Journal of Materials Physics and Chemistry, 2026, Vol. 14, No. 1, 1-10
Available online at http://pubs.sciepub.com/jmpc/14/1/1
Published by Science and Education Publishing
DOI:10.12691/jmpc-14-1-1

Theoretical Study of the Toxicity of a Series of Amides Herbicides Using Quantitative Structure-activity Relationships

DIOMANDE Sékou[1,*], DJABAN Akoua Déborah[1], KONATE Mory Latif[1], KONE Soleymane[2]

[1]Department of Agro-Industrial Sciences and Technologies (AIST), UFR Agriculture, Halieutic Resources and Agro-Industry (AHRAI), University of San Pedro, San Pedro, Ivory Coast
[2]Department of Sciences of Structure and Matter (SSMT), Laboratory of Constitution and Reaction of Matter (LCRM), University of Félix Houphouët-Boigny, Abidjan, Ivory Coast
*Corresponding author: sekou.diomande@usp.edu.ci

Received March 28, 2026; Revised April 30, 2026; Accepted May 07, 2026

Abstract This QSAR study was carried out on a series of twenty-five (25) amides herbicides and highlights the importance of four (4) key descriptors that contribute to the lethal dose LD_{50}. These are polarizability (Pol), lipophilicity (LogP), total energy (E_{Tot}), and chemical potential (μ). First, the molecular descriptors were determined using the DFT method with the B3LYP/6-31+G(d,p) theory level. Next, the theoretical lipophilicity was calculated using the open-source software A/LogPS 2.1. These descriptors were combined with biological activity using multiple linear regression (MLR) to develop the model. Finally, the domain of applicability (DA) was defined to avoid any hazardous extrapolation, and it appears that all molecular structures can be used through modulation for the prediction of new analogs. These must contain key groups such as halogens (I, Br, Cl), delocalized π systems (benzene ring, conjugated double bonds), and sulfur- or phosphorus-containing groups in their respective structures in order to exhibit optimal activity.

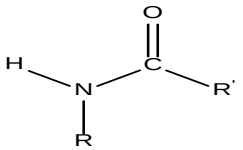
Keywords: *Amide herbicide, Lethal dose LD_{50}, QSAR, DFT*

Cite This Article: DIOMANDE Sékou, DJABAN Akoua Déborah, KONATE Mory Latif, and KONE Soleymane, "Theoretical Study of the Toxicity of a Series of Amides Herbicides Using Quantitative Structure-activity Relationships." *Journal of Materials Physics and Chemistry*, vol. 14, no. 1 (2026): 1-10. doi: 10.12691/jmpc-14-1-1.

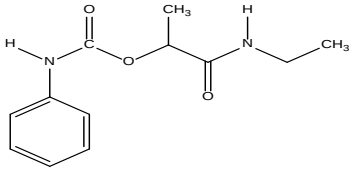
1. Introduction

Today's intensive agriculture use 90% of the pesticides available on the market, which comprise a range of more than 8,000 products sold worldwide [1]. In particular, herbicides play a crucial role in weed control and in optimizing crop yields [2]. In fact, herbicide residues can persist in the soil, seep into groundwater, and accumulate in the food chain, thereby exposing humans to various health risks [3]. Several studies have identified links between chronic exposure to certain herbicides and adverse health effects, such as endocrine disruption, carcinogenic effects, and immune system dysfunction [4].

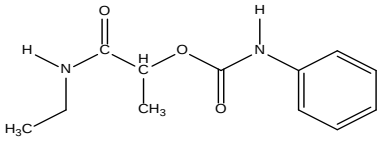
Given that there is a correlation between molecular structures and biological activity or toxicity, a QSAR study is conducted in this research to develop a mathematical model capable of predicting or explaining the toxicity of new compounds. This helps limit the excessive number of experiments—which can be time-consuming and costly—and reduces production costs [5,6]. The agrochemists can then use the model to synthesize new herbicides with improved biological activity. This descriptive and predictive study was conducted on a series of 25 amide herbicide molecules with lethal doses (LD_{50}) [7]. Improving the LD_{50} of the amide series requires the determination of certain descriptors through calculations performed at the B3LYP/6-31+G(d,p) theoretical level using the DFT method. These descriptors will then be combined with the lethal dose to develop a predictive model.



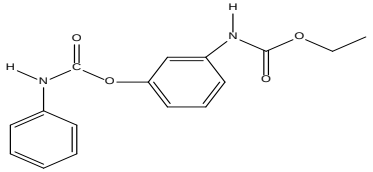
Basic structure



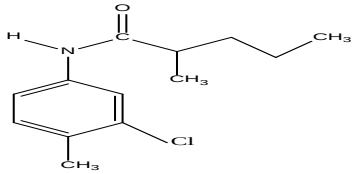
Mol 2



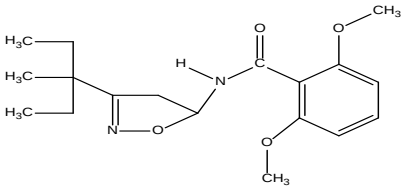
Mol 3



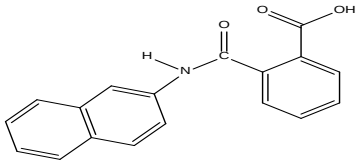
Mol 4



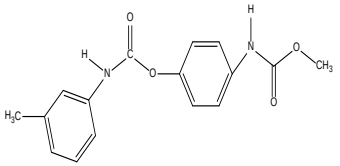
Mol 6



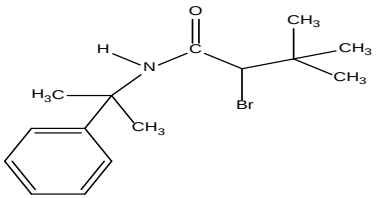
Mol 7



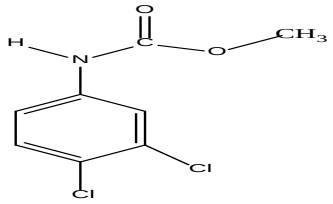
Mol 9



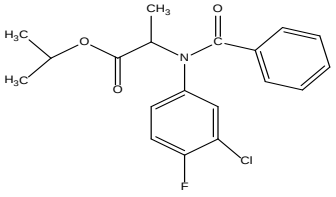
Mol 10



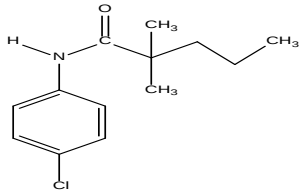
Mol 11



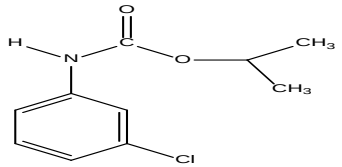
Mol 13



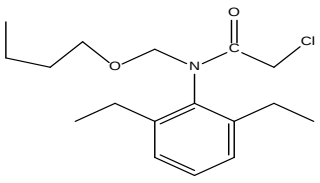
Mol 14



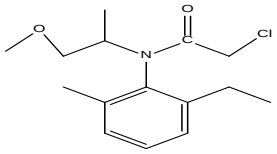
Mol 15



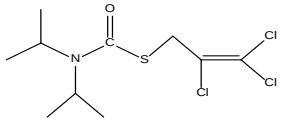
Mol 17



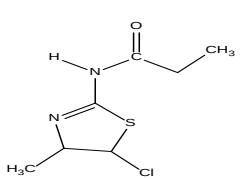
Mol 20



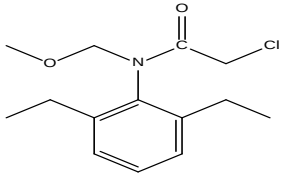
Mol 21



Mol 22

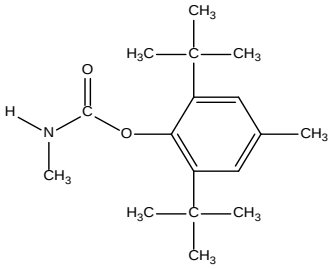


Mol 23

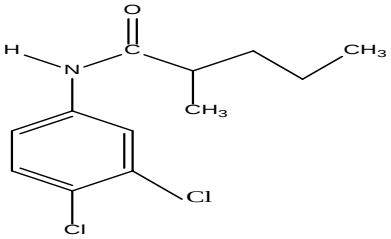


Mol 25

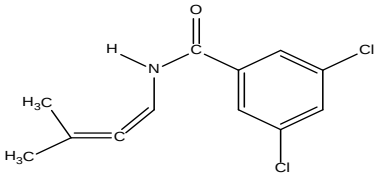
Training set



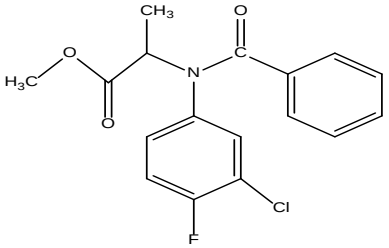
Mol 1



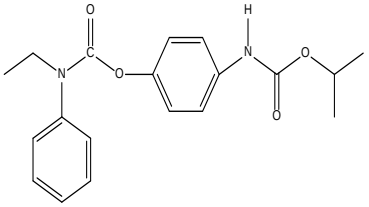
Mol 5



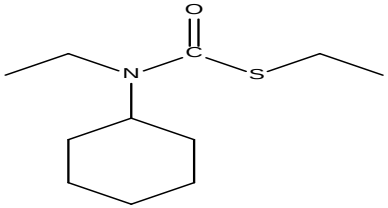
Mol 8



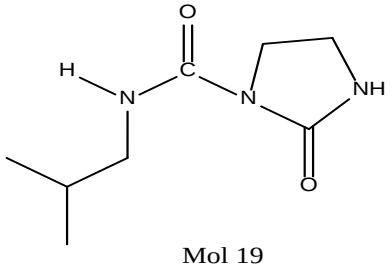
Mol 12



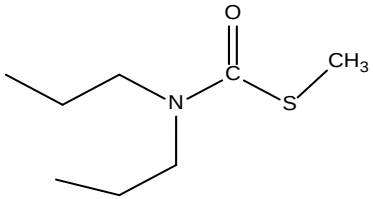
Mol 16



Mol 18



Mol 19



Mol 24

Validation set

Figure 1. Molecular structures of the training and validation sets of amides herbicides used in the model

2. Materials and Calculation methods

2.1. Materials

This study was carried out on a set of 25 molecules, whose structures are shown in Figure 1. Seventeen (17) molecules were used for the training set and eight (8) for validation. The LD_{50} is used to measure a substance's toxicity and is expressed in units of substance mass per body mass, i.e., mg/kg. It values range from 3.80 mg/kg to 416.87 mg/kg. Biological data are generally expressed as the negative of the decimal logarithm of the activity ($-\log_{10}(\mathrm{LD}_{50})$). This allows us to obtain high numerical values when these molecules are highly active [8,9]. The toxicity indicator is expressed by the potential pLD_{50} defined by equation (1):

$$pLD_{50} = \log_{10}\left(1/LD_{50} * 10^{3}\right) \quad (1)$$

Where LD_{50} is the dose of a substance that causes the death of 50% of the test animal population (mg/kg).

2.2. Calculation Methods

The calculations were performed using the Gaussian 09 software in the gas-phase mode [10]. DFT methods are generally known to generate various molecular properties [11,12,13,14,15] in QSAR studies. With the exception of lipophilicity, which was calculated using the A/logPS 2.1 software, all other descriptors are determined from a frequency-following optimization calculation at the B3LYP/6-31+G(d,p) theory level. As for the modeling, it was performed using the multiple linear regression method implemented in Excel [16] and XLSTAT [17].

2.3. Molecular descriptors

Thirteen (13) theoretical descriptors were calculated for the development of the QSAR model. This includes: total energy (E_{Tot}), HOMO energy (E_{HOMO}), LUMO energy (E_{LUMO}), energy gap (ΔE), chemical hardness (η), chemical potential (μ), electrophilicity index (ω), ionization energy (EI), dipole moment (μ_D), electronegativity (χ), lipophilicity (LogP), polarizability (Pol), and electronaffinity (AE). Among these descriptors, we found that a combination of four (4) key ones allowed us to develop a reliable model. These are lipophilicity (logP), polarizability (Pol), total energy (E_{Tot}), and chemical potential (μ). Lipophilicity is an important parameter that can aid in predicting a compound's pharmacological activity, as its transport, its passage through membranes, and its pharmacological activity may be influenced by its partition between a lipid phase and an aqueous phase [18,19]. Chemical potential is the variation in free enthalpy (G) upon the addition of a substance ; it plays a key role in the equilibrium of biological systems and the activity of molecules. It reflects a molecular system's tendency to attract electrons or electron-rich-systems. Polarizability is one of the parameters that reflect molecular properties related to hydrophobicity and, consequently, to biological activities [18]. The partial correlation coefficient calculated between the descriptors studied is less than 0.70 (aij< 0.70) ; this means that these different descriptors are independent of one another [20].

3. Estimation of Predictive Power

3.1. Internal Validation

The quality of a model is determined by certain criteria, such as the coefficient of determination R^2, the standard deviation S, the cross-validation correlation coefficients Q^2_{CV}, and Fischer's F-coefficient. The statistical indicators R^2, S, and F relate to the fit between the calculated and experimental values. They describe the model's predictive capability within its parameters and allow for an estimation of the accuracy of the values calculated on the training set [21,22]. The cross-validation coefficient provides information on the model's predictive power. R^2 indicates the dispersion of the theoretical values around the experimental values. The quality of the model is better when the data points lie close to the fitted line [23]. The fit of the data points to the line can be assessed using the coefficient of determination

Where

$Y_{i,\mathrm{exp}}$: Experimental value of the LD_{50}

$\hat{y}_{i,the'o}$: Theoretical value of the LD_{50}

$\overline{y}_{i,\mathrm{exp}}$: Experimental mean value of the LD_{50}

The closer the value of R^2 is to 1, the better the correlation between the experimental and theoretical values.

Furthermore, the variance σ^2 isdetermined by equation 1:

$$\sigma^2 = S^2 = \frac{\sum\left(y_{i,\mathrm{exp}} - y_{i,th?o}\right)^2}{n-k-1} \quad (3)$$

Where k is the number of independent variables (descriptors) and n is the number of molecules in the test or training set, and n-k-1 is the number of degrees of freedom. The standard deviation S is another statistical measure used. It assesses the reliability and accuracy of a model:

$$S = \sqrt{\frac{\sum\left(y_{i,\mathrm{exp}} - y_{i,the'o}\right)^2}{n-k-1}} \quad (4)$$

The Fischer F-coefficient is also used to measure the statistical significance of the model, that is, the quality of the choice of descriptors that make up the model.

$$F = \frac{\sum\left(y_{i,the'o} - y_{i,\mathrm{exp}}\right)^2}{\sum\left(y_{i,\mathrm{exp}} - y_{i,the'o}\right)^2} * \frac{n-k-1}{k} \quad (5)$$

The cross-validation coefficient of determination (Q^2_{CV}) is used to evaluate the accuracy of predictions on the test set and is calculated using the following equation:

$$Q^2_{CV} = \frac{\sum\left(y_{i,the'o} - \bar{y}_{i,\exp}\right)^2 - \sum\left(y_{i,the'o} - y_{i,\exp}\right)^2}{\sum\left(y_{i,the'o} - \bar{y}_{i,\exp}\right)^2} \quad (6)$$

According to Eriksson and al. [24], a model's performance is characterized by a Q^2_{CV} value greater than 0.5 for a satisfactory model, and for an excellent model, Q^2_{CV} is greater than 0.9. A model's training set is considered to perform well if the acceptance criterion R^2-$Q^2_{CV} < 0.3$ is satisfied.

3.2. External Validation

This validation is characterized by the R^2 (test) and R^2_{cv} (test) statistics. Recently, several studies [25,26] have shown that the R^2, R, and R^2_{cv}statistics are insufficient for assessing the predictive power of RQSA models. Consequently, other statistics must be examined. These parameters are known as "external validation criteria." They are the criteria proposed by Tospha and Roy and Roy.

3.2.1. Trospha's Criteria

1) $R^2_{Test} > 0.7$,

2) $Q^2_{CV\ Test} > 0.6$,

3) $\left|R^2_{Test} - R^2_0\right| \le 0.3$

4) $\frac{\left|R^2_{Test} - R^2_0\right|}{R^2_{Test}} < 0.1 et 0.85 \le k \le 1.15$,

5) $\frac{\left|R^2_{Test} - R'^2_0\right|}{R^2_{Test}} < 0.1 et 0.85 \le k' \le 1.15$

3.2.2. Roy's Criteria

In addition, Roy and Roy refined the prediction method for an RQSA model [27,28]. They developed the metrics r^2_m and Δr^2_m These are metric values ; r^2_m measures the proximity between the observed activity and the prediction. The metric values are calculated based on the observed and predicted activities.The advantage of metric values is that they can be used for both internal and external validation.

According to Roy and Roy, an RQSA model is acceptable if both criteria are met :

1) $\bar{r}^2_m = \frac{r^2_m + r'^2_m}{2} > 0,5$

2) $\Delta r^2_m = \left|r^2_m - r'^2_m\right| < 0,2$

With: $r^2_m = r^2 * (1 - \sqrt{\left(r^2 - r^2_0\right)})$

and $r'^2_m = r^2 * (1 - \sqrt{\left(r^2 - r'^2_0\right)})$

r : Coefficient of determination for molecules in the test or training set.

r_0 : Coefficient of determination for the regression between the experimental values and the predicted values for the test or training set.

r'^2_0 : Coefficient of determination for the regression between predicted values and experimental values for the test or training set.

4. Applicability domain (AD)

The final step in model development is to define the domain within which a compound can be predicted with certainty [29,20]. The domain of applicability (DA) allows us to define the range within which a compound can be predicted with certainty and to avoid any risky extrapolation. The AD can also be defined in terms of activity values and molecular types [29,30]. The Cook's distance method is used. This is a measure of the influence of a suspect point (outlier) on the results of a given regression, as described in [31,32].

$$D_i = \frac{\sum_i \left(\hat{y} - \hat{y}_i\right)^2}{k\sigma^2}$$

Where $\hat{y}$ and $\hat{y}_i$ are the $n \times 1$ vectors of predicted observations for the entire dataset and for the dataset excluding observation i, respectively, and k is the number of parameters estimated by the linear model with variance σ^2.The specific criteria used to exclude a suspected outlier was Di > 4/(n – k – 1), where n is the number of experimental data points. This threshold is determined by the Cook's distance, which is expressed as 4/((n-p-1)). Data points with a Cook's distance greater than this threshold are considered highly influential in the model.

5. Results and Discussion

5.1. Molecular Descriptors and Lethal Dose

Of the thirteen (13) calculated descriptors, only five (5) were significant in the model development using multiple linear regression (MLR). All descriptor values for the seventeen (17) molecules in the test set and the eight (8) molecules in the validation set are presented in Table 1

Table 1. Quantitative descriptors and experimental lethal dose LD_{50} for the test and validation datasets

Molecules	pLD_{50}	EI(eV)	Pol(u.a)	LogP	E_{Tot} (kcal/mol)	μ (eV)	LD_{50}
Training set							
MOL 2	1.6700	0.2381	169.4190	1.2600	-503400.1237	-0.1334	21.38
MOL 3	1.6700	0.2376	171.3200	1.2600	-503402.2429	-0.1327	21.38
MOL 4	2.4200	0.2276	226.5143	2.7500	-646262.0598	-0.1288	3.80
MOL 6	0.5300	0.2429	181.4867	3.7200	-688071.0898	-0.1413	295.120
MOL 7	1.4800	0.2284	251.2937	3.9200	-697053.3603	-0.1369	33.11

Molecules	$p\mathbf{LD_{50}}$	EI(eV)	Pol(u.a)	LogP	$\boldsymbol{E_{Tot}}$ (kcal/mol)	μ (eV)	$\mathbf{LD_{50}}$
Training set							
MOL 9	1.4500	0.2146	245.4847	3.3900	-611386.2153	-0.1436	35.48
MOL 10	1.4300	0.2265	230.7663	3.3900	-646261.7322	-0.1264	37.15
MOL 11	1.2000	0.2265	205.0253	4.2700	-2062465.185	-0.1450	63.10
MOL 13	0.3800	0.2402	138.9887	2.8400	-900321.2432	-0.1417	416.87
MOL 14	1.0500	0.2525	245.5093	3.8800	-988998.4739	-0.1540	89.13
MOL 15	1.2200	0.2556	177.1110	3.9900	-688061.2031	-0.1049	60.26
MOL 17	1.2500	0.2373	148.2903	2.6300	-661267.5998	-0.1341	56.23
MOL 20	1.0200	0.2529	226.3930	4.2000	-833948.8604	-0.1434	95.50
MOL 21	0.9900	0.2569	201.4440	3.3700	-784597.6866	-0.1477	102.33
MOL 22	0.7400	0.2499	194.2197	4.4100	-1442994.313	-0.1454	181.97
MOL 23	1.0000	0.2599	135.7813	1.8000	-826162.7286	-0.1511	100
MOL 25	0.7500	0.2566	191.4920	3.0200	-759928.4701	-0.1452	177.83
Validation set							
MOL 1	2.1000	0.2312	216.4007	5.3300	-545551.0339	-0.1207	7.94
MOL 5	1.5800	0.2489	182.8743	4.4000	-951801.2629	-0.1506	26.30
MOL 8	1.5900	0.2374	195.4933	3.7900	-950252.2126	-0.1522	25.70
MOL 12	1.1700	0.2540	222.7393	3.2100	-939647.339	-0.0989	67.61
MOL 16	1.0700	0.2178	261.1820	3.8800	-720275.9055	-0.0947	85.11
MOL 18	1.2200	0.2367	161.0007	3.9700	-602462.6483	-0.1213	60.26.
MOL 19	1.1300	0.2513	123.0807	-0.2400	-394526.4844	-0.1302	74.13
MOL 24	0.9400	0.2379	155.2333	3.7300	-578554.5209	-0.1197	114.82

5.2. QSAR model

The contribution of a descriptor to $\mathbf{LD_{50}}$, when correlated with other descriptors in the regression equation, depends not only on the sign of its coefficient but also on the sign of the descriptor it self. When the descriptor and its coefficient have the same sign, the descriptor enhances biological activity. Conversely, if they have opposite signs, the descriptor weakens the activity. Equation (2) below represents the best model based on the data in Table 1.

$$pDL_{50} = 3,7513 + 1,2089E-02 * \mathbf{Pol} - 0,5043 * \mathbf{LogP}$$
$$-5,3631E-07 * \mathbf{Electr} + 23,9594 * \boldsymbol{\mu} - 1,8915 * \mathbf{EI}$$

N=17; R^2= 0.824; Q^2_{CV} = 0.653; S= 0.921; F= 10.21; $R^2 - Q^2_{CV}$ = 0.171

According to the model, a high polarizability (Pol) value helps improve the $\mathbf{LD_{50}}$ .Additionally, lipophilicity, total energy, and chemical potential must be low to enhance this activity.

5.3. Model Validation

The values of the partial correlation coefficients for the descriptors (aij) in the model are shown in Table 2.

Table 2. Pearson correlation matrix between the various descriptors

Variables	$p\mathbf{LD_{50}}$	EI (eV)	Pol(u.a)	LogP	$\boldsymbol{E_{Tot}}$ (kcal/mol)	μ (eV)
$p\mathbf{LD_{50}}$	1	-0.437	0.295	-0.009	0.252	0.153
EI (eV)	-0.437	1	-0.404	-0.094	-0.411	-0.303
Pol(u.a)	0.295	-0.404	1	0.537	-0.161	0.138
LogP	-0.009	-0.094	0.537	1	-0.411	0.036
$\boldsymbol{E_{Tot}}$ (kcal/mol)	0.252	-0.411	-0.161	-0.411	1	0.321
μ (eV)	0.153	-0.303	0.138	0.036	0.321	1

Values in bold differ from 0 significantly at p < 0.05. Highly significant at p < 0.01. Highly significant at p < 0.001.

The partial correlation coefficients between the descriptors studied are less than 0.70, indicating that the descriptors are independent of one another.

The regression line between the experimental $\mathbf{LD_{50}}$ and the theoretical $\mathbf{LD_{50}}$ for the training set (blue points) and the test or validation set (red points) is shown in Figure 2.

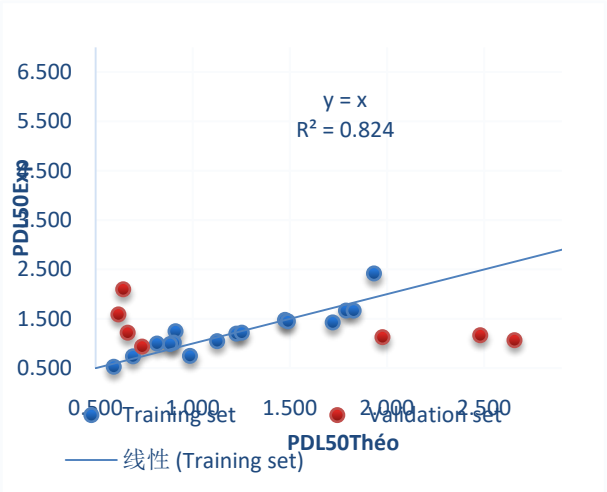


Figure 2. Graph showing the actual and predicted lethal dose for the entire set of training and validation data used in the model

The robustness and predicted reliability of the developed QSAR model were evaluated using the criteria proposed by Trospha and Roy.

Trospha's Criteria

1) $R^2_{Test} = 0.836 > 0.7$,

2) $Q^2_{CV\ Test} = 0.811 > 0.6$,

3) $\left|R^2_{Test} - R^2_0\right| = 0.015 \leq 0.3$

4) $\dfrac{\left|R^2_{Test} - R^2_0\right|}{R^2_{Test}} = 0.018 < 0.1 \; and \; 0.85 \le k = 1 \le 1.15$,

5) $\dfrac{\left|R^2_{Test} - R'^2_0\right|}{R^2_{Test}} = 0.04 < 0.1 \; and \; 0.85 \le k' = 0.976 \le 1.15$

All criteria have been verified and confirm that the model is suitable for predicting $\mathbf{LD}_{50}$. Let's check Roy's criteria.

Roy and Roy's Criteria

1) $\overline{r}^2_m = \dfrac{r^2_m + r'^2_m}{2} = 0.601 > 0.5$

2) $\Delta r^2_m = \left|r^2_m - r'^2_m\right| = 0.16 < 0,2$

With: $r^2_m = 0.679$ et $r'^2_m = 0.523$

Criterias are also verified.

External validation also confirms the validity of the QSAR model.

The model highlights five (5) descriptors, four (4) of which are significant: lipophilicity, total energy, and chemical potential. Figure 3 below shows the coefficients (sign and importance) assigned to these descriptors.

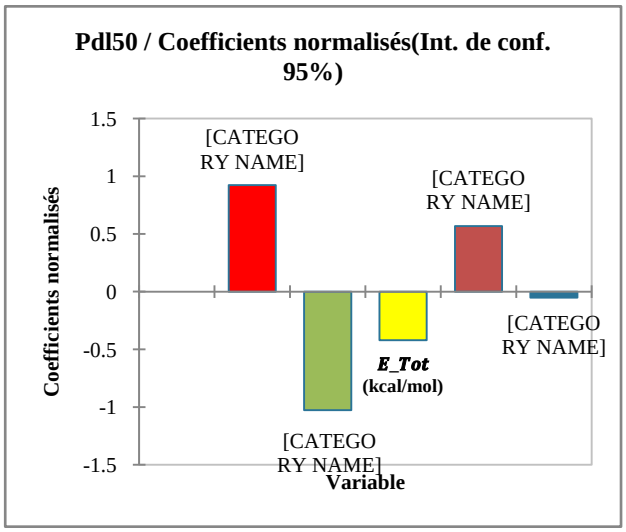


Figure 3. Coefficients defining the contributions of the three major descriptors to $\mathbf{LD}_{50}$ in the model

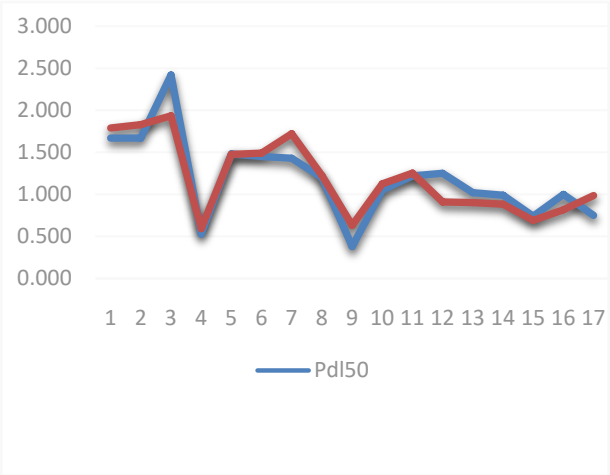


Figure 4. Similarity curve between experimental values and predicted values

An order of decreasing importance of the descriptors gives us: lipophilicity (LogP) > polarizability (Pol) > chemical potential (μ) > total energy.

For a better understanding of the model's predictive power, plots showing the agreement between the experimental and theoretical values have been generated. Figure 4 shows these plots.

These curves show minimal deviation. The model is therefore acceptable for prediction purposes.

To clarify the model's limits of application, its domain of applicability has been determined. The Cook distance provides further details.

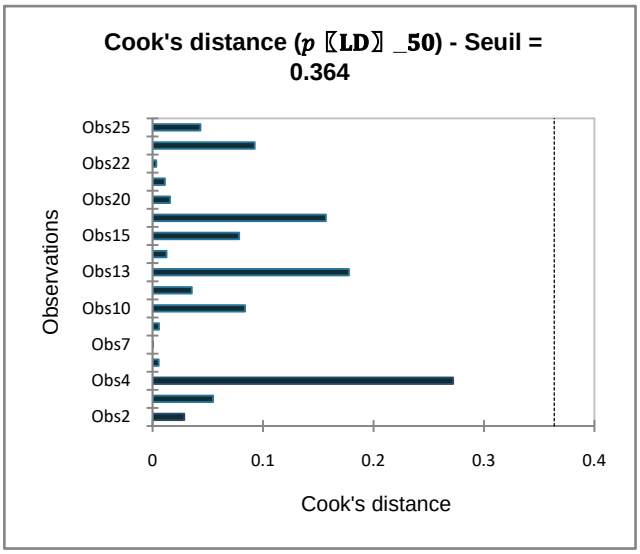


Figure 5. Cook's Distance

With p denoting the number of descriptors in the model and n denoting the number of molecules in the training set.

$$D = \frac{4}{(n-p-1)} = \frac{4}{(17-5-1)} = 0.364$$

This demonstrates the absence of outliers. The results of the various validations and the analysis of the applicability domain show that these molecules have a significant structural influence on the model. They can be used to predict the $\mathbf{LD}_{50}$ of amides herbicides and their derivatives using the established model.

6. Conclusion

The QSAR study on the series of amide herbicides allowed us to identify four (4) key descriptors for predicting the of $\mathbf{LD}_{50}$ these molecules and the results of the internal and external validation criteria suggest that the model is predictive. Among these descriptors are polarizability (Pol), lipophilicity (LogP), total energy (E_{Tot}), and chemical potential (μ). In addition, it is clear that lipophilicity plays a key role in this model, followed by polarizability. Since the polarizabilities of the molecules must be high, modulation with electron-rich groups such as halogens (I, Br, Cl) and delocalized π systems (benzene ring, conjugated double bonds), as well as groups containing sulfur or phosphorus, could improve activity. The domain of applicability indicates the absence

of outliers. All of these molecules can therefore be used in the prediction. Thus, this study opens the way for the synthesis of new amide herbicides with improved activity.

References

[1] S.M. Kumar, D.S. Kumar, S. Kumargupta, S.P. Pandey, R.yadav, Persistent organochlorine pesticides and polychlorinated biphenyls in intensive agricultural soils from North India. *Asian Journal of Pharmaceutical Research*, 2011, 1, 62.

[2] Powles, S. B., & Yu, Q. Evolution in action: Plants resistant to herbicides. *Annual Review of Plant Biology*, 61, 317–347, 2010.

[3] Carvalho, F. P. Pesticides, environment, and food safety.*Food and Energy Security*, 6(2), 48–60, 2017.

[4] Mostafalou, S., & Abdollahi, M. Pesticides: an update of human exposure and toxicity. *Archives of Toxicology*, 91(2), 549–599 , 2017.

[5] Chtita S., Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments*. PhD Thesis*, 2017.

[6] Patrick GL, Chimie pharmaceutique. Paris: *De Boeck*, 2003.

[7] DIDI Mabrouka, Prédiction de la toxicité d'une série d'amides herbicides, *Mémoire de Magister*, 2010.

[8] S.Chaltterjee, A.Hadi, B. Price, Regression Analysis by Examples; *Wiley VCH*: New York, USA, 2000.

[9] Phuong HTN, thèse de doctorat «Synthèse et étude des relations structure/activité quantitatives (QSAR/2D) d'analogues Benzo [c] phénanthridiniques», *Université d'Angers*, (France), 2007.

[10] Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

[11] P.K. Chattaraj, A.Cedillo, and R.G Parr, *J. Phys.Chem*., 103:7645, 1991.

[12] P.W.Ayers, and R.G.Parr, *J.Am Chem., Soc*., 122:2010, 2000.

[13] F. De Proft, J.M.L.Martin, and P. Geerlings, *Chem. Phys. Let*., 250:393, 1996.

[14] P.Geerlings, F. De Proft, J.M.L.Martin, In Theoretical and Computational Chemistry; *Seminario*, J., Ed.,; Elsevier; Amsterdam, , 4 (Recent Developments in Density Functional Theory): 773, 1996.

[15] F.DeProft, J.M.L. Martin, and P. Geerlings, *Chem. Phys.Let*., 256: 400, 1996.

[16] Microsoft ® Excel ® 2016 (16.0.11029) MSO (16.0.11029) 64 Bits (2016) Partie de *Microsoft Office Professionnel Plus.*

[17] XLSTAT Version 2021.2.2 (64 bit) Copyright 1995-2025 (2021) XLSTAT and AddinsoftwareRegistrered Trademarks of Addinsoft. https: //www.xlstat.com.

[18] Chtita S. Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments [*thèse*]. [Meknès]: Moulay Ismail; 2017.

[19] Numbury Surendra Babu, Didugu Jayaprakash. Global and Reactivity Descriptors Studies of Cyanuric Acid Tautomers in Different Solvents by using of Density Functional Theory (DFT). *Int J Sci Res IJSR*; 4(6): 615-20, 2015.

[20] A. Vessereau, Méthodes statistiques en biologie et en agronomie. Lavoisier (Tec and Doc). Paris: 538, 1988.

[21] G.W. Snedecor, W.G. Cochran, Statistical Methods; *Oxford and IBH: New Delhi,India*; 1967: 381.

[22] M.V.Diudea, QSAR/QSAR Studies for Molecular Descriptors; *Nova Science*: Huntingdon, New York, USA, 2000.

[23] E.X.Esposito, A.J.Hopfinger, J.D.Madura, Methods in Molecular Biology, 275: 131-213, 2004.

[24] L.Eriksson, J. Jaworska, A. Worth, M.T.D. Cronin, R.M.Mc Dowell, P.Gramatica, Methods for Reliability and Uncertainly Assessment and for Applicability Evaluations of Classification and Regression-Based QSARs, *Environmental Health Perspectives*, 111(10): 1361-1375, 2003.

[25] A. Golbraikh, A. Tropsha, Beware of q 2! *J. Mol. Graph. Model*., 20, 269–276, 2002.

[26] T. M. Martin, P. Harten, D. M. Young, E.N. Muratov, A. Golbraikh, H. Zhu and A.Tropsha, Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling? *J. Chem. Inf. Model*., 52, 2570−2578, 2012.

[27] Roy, P. P. and Roy, K**.** On some Aspects of Variable Selection for Partial Least Squares Regression Models. *QSAR & combinatorial Science*, 2008, 27, 302-313.

[28] Partha Pratim Roy, Somnath Paul, Indrani Mitra and Kunal Roy, On Two Novel Parameters for Validation of Predictive QSAR Models. *Molecules*, 14, 1660-1701, 2009.

[29] Ghamali M, Chtita S, Bouachrine M, Lakhlifi T. Méthodologie générale d'une étude RQSA/RQSP. *Rev Interdiscip*.; 1(1): 1-7, 2016.

[30] Vinca P. Approches structure-propriété pour la prédiction des propriétés physico-chimiques des substances chimiques [Chimie théorique et informaique]. [Paris VI]: Pierre et Marie Curie,. 208p. 2013.

[31] Díaz-García, J.A.; González-Farías, G. A note on the cook's distance. *J. Stat. Plan. Inference*, 120, 119-136, 2004.

[32] Militino, A.F.; Palacios, M.B.; Ugarte, M.D. Outliers detection in multivariate spatial linear models. *J. Stat. Plan. Inference*, 136, 125-146, 2006.