



Evaluation of the Model Prediction Toxicity (LD50) for a serie of 42 organophosphate Pesticides

H FIKRI*, A. OUBIBI¹, T. FECHTALI

Laboratory of Neurosciences, Integrated Physiopathology and Natural Substances - F.S.T. Mohammedia

* Corresponding author: fikrane@hotmail.com

ABSTRACT: Relationships structure-toxicity have been studied for a set of 42 organophosphate pesticides (OPs) through multiple linear regression (RLM) and artificial neural networks (RNA). A model with three descriptors, including: total lipophilicity [$\log(P)$], widths radicals R_1 [(LR₁)] and R_2 [(LR₂)] and has achieved good results in phase Training and phase prediction of toxicity [$\log LD_{50}$ (lethal dose 50, Oral rat)]. The linear model (RLM: $n=40$, $r^2=0.86$, $s=40$ and $q^2 = 0.66$) and non-linear model with a configuration [3-6-1] (RNA: $r^2=0.95$, $s=0.73$ and $q^2 = 0.17$) have proved very successful and complementary. The selected descriptors indicate the importance of lipophilicity and widths radicals R_1 and R_2 in the contribution of the toxicity of pesticides derived from OPs used in this study. This information is relevant to the design of a new model of non-toxic pesticides OPs.

[H FIKRI, A. OUBIBI, T. FECHTALI. **Evaluation of the Model Prediction Toxicity (LD50) for a serie of 42 organophosphate Pesticides.** *Biomedicine and Nursing* 2021;7(4):61-66] ISSN 2379-8211 (print); ISSN 2379-8203 (online). <http://www.nbmedicine.org>. 8. doi:[10.7537/marsbnj070421.08](https://doi.org/10.7537/marsbnj070421.08).

KEYWORDS: Multiple Linear Regression (RLM), artificial neural networks (RNA), Organophosphates (OPs), LD₅₀, Descriptors.

INTRODUCTION

The pesticide is a limit generic for a variety of classes such as chemical herbicides, fungicides and nematicides. The techniques of computer simulation still more effective means potentially offering alternatives to probe relations structure-toxicity. The objectives of our study are to provide additional information on the behaviour of organophosphorus compounds and set in the future the necessary criteria for designing a model for a new generation of organophosphate pesticides.

The use of quantitative relationships structures QSAR activities currently has considerable attention [1] for pharmaceutical needs [2], as well as the study of the toxicological mechanisms of chemical environmental pollutants (products Endocrine disrupting phytochemicals) [3]. Today, a large number of families of compounds have already been the subject of such research. Among these families is organophosphate pesticides.

Pesticides especially organophosphate (OPS) are the most used in the world of agricultural production. Of these, 70% are highly toxic, although they have a low persistence and are easily biodegradable, they are subject to ecological concerns because they are toxic to non-target species even at low concentrations. [4].

In recent years, chemists and biologists have paid great attention to quantitative structure-activity relationships, known as QSAR. The development of such relationships successfully predicts certain properties and activities of chemical structures without recourse to synthesize or test them [5].

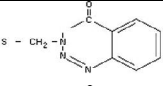
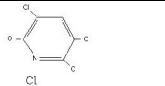
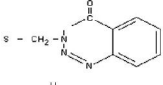
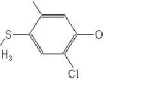

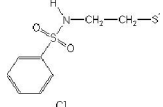
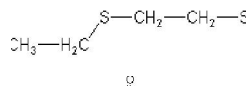
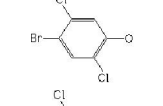
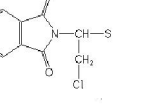
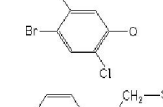
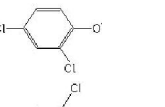
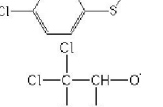
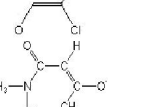
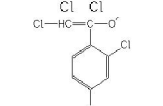
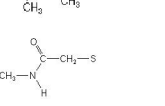
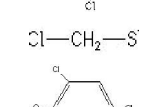
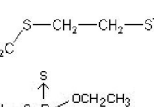
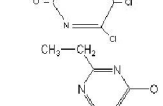
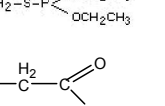
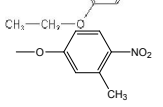
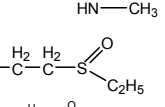
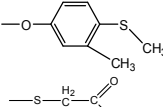
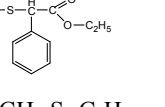
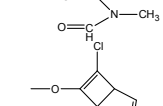
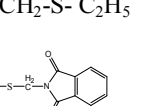
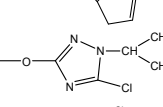
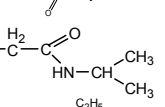
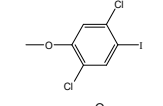
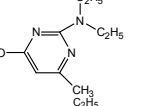
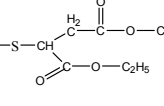
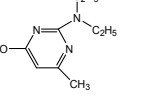






Indeed, one of the major concerns of scientists is to have a pesticide that is more effective and less toxic. Only the achievement of these objectives will guarantee the development of the pesticide market. It is within the framework of this perspective that our approach, which allows not only to distinguish between the variation of the toxicity between the different pesticide molecules, but also to establish an effective model on a data bank of 42 pesticides Organophosphorus.

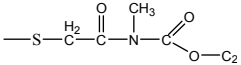
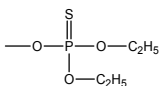
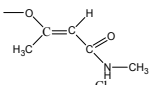
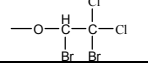
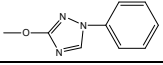
1. Materials and methods

1.1. Experimental data

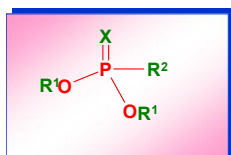
In order to establish our model of quantitative structure-toxicity relationship, we collected a sample, as wide possible, of 42 organophosphate compounds described in FOOTPRINT's PPDB database [6], with activity values Specific ecotoxicologicals of acute oral toxicity for rats (mammalian test organisms).

Table 1: Chemical structures of the compounds studied and experimental values of logDL50 (lethal dose 50, oral, rat).

N	X	R ₁	R ₂	log LD ₅₀ (Exp)	N°	X	R ₁	R ₂	log LD ₅₀ (Exp)
1	S	C ₂ H ₅		1,08	11	S	CH ₃		3,45
2		CH ₃		0,95	12	S	C ₂ H ₅		1
3	S			2,43	13	O	CH ₃		1,48
4	S	CH ₃		3,20	14	S	C ₂ H ₅		1,85
5	S	C ₂ H ₅		1,72	15	S	C ₂ H ₅		2,23
6	S	C ₂ H ₅		1	16	O	CH ₃		1,90
7	S	C ₂ H ₅		1,64	17	O	CH ₃		1,23
8	O	C ₂ H ₅		1,08	18	O	CH ₃		2,40
9	S	C ₂ H ₅		0,84	19	S	C ₂ H ₅		1,59
10	S	C ₂ H ₅		1,82	20	S	C ₂ H ₅		2,32
21	O	CH ₃		3,25	32	O	CH ₃		1,81
22	S	CH ₃		2,52	33	O	CH ₃		1,68
23	S	CH ₃		2,30	34	S	CH ₃		2,40
24	S	CH ₃		2,56	35	S	C ₂ H ₅		0,66
25	O	CH ₃		1,98	36	S	CH ₃		2,05
26	S	C ₂ H ₅		1,43	37	S	C ₂ H ₅		0,90
27	S	CH ₃		3,37	38	S	C ₂ H ₅		2,15
28	S	CH ₃		3,07	39	S	CH ₃		3,15

29	S	C ₂ H ₅		1,56	40	S	C ₂ H ₅		1
30	O	CH ₃		1,23	41	S	CH ₃	S-CH ₂ -CH ₂ -S- C ₂ H ₅	1,60
31	O	CH ₃		1,96	42	S	C ₂ H ₅		1,82

The chemical structures of the series of the organophosphate compounds studied are grouped in table 1 with the logDL50 values used as a dependent variable. All these compounds have the motif [P (o) (o)] as a common chemical structure (Figure 1).



R¹ = Methyl or Ethyl.

R² = alkyl, heterocyclic or aromatic groupment.

X = O ou S.

Figure 1: General structure of the organophosphorus pesticides.

1.2. Analysis by Multiple Linear Regression (MLR)

A total of 63 physicochemical descriptors were generated to code each molecule by the MMP program [7].

The experimental data, containing all the values by n compounds (n = 42), was subjected to Multiple Linear Regression analysis [8]. We used two main approaches in the MLR: the stepwise regression approach [9]; and the "backward" approach, [10]. The contributions of the relevant descriptors are calculated using the Gore method [11] [12], which excludes a single example of all the data for build the entire test."leave-one-out". The predictive power of the model has been evaluated by the cross-validation method.

1.3. RNA analysis

In neural networks, the units are organized in indexed layers. Each unit receives as input the outputs of the previous layer. The evaluation of the impact of the relevant descriptors [13] is to remove each time a descriptor i of the initial architecture and train the network, then calculate its contribution.

2. RESULTS AND DISCUSSIONS

2.1. Establishment of RLM models

Given the large number of 57 descriptors used to code each molecule, we subjected our data matrix to Stepwise stepwise selection [14, 15, 16], in order to highlight the most relevant descriptors. The two steric parameters width of radical R1 [L(R1)], radical R2 [L(R2)], and lipophilicity [Log P], represent the relevant descriptors in the determination of the toxicity of these organophosphorus pesticides studied. The best model obtained by the multiple linear regression is the following (equation1):

Equation 1: $\text{Log}_{10}(\text{LD}_{50}) =$

$$(3,858 \pm 0,790) + (0,081 \pm 0,023) * \text{LogP} - (0,938 \pm 0,133) * [L(R_1)] + (0,257 \pm 0,062) * [L(R_2)]$$

$$n = 4 ; r = 0,84 (r^2 = 0,64) ; s = 0,45 ; F = 29,52$$

The contributions of the parameters [L(R1)], [L(R2)] and Log P are respectively 45.78% 33.33 and 20.89%. The model is statistically significant and accounts for up to 63% of initial information. The residual standard deviation is relatively small (s = 0.45), it is of the order of the error generally committed in biological tests of this type. The large contribution of the width of the substituent R1 reflects the great influence of the size of this substituent in particular on the toxicity of organophosphorus pesticides.

To get closer and closer to the experimental error that is generally of the order of 5% and to check the prediction limits in our sample, we applied (which is of the order of 0.4781) to the sample of 42 molecules. We obtained a new sample of 40 molecules. The application of multiple linear regression to this sample leads to equation (2) after elimination of molecules 39 and 40:

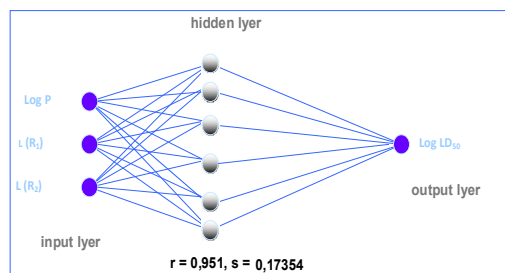


Figure 1 : Description of architectural Neural network

The comparison of this classification with that obtained in the linear model reveals a change of order between Log P, L (R1) and L (R2). This change can be explained by the possible existence of a non-linear relationship between activity and lipophilicity, which is not the case for widths.

Equation 2: $\text{Log}(\text{LD}_{50}) = (1,941 \pm 0,064) + (0,360 \pm 0,082) * \text{LogP} - (0,581 \pm 0,077) * [\text{L}(\text{R}_1)] + 0,273 (\pm 0,070) * [\text{L}(\text{R}_2)]$

$n = 40 ; r = 0,86 (r^2 = 0,728); s = 0,40; F = 32,18$

The statistical quality of the equation is very good, it explains 73% of the total variance, and higher than the other models described in the literature if we take into account the number of descriptors used. It explains up to 73% of the total variance with a standard error "s" much lower than the average error made on the observed values of Log (LD50) which is of the order of 0.741 for an interval ranging from 0, 66 to 3.45. to show that our model is not due to chance, we applied the experiment of changing the column of the dependent variable randomly so that each molecule does not find its true activity but the activity of a another molecule, without touching the columns of the independent variables. The result of this test on the sample of 42 molecules shows that the statistical quality of our model decreases very remarkably, it goes from $r = 0.86, s = 0.40$ and $F = 32.18$ at $r = 0.46, s = 0.60$ and $F = 0.07$. This result clearly indicates that the descriptors selected for this study describe well the activity of the series of organophosphorus compounds.

2.2. ESTABLISHMENT OF NON-LINEAR MODELS:

Neural network. In order to improve the linear model obtained and test the possibility of non-linear effects that may possibly exist between the activity and the descriptors of the linear model, we used the artificial neural network. The neural network employed therefore has the architecture [3-6-1], (three neurons in the input layer, six neurons in the hidden layer and a neuron in the output layer).

This preliminary study (Table 2) allowed us to conclude that the neural network with the architecture [3-6-1] schematized in figure 2 is able to establish a satisfactory relationship between the relevant descriptors and the Log activity (LD50). We observe an improvement in the statistical parameters of the non-linear model ($r = 0.95$ and $s = 0.17$) compared to the linear model ($r = 0.86$ and $s = 0.40$), which suggests the existence possible non-linear relations between the activity and the descriptors. The evaluation of the contribution of the relevant descriptors gives the following classification: $\text{LogP} > \text{L}(\text{R}_1) > \text{L}(\text{R}_2)$.

Table 2: variation of r^2 and s with number of neurons

NNC	s	r
2	0,27564299	0,75056049
3	0,25787559	0,82837559
4	0,1948047	0,92550181
5	0,17503741	0,94251181
6	0,17354904	0,95147377
7	0,17417838	0,94901349

Conclusion

Because some organophosphorus compounds are known for their pesticidal properties, [18]. we have established a study of the quantitative relationships between the molecular structure and toxicity of a series of organophosphorus pesticides, with the aim of predicting the activity of new products at high levels added value.

QSAR models are expected to play an important role in the risk assessment of chemicals on humans [19]. The most relevant descriptors were evidenced by stepwise linear multiple regression analysis. The analysis of the statistical parameters of models published in the literature shows that our model is more efficient [20,21, 22]. Its generate the steric aspect, expressed both by the widths of the radicals R 1 and R 2 and the however, the nature of the heteroatom (X) does not seem to have a decisive effect on the activity of organophosphorus pesticides. The structure-toxicity relationship model thus produced during this work can be classified as a predictive model that assists in the selective design of new molecules with low toxicity, in the family of organophosphorus compounds. Indeed, the analysis of the schematized model, based on a molecular description of the organophosphorus compounds, shows that a low toxicity (high LD50) is associated with a low value in [L(R1)], a global lipophilicity and a width. [L(R2)] more important. In conclusion, this prediction model makes it possible to avoid experimental tests and provide an immediate result.

For future developments, lazy structure–activity relationships [23] will be established and tested algorithms, semantic web aware web services, and language bindings, which can serve as building blocks for new algorithms and applications. We hope that these facilities will speed up the development cycle of future predictive toxicology applications, and will ultimately lead to improved and more relevant applications in this area.

REFERENCES

- [1] S. Weaver, M.P. Gleeson / Journal of Molecular Graphics and Modelling 26 1315–1326 (2008).
- [2] United States Patent 6503938, 2-phenylindoles as antiestrogenic pharmaceutical agents, PatentStorm, (2003).
- [3] R. McKinly / Environment International 34 168–183, (2008).
- [4] S.Penas., Toxicol. Appl. Pharm. 196, 108–113, (2004).
- [5] United States Patent US042189, one-dimensional QSAR Models, (2005).
- [6] L. Liu, Y. Cheng, Y. Fu, R. Chen and Q. Guo, J. Chem. Inf. Comput. Sci. 42, 1164, (2002).
- [6] FOOTPRINT pesticide properties database PPDB; Base de données rassemblée par l'Université de Hertfordshire, dans le cadre de projet FOOTPRINT financé par l'UE (FP6-SSP-022704) (2007, 2008).
- [7] MMP, molecubr modeling pro-Demo™ Revision 301 demo published by ChemSW Software TM.
- [8] D. J. Livingstone, D. T. Manallick, J. Med. Chem., 36, 1295 (1993).
- [9] [STATGRAPHICS Plus for Windows Version 42 / 15, (1999).
- [10] MATLAB Version.7, the mat works, Inc, Natick, MA, (2002).
- [11] W. L. Gore, Inter-science, New York, 14, (1952).
- [12] M. Ferandez, journal of molecubr graphics and modelling 25 410-422 (2006).
- [13] Cherquaoui, D., Esseffar, M., Villemin, D., Cence, J. M., Chastrette, M., Zakarya, New J. Chem., 839 (1998).
- [14] Zahouily, M., Rhihil, A., Bazoui, H., Sebti, S., Zakarya, D. J. Mol. Model. 8 168 (2002).
- [15] Zahouily, M., Lzar, A., Elmakssoudi, A., Rakik, J., Elyachi, S., Rayadh, A. J. Mol. Model, (2005).
- [16] M. Zahouily, J. Rakik, M. Lzar, M. A. Bahloui, A. Rayadh, and N. Komiha

- ARKIVOC 245-256, (2007).
- [17] M.Hamadache et al. Journal of Hazardous Materials Volume 303, Pages 28-40 February (2016),
- [18] . J. Hardy, B., Douglas, N., Helma, C., Rautenberg, M., Jeliaskova, N., Jeliaskov, V., et al. Chem. Inf. Model. 49, 2077–2081. (2010)
- [19] Maunz, A., Helma, C., Cramer, T., and Kramer, S] SAR QSAR Environ. Res. 19, 413–431. (2010)
- [20] JB Veselinović, GM Nikolić, NV Trutić- SAR and QSAR - Taylor & Francis, (2015)
- [21] Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Alken, P., etal. Toxicol. Sci. 95,5–12 GNU Scientific Library (2009).
- [22] VM Alves, EN Muratov, A Zakharov, NN Muratov- Chemical Toxicology, -Elsevier (2017)
- [23] Maunz A, Gütlein M, Rautenberg M, Vorgrimmler D, Gebele D and Helma. Front. Pharmacol (2013).

3/12/2021