and the observed values are given in table 4.

The obtained squared correlation coefficient (R^2) value confirms that the ANN results were the best to build the quantitative structure activity relationship models.

In this part, we investigated the best linear QSAR regression equations established in this study. Based on this result, a comparison of the quality of ACP, MLR, MNLR and ANN models shows that the ANN models have substantially

better predictive capability because the MNLR approach gives better results than MLR, and MNLR. ANN was able to establish a satisfactory relationship between the molecular descriptors and the activity of the studied compounds.

The values of predicted activities calculated using ANN and the observed values are illustrated in figure 6.

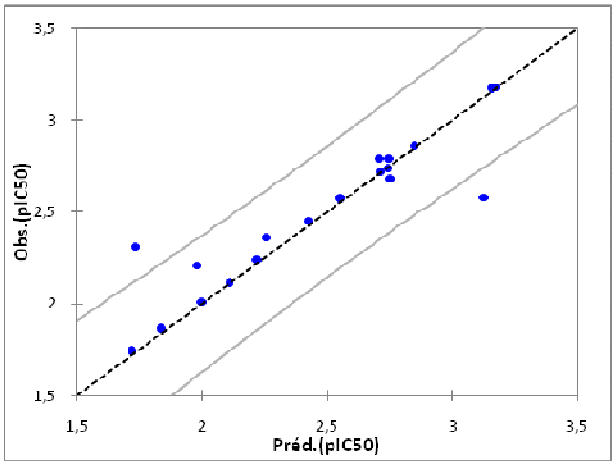


Figure 6. Correlations of observed and predicted activities calculated using ANN

The values of predicted activities calculated using ANN and the observed values are given in table 4.

In the above these QSAR models, ANN is the best model, that indicate the effects of these descriptors on the herbicidal activity of the studied pyrazinecarboxamide derivatives. A unified ANN models with high statistical quality (R = 0.997, R^2 = 0.994) was obtained from the pool of all type of descriptors. In order to validate the generated ANN models leave one out (LOO) method was used to check their predictivity and robustness, test sets of new compounds, not included in the model development set, must be used. The LOO is an approach particularly well adapted to the estimation of that ability. In this procedure, one compound is removed from the data set, the network is trained with the remaining compounds and used to predict the discarded compound. The process is repeated in turn for each compound in the data set. In this paper the 'leave-one-out' procedure was used to evaluate the predictive ability of the ANN.

$$\mathbf{N = 16 \quad R = 0.997 \quad R^2 = 0.994 \quad R_{LOO} = 0.999 \quad R^2_{LOO} = 0.998}$$

Table 4. Observed, predicted activities and residue according to different methods

	pIC_{50}						
N	**Observed**	**MLR**		**MNLR**		**ANN**	
		Pred.	**Resid.**	**Pred.**	**Resid.**	**Pred.**	**Resid.**
4	2.11	2.186	-0.076	2.188	-0.078	2.109	0.001
5	3.17	3.107	0.063	3.101	0.069	3.155	0.015
6	2.79	2.680	0.110	2.695	0.095	2.749	0.041
7	2.58	2.737	-0.157	2.707	-0.127	3.130	-0.550
8	2.74	2.721	0.019	2.709	0.031	2.742	-0.002
9	2.36	2.359	0.001	2.361	-0.001	2.255	0.105
10	3.18	3.134	0.046	3.166	0.014	3.175	0.005
11	2.72	2.665	0.055	2.689	0.031	2.713	0.007
12	1.74	1.764	-0.024	1.793	-0.053	1.718	0.022
13	2.45	2.331	0.119	2.299	0.151	2.423	0.027
14	2.21	2.343	-0.133	2.331	-0.121	1.981	0.229
15	2.79	2.898	-0.108	2.886	-0.096	2.708	0.082
16	2.01	1.930	0.080	1.924	0.086	2.001	0.009
17	2.31	2.310	0.000	2.310	0.000	1.735	0.575
18	2.24	2.269	-0.029	2.286	-0.046	2.223	0.017
19	1.86	1.825	0.035	1.815	0.045	1836	0.001

In this part, we investigated the best linear QSAR regression equations established in this study. Based on this result, a comparison of the quality of the CPA, MLR, MNLR and ANN models shows that the ANN model has

substantially better predictive capability because the ANN approach gives better results than MLR and MNLR. ANN was able to establish a satisfactory relationship between the molecular descriptors and the activity of the studied compounds.

From the values of correlation coefficient of the tree compounds (test set) (R^2_{test}), the 'leave one out' (LOO) cross-validated coefficient (R^2_{LOO}) and other statistical parameters of these methods (MLR, MNLR and ANN), it is clear that the predictive power of this model is high.

The predicted activity values, pyrazinecarboxamides derivatives of this set, obtained from above three QSAR models are listed in Table 4 along with their observed activity.

CONCLUSION

In this study, QSAR models were used to search for molecular descriptors closely related to the herbicidal activity of pyrazinecarboxamides derivatives. Furthermore, internal and external validations were conducted to check the reliability and the stability of the QSAR models elaborated by the MLR, MNLR and ANN methods, and the resulting models were compared. This results show that the artificial neural network ANN had a good predictive ability and strong robustness than the MLR and MNLR, yields a regression model with improved predictive power, we have established a relationship between several descriptors and the herbicidal activity of pyrazinecarboxamides derivatives. Thus, the model could be efficiently employed for estimating the herbicidal activity of pyrazinecarboxamides derivatives and for select the descriptors which have a impact on this activity and are sufficiently rich in chemical, electronic and topological information to encode the structural feature, may be used with other descriptors for the development of predictive QSAR models.

The QSAR model is statistically significant, robust and can be used for prediction purposes, it may be help full for a better understanding of the activity of this class of compounds and useful as a guidance for the proposal of new herbicidal agents.

Acknowledgment
We are grateful to the "Association Marocaine des ChimistesThéoriciens" (AMCT) for its pertinent help concerning the programs.

REFERENCES

[1] P Böger; K Wakabayashi; K Hirai, Springer-Verlag, Berlin, Germany, **2002**, 364.
[2] I Heap, *http://www.weedscience.org*, accessed May 11, **2008**.
[3] K Laarej; M Bouachrine; S Radi; S Kertit; B Hammouti, *E-J. Chem.*, **2010**, 7 (2), 419–424.
[4] S Chtita; M Larif; M Ghamali; A Adad; R Hmamouchi; M Bouachrine; T Lakhlifi, *International Journal of Innovative Research in Science, Engineering and Technolog*, **2013**, (11), pp. 6586-6601.
[5] M Dolzal; PP Cmedlova; J Vinsova; J Kunes; V Buchta; J Jampilek; K Kralova; J.*med. Chem.*, **2008**, 43 (5), 1105-1113.
[6] K Prasad; S Rajesh, *J. Comput. Method. Mol. Design*, **2011**, 1 (1): 7-13
[7] Gaussian 03; MJ Revision B.01; MJ Frisch and al., *Gaussian, Inc.*, Pittsburgh, *PA*, **2003**.
[8] C Lee; W Yang; RG Parr; *Phys. Rev.*, **1988**, 37, 785-789.
[9] S Chtita; M Larif; M Ghamali and al., *IJIRSET*, **2013**, 2(11), 6586-6601.
[10] JN Hogarh; N Seike; Y Kobara; YA Habib; J Jnamd; JS Lee; X Liu; L Jun; G Zhang; S Masunaga, *Chemosphere*, **2012**, 86: 718–726.
[11]M Ghamali; S Chtita; R Hmamouchi; A Addad; M Bouachrine; T Lakhlifi, *IJARCSSE*, **2014**, 4(1), 536–546.
[12] S Huang; R Li; PJ Connolly; S Emanuel; SA Middleton, *Bioorganic & Med. Chem. Lett*, **2006**, 16, 4818–4821.
[13] J Zupan; J Gasteiger, Weinheim, **1999**.
[14] R Chimizou; H Iwamura; T Fujita, *Agric. Food Chem*, **1988**, 36, 1276.
[15] N Turkkan, Revue de l'Université de Moncton, **1993**, 26 (1), 205–221.
[16] G Jing; Z Zhou; J Zhuo, *Chemosphere*, **2012**,86, 76-82.
[17] PY Lee; CYJ Chen, *Hazard. Mater.*, **2009**, 165, 156-161.
[18] PY Lee; CYJ Chen, *Hazard. Mater.*, **2009**, 165, 156-161.
[19] B Elidrissi; A Ousaa; M Ghamali; S Chtita; MA Ajana; M Bouachrine; T Lakhlifi, *JCMMD*, **2014**, 4(3), 28-37.
[20] B Elidrissi; A Ousaa ; M Ghamali; S Chtita; MA Ajana; M Bouachrine; T Lakhlifi, *JCMMD*, **2014**, 4(4), 140-149.