http://dx.doi.org/10.21577/0103-5053.20200039

J. Braz. Chem. Soc., Vol. 31, No. 7, 1531-1540, 2020
Printed in Brazil - ©2020 Sociedade Brasileira de Química

Article

Pyrimidine Derivatives: QSAR Studies of Larvicidal Activity against *Aedes aegypti*

Zenaide S. Monte, [a] ***Daniela M. A. F. Navarro,*** [b] ***Júlio C. R. O. F. Aguiar,*** [b] ***Jessica S. Nascimento,*** [b] ***Marcus T. Scotti,*** [c] ***Luciana Scotti,*** [c] ***Renata P. C. Barros,*** [c] ***Aline C. S. Santos,*** [d] ***Valéria R. A. Pereira,*** [d] ***Emerson P. S. Falcão*** [e] ***and Sebastião J. Melo*** *[,a]

[a]*Programa de Pós-Graduação em Ciências Farmacêuticas e em Ciências Biológicas, Universidade Federal de Pernambuco, Avenida Professor Moraes Rego 1235, 50740-560 Recife- PE, Brazil*

[b]*Departamento de Química Fundamental, Universidade Federal de Pernambuco, 50740-560 Recife-PE, Brazil*

[c]*Programa de Pós-Graduação em Produtos Naturais e Sintéticos Bioativos, Universidade Federal da Paraíba, 58051-900 João Pessoa-PB, Brazil*

[d]*Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, 50670-420 Recife-PE, Brazil*

[e]*Núcleo de Nutrição do Centro Acadêmico de Vitória (CAV), Universidade Federal de Pernambuco, 55608-680 Vitoria de Santo Antão-PE, Brazil*

The present study investigated the activity of pyrimidine derivatives against *Aedes aegypti*. Two compounds, **3c** and **3d** showed excellent larvicide activity. Additionally, quantitative structure-activity relationship (QSAR) models were built using multiple-linear regression and partial least squares with descriptors generated from Dragon and VolSurf+ software, respectively. The best model is obtained with multiple linear regression (MLR), leading to a robust model. Moreover, the QSAR model is validated by means of some internal validation techniques in order to check its reliability, quality and robustness for predicting the larvicidal activity against *A. aegypti*. The models confirmed that the three-dimensional structure of molecules, steric properties, hydrophobic polar surface area, log partition (logP) and a simple pattern of substituent groups as methyl, methoxy, and succinimide in the pyrimidine derivatives are responsible for the larvicidal activity of the pyrimidine derivatives. Even more, the activity decreases by an electron-withdrawing group in R_1 and increases when it is replaced by an aromatic ring activator group. These findings will aid in further studies of new pyrimidine derivatives active against *Aedes aegypti*.

Keywords: pyrimidine derivative, *Aedes aegypti*, QSAR model, robust model, MLR model

Introduction

Mosquitoes are a means of transmission of many neglected diseases, with millions of people being threatened vector-borne in the world,[1-3] particularly in tropical and subtropical regions.[4-6] According to Gorle *et al.*,[7] these diseases affected chiefly the tropical and subtropical regions of countries that resist chemical vector control programs because the population refused to prevent mosquito control by using chemical treatment with synthetic insecticides. According to the World Health Organization (WHO), in recent years, the transmission of dengue hemorrhagic fever, zika virus, dengue fever and chikungunya has increased dramatically in those regions that are conducive to mosquito proliferation, as these mosquitoes adapt in environments characterized by irrigation systems and heavy rains.[7-12] Chikungunya produces fever and rheumatic pain, interfering with people's quality of life for days, months or even years in more serious cases. Zika presents headache, irritation of the skin, redness of the eyes or vomiting, fatigue, fever, chills, loss of appetite or sweating. In some cases, zika may cause paralysis (known as Guillain-Barré syndrome) and, in pregnant women, there may be hemorrhagic dengue: damp, pale and cold skin, as well as a decrease in blood pressure, high fever and malaise up to three days after the mosquito bite.[13-16] Unfortunately, the use of synthetic insecticides

*e-mail: melosebastiao@yahoo.com.br

to combat the disordered growth of *Aedes aegypti* has not been effective in combating these mosquitoes since they are becoming resistant to conventional poisons, as well as increasing environmental problems and presenting serious damage to human health.[17-19] Twenty-seven pyrimidine derivatives have already been reported in the literature by our research group.[20,21] Because pyrimidines may have insecticidal activity,[22] we investigated its biological activity against *A. aegypti*, with respect to the dengue vector. Additionally, quantitative structure-activity relationship (QSAR) studies were performed to understand the main physicochemical features responsible for the larvicidal activity of the studied compounds.

Experimental

Maintenance and creation of the *Aedes* colony

The beginning of the cycle was the hatching of *Aedes*' eggs in glass cups containing enough distilled water to cover them completely and a little cat food to stimulate the hatching more quickly. The development of the larval cycle continued by changing water in the basins and adding more food. The end of one cycle and the beginning of the next cycle was accomplished by putting the larvae and pulp in a glass beaker with a little distilled water and food (cat food) into cages. The following week, the cups were removed and a blood repast was added to the liquid. New egg collection was conducted with filter paper and dried at room temperature of 27 ± 1 °C and 75% humidity. The mosquitoes were kept alive with a piece of cotton soaked with 10% sucrose solution.

Experimental procedure for larvicidal bioassays

Preliminary solubility tests were initially conducted for each sample aliquot to be tested in co-solvents such as ethanol, tween 80 or acetone, in order to select the co-solvent which best solubilized the sample in water. After this test, a stock solution concentration equal to 100 parts *per* million (ppm) was prepared by dissolving 5.0 mg of the compound to be tested in a volume of 0.7 mL of the chosen cosolvent.[23,24] The contents were then transferred from the beaker into a 50 mL volumetric flask and the volume topped with distilled water. Preliminary larvicidal tests were performed in concentrations of 10, 50 and 100 ppm in order to observe the concentration range where the compound was most active. The tests were carried out in triplicate which, in each replica, 20 larvae of *A. aegypti* were used and larvicidal activity was observed after 24 and 48 h from the beginning of the test. The larvae were considered dead when they did not respond to the stimulus or did not emerge on the surface of the solution. Negative controls (solution containing only co-solvent and distilled water) were carried out simultaneously with the tests. For the determination of the lethal concentration (LC_{50}) values for 50% of the larval population, the data was obtained using Probit software[25] with the statistical program StatPlus Pro 6.2.5.0,[26] at a 95% confidence level.

Evaluation of cytotoxic activity

Macrophage cell lines RAW 264.7 were cultured in a complete Dulbecco's modified Eagle medium (DMEM, 100 mg mL^{-1} streptomycin, 100 U mL^{-1} penicillin and 10% fetal calf serum (FCS)), (Cultilab, Campinas, São Paulo, Brazil). Cells were maintained at 37 ºC, in a 5% CO_2 atmosphere. Macrophages were seeded at 10^5 cells *per* well in 96-well plates and incubated for 24 h (37 °C and 5% CO_2). Compounds were then added to six different concentrations (6.25 to 200 μg mL^{-1}). Wells containing only the culture medium and cells (without treatment) were used as the negative control. Cells were incubated for 48 h. Then a MTT solution (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromide (Sigma, St. Louis, MO, USA) at 5 mg mL^{-1} in phosphate buffered saline (PBS) was added. The plates were incubated again for 2 h. The remaining culture medium and the unreduced MTT were removed and 100 μL of dimethyl sulfoxide (DMSO) was added for solubilization of formazan. The amount of formazan was determined by measuring the absorbance at 570 nm. The assay was performed in triplicate. The cytotoxicity concentration (CC_{50}) was calculated by regression analysis with GraphPad Prism Software 5.0 (San Diego, CA, USA).[27]

Statistical analysis

Statistical analysis was performed using non-parametric tests. A simple linear regression test was performed to obtain CC_{50}.

Molecular descriptors

Initially, the SMILES (simplified molecular line entry system) code of the molecules was obtained, generating a single file, .smi format, of all molecules. This file was introduced into the software Standardizer, ChemAxon,[28] to canonize structures, add hydrogens, perform aromatic form conversions, clean the molecular graph in three dimensions. The process uses a divide-and-conquer approach. The structure is split into small fragments which

are organized into a tree using connectivity information. Conformers generated for the initial structure (represented by the root node in the tree) are optimized. The tree building process uses a proprietary extended version of the Dreiding force field.[29] Finally, the compounds were saved in SDF format that was used as input in two software programs to generate the molecular descriptors. The first program was the VolSurf+ v. 1.0.7,[30,31] which produced the file containing the structures in 3D, by calculating molecular descriptors through the intrusion of these molecules to the molecular interaction fields (MIFs), using 4 different probes: amide nitrogen (N1), carbonyl oxygen (O), hydrophobic probe (DRY) and water probe (H_2O). It was also calculated some molecular descriptors not derived from the field of molecular interaction (non-MIF), generating a total of 128 descriptors; for example, quantifiable descriptors such as hydrophobic-lipophilic balance, capacity factors, molecular size, hydrophobic and hydrophilic regions, amphiphilic moments and moments of energy interaction.[31] Due to ease of use, understanding and interpretation of these descriptors, they were chosen in the present study.

The second program used was Dragon v. 7.0.10,[32] to calculate a total of 5255 molecular descriptors through 29 descriptor blocks. After calculating these descriptors, a data treatment was performed in which descriptors of constant variables were excluded for each descriptor block, besides the exclusion of those descriptors that had a high degree of correlation ($r < 0.95$).

Model generation

The QSAR models were developed using the partial least squares (PLS) method with VolSurf+ software.[30,31] The PLS is based on linear regression, making it possible to extract and rationalize multivariate information to explain the maximum correlation between the descriptors, matrix **x** and matrix **y**, to then calculate a new set of orthogonal variables that had not yet been correlated, the latent variables (LVs). Whenever the number of variables is greater than the number of samples and multicollinearity occurs among independent variables, this method becomes appropriate.[33,34] The self-scale pre-processing was performed and applied to all independent variables, by subtracting the mean by the values of each variable and then dividing the resulting values by the standard deviation. A variable influence in projection (VIP) plot was used to select the number of original variables. The VIP parameter has the ability to condense the importance of the variables in the PLS model and to quantify this importance by the coefficient values and indicate the contribution of each descriptor to the model.[35] The model performance was estimated by the variance explained in the coefficient of determination (r^2) and the coefficient of determination in the cross-validation by leave-one-out (Q_{cv}^2). The model and number of ideal LVs, that is, orthogonal linear combinations of the original variables, were determined by the highest value of Q_{cv}^2.[36,37] MobyDigs software[38] was also used to calculate and generate regression models using a genetic algorithm, in which the models for the activity were constructed and internally validated using leave-one-out Q_{cv}^2.

Results and Discussion

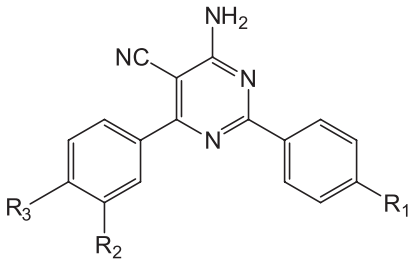
Insecticidal activities

Insecticidal resistance is the main problem in the control of the mosquito population in the world, principally for *A. aegypti*. In view of this, compounds of pyrimidine derivatives with potential anti-larvicidal activity were investigated to detect mortality of larvae in the presence of these compounds. The 27 compounds **3a**-**3r**, **5a**-**5f** and **6g**-**6i** screened for larvicidal activity against four instar larvae of *A. aegypti* are listed in Table 1. Considerable variation was observed in the susceptibility of the larvae to the different pyrimidines, indicating those which proved to be the most promising agents, with impressive larvicidal activity. Thus, 17 compounds **3e**-**3l**, **3n**, **3q**, **3r**, **5a**, **5e**, **5f**, **6g**-**6i** of the 27 pyrimidines showed ≥ 50% mortality against the mosquito larvae in concentration of 100 ppm; the other 10 compounds **3a**-**3d**, **3m**, **3o**, **3p** and **5b**-**5d** achieved ≤ 50% mortality level of larval mortality in concentration of 100 ppm.

Among all the synthesized compounds, **3l** and **5a** showed highest mortality results (56.7 ± 15.4, 65.7 ± 15.9, respectively) and reached standard value at 100 ppm concentration at 72 h of exposure; other results were higher than the standard value at 100 ppm concentration. Impressive mortality was observed for **3c** and **3d**, with values 5.4 ± 1.8 and 4.4 ± 0.5, respectively. The mortality rates were considered at 100 ppm concentration at 72 h.

In parallel to the evaluation of larvicidal activity, the toxicity of the compounds on mammalian cells was determined, using the murine macrophage RAW 264.7. A variation in the activity of the pyrimidines was observed. Nine of the 27 pyrimidines in Table 1 showed lower cytotoxicity, with high CC_{50} values, above 100 μM: **3a**, **3b**, **3f**, **3j**, **3k**, **3m**, **3o** and **5a**, **5d**. Another nine, showed higher toxicity, with CC_{50} values below 50 μM: **3d**, **3e**, **3g**, **3l**, **3n**, **3p**, **3q**, **5c** and **6i**. Intersecting the results obtained in the evaluation of larvicidal activity and mammalian cell toxicity, the pyrimidines which showed the best results, i.e.,

Table 1. Results of biological activity for compounds **3a-3r**, **5a-5f** and **6g-6i**



Compound	R_1	R_2	R_3	Concentration range / ppm	LC_{50} / μM	Confidence interval / %	Standard error for LC_{50}	Cytotoxicity CC_{50} / μM
3a	H	NO_2	H	50-20	46.3	42.8-52.5	2.4	112.1
3b	H	H	NO_2	50-15	18.6	14.3-21.6	1.95	124.55
3c	OMe	NO_2	H	25-4	5.4	4.3-6.4	0.5	88.56
3d	OMe	H	NO_2	45-5	4.4	1.4-7.0	1.8	16.37
3e	C_2H_6N	NO_2	H	–	> 100	–	–	33.29
3f	C_2H_6N	H	NO_2	–	> 100	–	–	203.56
3g	C_2H_6N	H	Cl	–	> 100	–	–	31.07
3h	OMe	H	Cl	–	> 100	–	–	57.06
3i	NH_2	H	OMe	–	> 100	–	–	82.69
3j	NH_2	NO_2	H	–	> 100	–	–	188.06
3k	NH_2	H	NO_2	–	> 100	–	–	264.78
3l	NO_2	NO_2	H	100-50	56.7	24.0-68.7	15.4	49.81
3m	NO_2	H	NO_2	50-20	31.4	28.3-34.9	1.7	131.47
3n	NO_2	H	OMe	–	> 100	–	–	46.89
3o	Me	NO_2	H	40-5	17.0	14.1-19.4	1.4	175.85
3p	Me	H	NO_2	10-45	18.5	16.8-20.4	0.9	30.02
3q	Me	H	Cl	–	> 100	–	–	17.39
3r	Cl	H	H	–	> 100	–	–	87.8
5a	H	succinimide	H	120-70	65.7	41.0-105.3	15.9	136.1
5b	H	H	succinimide	70-20	48.5	36.8-109.3	13.64	52.36
5c	OMe	succinimide	H	35-10	11	7.0-13.6	1.9	17.32
5d	OMe	H	succinimide	40-10	12.99	9.5-15.5	1.64	128.85
5e	C_2H_6N	succinimide	H	–	> 100	–	–	99.3
5f	C_2H_6N	H	succinimide	–	> 100	–	–	84.4
6g	pyrrol-2-one	H	C_2H_6N	–	> 100	–	–	60.94
6h	pyrrol-2-one	H	OMe	–	> 100	–	–	56.03
6i	OMe	H	maleimide	–	> 100	–	–	37.2

LC_{50}: lethal concentration; CC_{50}: cytotoxicity concentration.

low toxicity in macrophages and high mortality rate at the concentration of 100 ppm for *A. aegypti* larvae were: **3f**, **3j**, **3k** and **5a**.

QSAR models

In order to minimize errors in the studies by QSAR, Kubinyi[39] recommends that the biological activity values should be standardized (inhibition concentration-IC_{50}, LC_{50}, lethal dose-LD_{100}, etc) and that the variation in the values of activity between the most active and least active compounds is at least one log unit. It can be observed that the difference of logarithmic units between the highest activity compound, **3d** $pLC_{50} = 4.9$, and the lowest activity compound, **5a** $pLC_{50} = 3.75$, is 1.15, in accordance with the basic principle of the QSAR study.

Regression analysis of the training set generated equation 1, with TDB05v (3D Autocorrelations) and Mor04u (3D MoRSE) descriptors, which are able to explain almost 94.4% of the variance in antilarval activity.[40]

$$pLC_{50} = -0.317\ (\pm 0.070)\ TDB05v - 0.218\ (\pm 0.558)\ Mor04u + 4.24\ (\pm 0.067) \quad (1)$$
$$(n = 12;\ r^2 = 0.944;\ s = 0.102;\ F = 75.94;\ Q^2_{cv} = 0.920;\ SDEP = 0.106)$$

where n is the number of samples, r^2 is the coefficient of determination, s is the mean square error, F is the Fisher function, Q^2 is the cross-validated r^2, and SDEP is standard deviation error of prediction.

An analysis of Table 2 and Figure 1 indicates that each value represents the adjustment in relation to a line of the points that had been used for the calibration of the model. Equation 1, described above, shows that the value of the internal prediction (leave-one-out) coefficient (Q^2_{cv}) is considerable (0.920), indicating how robust the model was. The value of F (75.94) was highly expressive, with 95% significance and with 2 and 9 degrees of freedom, where the minimum required value is 4.26. In the conception of obtaining information from 3D (three-dimensional) atomic coordinates through the transformation technique used in electron diffraction studies, one of them Mor04u, is based on the 3D-morse descriptor (3D-molecule representation of structures based on electron diffraction) selected in equation 1. The descriptor TDB05v is a 3D-topological distance-based descriptor, which is based on Moreau-Broto's 2D (two-dimensional) autocorrelation that portrays how a certain property is spread throughout a topological molecular structure, also encoding information about the separation space between two atoms.[40] The

Table 2. Predicted by cross-validation (leave-one-out) and experimental values of pLC_{50} from equation 1, and the respective errors

Compound	pLC_{50}[a] / M	pLC_{50} Dragon / M	Error Dragon	pLC_{50} Vols / M	Error Vols	pLC_{50} consensus / M	Error consensus
3a	3.84	3.87	0.03	4.01	0.17	3.94	0.10
3b	4.23	4.01	–0.22	3.99	–0.24	4.00	–0.23
3c	4.81	4.85	0.04	4.54	–0.27	4.70	–0.11
3d	4.90	4.97	0.07	5.02	0.12	5.00	0.09
3l	3.81	3.87	0.06	3.79	–0.02	3.83	0.02
3m	4.06	3.98	–0.08	4.19	0.13	4.09	0.03
3o	4.29	4.23	–0.06	4.14	–0.15	4.19	–0.11
3p	4.25	4.38	0.13	4.59	0.34	4.49	0.23
5a	3.75	3.92	0.17	3.96	0.21	3.94	0.19
5b	3.88	3.95	0.07	3.67	–0.21	3.81	–0.07
5c	4.56	4.45	–0.11	4.50	–0.06	4.48	–0.09
5d	4.49	4.48	–0.01	4.58	0.09	4.53	0.04

[a]–logarithm of the lethal concentration values for 50% of the larval population in values of experimental activity (pLC_{50}).

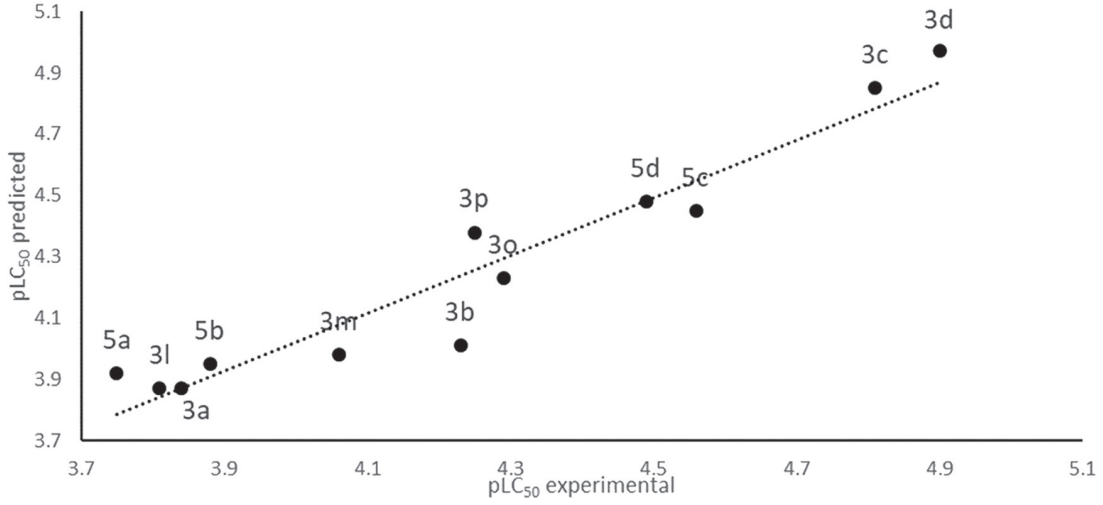


Figure 1. Graph of the MLR model generated from the Dragon descriptors of equation 1, representing the values of experimental activity (pLC_{50}) *versus* predicted (leave-one-out cross validation) activity values.

larvicidal activity is closely related to the three-dimensional structure of molecules as well as the steric properties of specific points of the molecule.

The PLS model selected with five LVs demonstrated the following statistical parameters: $Q_{cv}^2 = 0.741$ (Figure 2 and Table 2), standard deviation of errors of prediction (SDEP) = 0.190, $R^2 = 0.934$ and standard deviation of errors of calculation (SDEC) = 0.096.

The graphs of the t1-t2 scores and loadings with two latent variables of the PLS model can be observed in Figure 3, showing a separation between the most active compounds and less active. The compounds to the left of the graph indicate the lower pLC_{50} chemical structures, while the compounds to the right of the graph depict the chemical space of the molecules with the highest pLC_{50} values.

In the loading plot, (Figure 3b), the descriptors VolSurf+ that contributed more in the PLS model can be seen. The HSA (hydrophobic surface area) and log partition between cyclohexane and water (logPc-hex) contribute positively to the pLC_{50} values. These descriptors are related to the hydrophobic characteristics of these molecules. The HSA is computed through the hydrophobic region of the molecule and the logPc-hex.[41,42] The physical-chemical characteristics described by the HSA descriptors and logPc-hex are important for the discovery of new drugs since they provide pharmacokinetic properties in the initial phase, revealing molecules with greater chance of being transported through the circulatory system to the target tissues.[41,43] Since the LC_{50} values were obtained from the test against the larvae mosquito, these properties are related to potential toxicity. On the other hand, the descriptors PSA (polar surface area), PSAR (ratio between the PSA and the surface) and PHSA (polar hydrophobic surface area), the ratio between the polar surface area PSA and the HSA, contribute negatively to pLC_{50} values, therefore the more polar compounds are less active. Figure 4 shows the lipophilic and hydrophilic regions of the higher activity compound, **3d**, and the compound with the lower activity, **3l**. It can be observed that compound **3d** has a larger lipophilic area and smaller hydrophilic surface area when compared with molecule **3l**, thus justifying its better activity.

Given the results and analyzing the most active and least active compounds, we can observe the influence of the substituents on the radicals R_1, R_2 and R_3. The PLS model classified that molecules with higher lipophilicity (logPn-hex) and HSA have higher activities and those with smaller PSA, smaller radius between PSAR and the smallest radius between the polar surface area and the PHSAR has low biological activity. Thus, we can analyze the substituents R_1, R_2 and R_3 to understand the relationship between chemical structure and biological activity.

First looking at the radicals R_2 and R_3 when substituted by the nitro group (NO_2), we see that when located at R_3, position *para*, there is a slight improvement in activity, comparing, for example, molecules **3d** and **3c**, where the only difference is precisely the position of the group nitro. This slight improvement is explained by the decrease in PSAR and PHSAR when nitro is in the R_2 target position, Table 3. Similarly, it occurs when the substituent of R_2 and R_3 is succinimide, molecules **5c** and **5d**, where substitution in R_2, *meta* position, is favored.

Now looking at the radical R_1, which has been replaced by OCH_3, CH_3, H and NO_2, we see that when there is a methoxyl (OCH_3) the activity profile of the molecule is better than when replaced by a methyl, molecules **3c** and **3o**, for example. When substituted by methyl group there is a considerable decrease in PSA, PSAR and PHSAR (Table 3), causing the activity to be decreased, pLC_{50} from 4.81 for

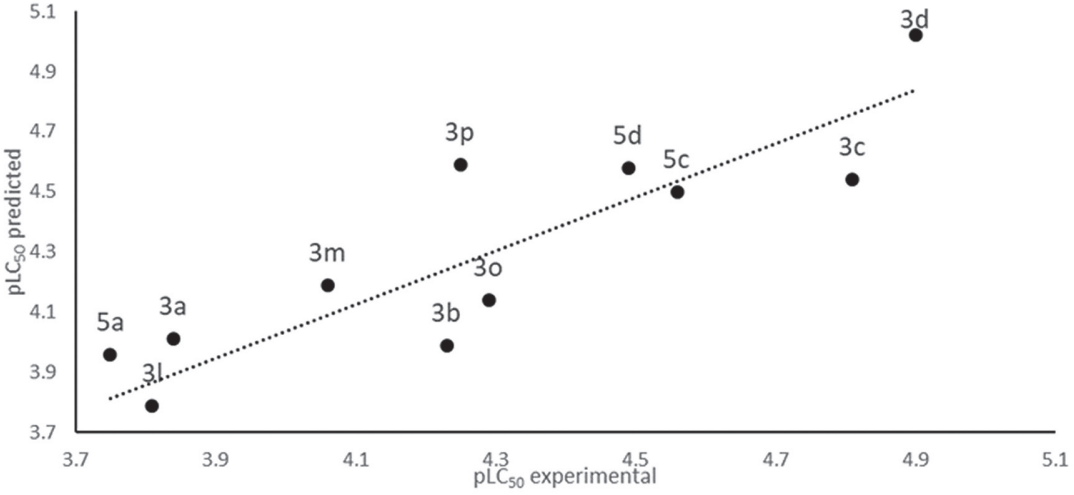


Figure 2. Plot of experimental activity values (pLC_{50}) *versus* the predicted values by cross-validation leave-one-out of the selected partial least squares (PLS) model.

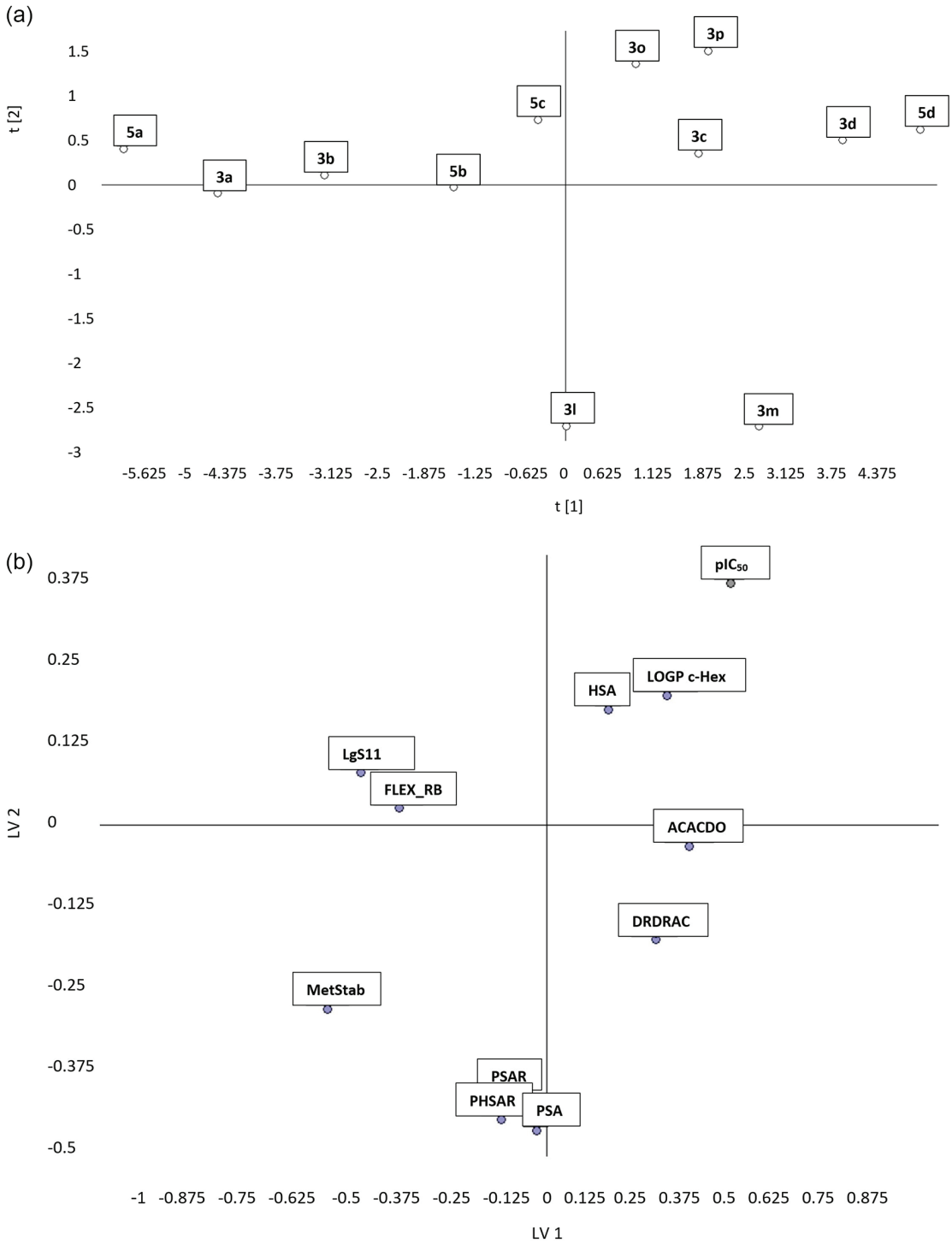


Figure 3. Graphs of the partial least squares (PLS) model. The (a) scores and (b) loadings plot with the two latent variables (LV1 *vs.* LV2).

molecule **3c** and from 4.29 for molecule **3d**. Comparing now the methoxyl with the nitro group, molecules **3c** and **3l**, we have that the methoxyl group has a higher lipophilicity due to the presence of methyl and there is also an increase in hydrophobic surface area, explaining why the molecule **3c** is more active than molecule **3l**, with **3l** being one of the least active (pLC_{50} = 3.81). Similar behavior occurs when R_1 is replaced by hydrogen compared to methoxyl substitution, molecules **5c** and **5a**, the increase in logPc-hex and HSA (Table 3) will be greater with methoxyl, so the molecule will have better activity.

An important fact that should also be considered in the analysis is that the vector sum of the individual bond dipole moments will change the polar and nonpolar profile

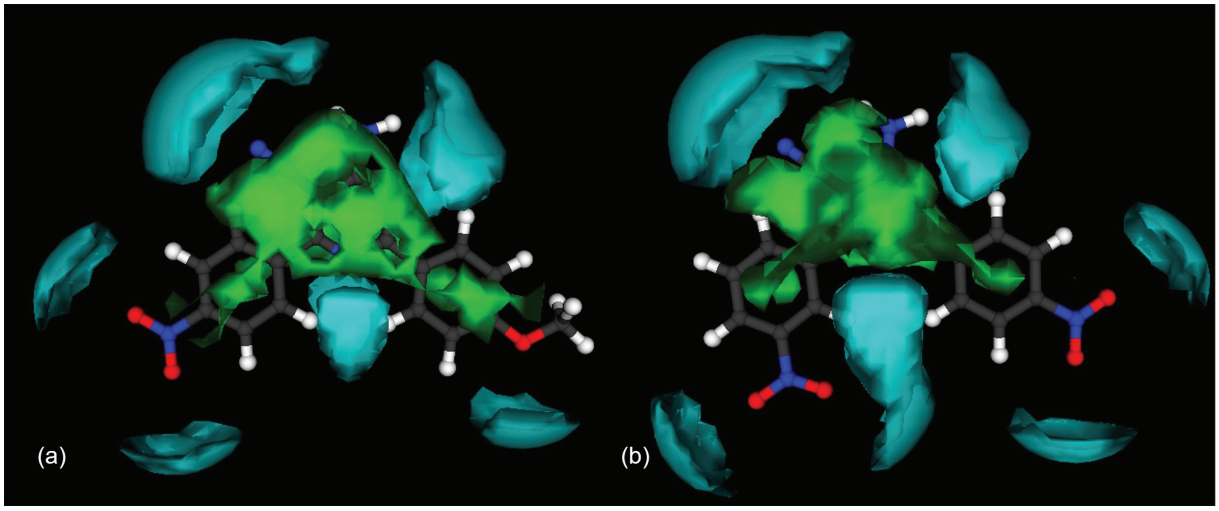


Figure 4. 3D molecular fields calculated with water (cyan) and lipophilic probe (green) for compounds **3d** (a) and **3l** (b). The zones shown in cyan are the hydrophilic regions contoured at −2.8 kcal mol^{-1} and the green zones which are the lipophilic regions contoured at −1.2 kcal mol^{-1}.

of the molecule, and that certain geometries make the resulting dipole moment null, resulting in nonpolarity of the molecule.[43-49]

A consensus model was obtained from calculating the average values of the two models generated by the combination of the MLR and PLS models with Dragon and VolSurf+ descriptors respectively (Table 2). The predicted activity was estimated by taking an average of the predicted pLC_{50} from both QSAR methods. This procedure usually provides better prediction accuracy than the majority of individual models since errant predictions are dampened by the predictions from the other methods. The performance of the MLR model is better for all samples, except for compounds **3l** and **5c**. The consensus model shows almost the same performance as the MLR model and better performance than the PLS model. Finally, the consensus shows the following parameters: Q_{cv}^2 = 0.879, standard deviation of errors of prediction (SDEP) = 0.130.

Although the descriptors related to the electronic effect of the substituent groups were not selected, we also observed that there is an electronic pattern relating the chemical structure of the compounds with the biological activity. This electronic influence is specifically perceived in the radical R_1, where there are variations of the substituent groups.

When R_1 is replaced by a ring deactivator group, activity decreases, as in the case of the nitro group which is a strong ring deactivator, as does molecule **3l** which has one of the smallest biological activities. When R_1 is replaced by an aromatic ring activator group, electro donor, biological activity also increases.

We can compare, for example, molecules **3l**, **3o**, **3c** and **3j**, to **3l** as already mentioned have an aromatic ring deactivating group, whereas molecules **3o**, **3c** and **3j** have a ring activating group. Molecule **3c** has a methoxy in R_1 which is a moderate activator leading to a good biological activity of the molecule, **3d** has a methyl which is a weak activator of the aromatic ring and there is a decrease in biological activity when compared to **3c**, and molecule **3j** has an amine at R_1 which is a strong ring activator, but it greatly elevates molecule's hydrophilia and therefore leaves it inactive.

Table 3. Values of descriptors selected in partial least squares (PLS) for the most active and least active molecules

Compound	logPc-hex[a]	HSA[b]	PSA[c]	PSAR[d]	PHSAR[e]	pLC_{50}[f]
3d	2.818	385.546	123.65	0.242	0.320	4.90
3c	2.705	377.263	123.65	0.246	0.327	4.81
5c	0.628	446.773	123.65	0.216	0.276	4.56
5d	0.741	456.489	123.65	0.213	0.271	4.49
3o	2.668	386.493	112.22	0.225	0.290	4.29
3l	2.057	359.925	149.37	0.293	0.415	3.81
5a	0.013	421.489	112.22	0.210	0.266	3.75

[a]logPc-hex: log partition between cyclohexane and water; [b]HSA: hydrophobic surface area; [c]PSA: polar surface areas; [d]PSAR: polar surface and surface; [e]PHSAR: polar hydrophobic surface area; [f]pLC_{50}: experimental activity.

Thus, we conclude that the substituent group on R_1 being activator of the aromatic ring is better for the activity, but this electro donor group needs to have an ideal hydrophilic character not to make the molecule with large polar surface area.

Conclusions

Compounds **3c** and **3d**, showed excellent activity as larvicides. In the evaluation of larvicidal activity and mammalian cell toxicity, pyrimidines which showed the best results, i.e., low toxicity in macrophages and high mortality rate at the concentration of 100 ppm for *A. aegypti* larvae, were: **3f**, **3j**, **3k** and **5a**. The QSAR models showed some physicochemical properties related to the pharmacokinetic behavior, such as hydrophobicity when the compounds were tested against mosquito larvae. A simple pattern of substituent groups, such as methyl and methoxy at R_1, and a succinimide at R_3, are responsible for the increase of larvicidal activity of the pyrimidine derivatives. This information can be used in further studies.

Supplementary Information

Supplementary data are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

The authors gratefully acknowledge the Brazilian agencies FACEPE, CNPq, and CAPES, for providing financial support under grant: FACEPE/BFP-0094-4.03/17, Z. S. do Monte thanks to the Postgraduate Program in Pharmaceutical Sciences and, Postgraduate Program in Biological Sciences Chemistry Graduate Program at UFPE and, FACEPE and, CNPq/INCT-INAMI fellowships.

Author Contributions

Zenaide S. Monte was responsible for investigation (lead); Daniela M. A. F. Navarro for conceptualization and investigation (lead); Júlio C. R. O. F. Aguiar for investigation (supporting); Jessica S. Nascimento for investigation (supporting); Marcus T. Scotti for validation (lead); Luciana Scotti for formal analysis (lead); Renata P. C. Barros for investigation (supporting); Aline C. S. Santos for investigation (lead); Valéria R. A. Pereira for conceptualization (lead); Emerson P. S. Falcão for conceptualization (lead); Sebastião. J. Melo for project administration (lead).

References

1. Adnan, I. Q.; *Zika Virus Disease from Origin to Outbreak*, 1st ed.; University of Missouri: Columbia, 2018.
2. https://www.cdc.gov/zika/about/index.html, accessed in February 2020.
3. https://www.who.int/emergencies/zika-virus/timeline/en/, accessed in February 2020.
4. Song, G. P.; Hu, D. K.; Tian, H.; Li, Y. S.; Cao, Y. S.; Jin, H. W.; Cui, Z. N.; *Sci. Rep.* **2016**, *6*, 22977.
5. Avula, S. H.; Boddu, S. H.; Yarasani, P.; *Der Pharma Chem.* **2017**, *9*, 68.
6. Yang, T.; Liang, L.; Guiming, F.; Zhong, S.; Ding, G.; Xu, R.; Zhu, G.; Shi, N.; Fan, F.; Liq, Q.; *J. Vector Ecol*. **2009**, *34*, 148.
7. Gorle, S.; Maddila, S.; Chokkakula, S.; Lavanya, P.; Singh, M.; Jonnalagadda, S. B.; *J. Heterocycl. Chem*. **2016**, *53*, 1852.
8. Curtis, C. F.; Pasteur, N.; *Bull Entomol. Res.* **1980**, *71*, 153.
9. Evidente, A.; Cimmino, A.; Berestetskiy, A.; Andolfi, A.; Motta, A.; *J. Nat. Prod*. **2008**, *71*, 1897.
10. Louca, V.; Lucas, M. C.; Green, C.; Majambere, S.; Fillinger, U.; Lindsay, S. W.; *J. Med. Entomol*. **2009**, *46*, 546.
11. Murray, N. E. A.; Quam, M. B.; Wilder-Smith, A.; *J. Clin. Epidemiol*. **2013**, *5*, 299.
12. Bhatt, S.; Gething, P. W.; Brady, O. J.; Messina, J. P.; Farlow, A. W.; Moyes, C. L.; Drake, J. M.; Brownstein, J. S.; Hoen, A. G.; Sankoh, O.; Myers, M. F.; George, D. B.; Jaenisch, T.; William Wint, G. R.; Simmons, C. P.; Scott, T. W.; Farrar, J. J.; Hay, S. I.; *Nature* **2013**, *25*, 504.
13. Agra-Neto, A. C.; Napoleao, T. H.; Pontual, E. V.; Santos, N. D. L.; Andrade Luz, L.; Oliveira, C. M. F.; Melo-Santos, M. A. V.; Coelho, L. C. B. B.; Navarro, D. M. A. F.; Paiva, P. M. G.; *J. Parasitol. Res*. **2014**, *113*, 175.
14. Autran, E. S.; Neves, I. A.; Silva, C. S. B.; Santos, G. K. N.; Camara, C. A. G.; Navarro, D. M. A. F.; *Bioresour. Technol*. **2009**, *100*, 2284.
15. Paixão, E. S.; Teixeira, M. G.; Rodrigues, L. C.; *BMJ Global Health* **2017**, *3*, 000530.
16. Karunamoorthi, K.; Sabesan, S.; *Health Scope* **2013**, *2*, 14.
17. Nicolopoulou-Stamati, P.; Maipas, S.; Kotampasi, C.; Stamatis, P.; Hens, L.; *Front. Public Health* **2016**, *4*, 148.
18. Azam, M. A.; Dharanya, L.; Mehta, C. C.; Sachdeva, S.; *Acta Pharm*. **2013**, *63*, 19.
19. Chen, L.; Jin, Y.; Fu, W.; Xiao, S.; Feng, C.; Fang, B.; Gu, Y.; Li, C.; Zhao, Y.; Liu, Z.; Liang, G.; *ChemMedChem* **2017**, *12*, 1022.
20. de Melo, S. J.; do Monte, Z. S.; da Silva Santos, A. C.; Silva, A. C. C.; Ferreira, L. F. G. R.; Hernandes, M. Z.; Silva, R. O.; Falcão, E. P. S.; Brelaz-de-Castro, M. C. A.; Srivastava, R. M.; Pereira, V. R. A.; *Med. Chem. Res*. **2018**, *27*, 2512.
21. Melo, S. J.; Monte, Z. S.; Srivastava, R. M.; Falcão, E. P. S.; Silva, R. O.; *BR pat. 102017 0151280*, **2017**.

22. Liu, X. H.; Wang, Q.; Sun, Z. H.; Wedge, D. E.; Becnel, J. J.; Estep, A. S.; Tan, C. X.; Weng, J. Q.; *Pest. Manage. Sci*. **2017**, *73*, 953.
23. da Silva, J. B.; Navarro, D. M.; da Silva, A. G.; Santos, G. K.; Dutra, K. A.; Moreira, D. R.; Ramos, M. N.; Espíndola, J. W.; de Oliveira, A. D.; Brondani, D. J.; Leite, A. C.; Hernandes, M. Z.; Pereira, V. R.; da Rocha, L. F.; de Castro, M. C.; de Oliveira, B. C.; Lan, Q.; Merz Jr., K. M.; *Eur. J. Med. Chem.* **2015**, *15*, 162.
24. Silva, M. F.; Bezerra-Silva, P. C.; de Lira, C. S.; de Lima Albuquerque, B. N.; Agra Neto, A. C.; Pontual, E. V.; Maciel, J. R.; Paiva, P. M.; Navarro, D. M.; *Exp. Parasitol*. **2016**, *165*, 64.
25. Finney, D. J.; *Probit Analysis*, 3rd ed.; Cambridge University Press: New York, UK, 1971.
26. https://www.analystsoft.com/br/, accessed in February 2020.
27. *GraphPad Prism*, version 5.00 for Windows; GraphPad Software, San Diego, California, USA, 2015. Available at www.graphpad.com, accessed in February 2020.
28. *Standardizer software*, version for Win32; Cambridge Innovation Center, Cambridge, USA, 2015. Available at www.chemaxon.com, accessed in February 2020.
29. Mayo, S. L.; Olafson, B. D.; Goddard, W. A.; *J. Phys. Chem*. **1990**, *94*, 8897.
30. Cruciani, G.; *VolSurf+ program*, version 1.0.7; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2012. Available at www.moldiscovery.com, accessed in February 2020.
31. Cruciani, G.; Crivori, P.; Carrupt, P.-A.; Testa, B.; *J. Mol. Struct.* **2000**, *503*, 17.
32. Kode, S. R. L.; *Dragon program*, version 7.0.10; Pisa, Italy, 2018. Available at www.chm.kode-solutions.net/index.php, accessed in February 2020.
33. Geladi, P. E.; Kowalski, B. R.; *Anal. Chim. Acta* **1986**, *185*, 1.
34. Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J.; *SIAM J. Sci. Comput*. **1984**, *5*, 735.
35. Cruciani, G.; Pastor, M.; Mannhold, R.; *J. Med. Chem.* **2002**, *45*, 2685.
36. Cramer, R. D.; Bunce, J. D.; Patterson, D. E.; Frank, I. E.; *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
37. Wakeling, I. N.; Morris, J. J.; *J. Chemom*. **1993**, *7*, 291.
38. Todeschini, R.; Ballabio, D.; Consonni, V.; Mauri, A.; Pavan, M.; *Mobydigs*, version 1.1; Milano, Italy, 2009. Available at www.talete.mi.it/, accessed in February 2020.
39. Kubinyi, H.; *Pharmacologyonline* **2011**, *1*, 306.
40. Todeschini, R.; Consonni, V.; *Molecular Descriptors for Chemoinformatics Weinheim*, 2nd ed.; Wiley-VCH: Weinheim, 2009.
41. Martins, C. R.; Lopes, W. A.; Andrade, J. B.; *Quim. Nova* **2013**, *36*, 1248.
42. Imre, G.; Veressc, G.; Volfordd, A.; Farkas, Ö.; *J. Mol. Struct.: THEOCHEM* **2003**, *666*, 51.
43. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.; *Adv. Drug Delivery Rev.* **2001**, *46*, 3.
44. Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S.; *Quant. Struct.-Act. Relat*. **1993**, *12*, 9.
45. Hansch, C.; Leo, A.; *J. Am. Chem. Soc*. **1996**, *18*, 10678.
46. Ernandi, G.; Caron, G.; Pintos, I. G.; Gerbaldo, M.; Perez, D. I.; Gandara, Z.; Martinez, A.; Gomes, G.; Fall, Y.; *Eur. J. Med. Chem*. **2011**, *46*, 860.
47. Silva, L. M.; Alves, M. F.; Scottic, L.; Lopes, W. S.; Scottic, M. T.; *Ecotoxicol. Environ. Saf.* **2018**, *153*, 151.
48. Klein, C. T.; Kaiser, D; Ecker, G.; *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 200.
49. Zamora, I.; Oprea, T.; Cruciani, G.; Pastor, M.; Ungell, A.-L.; *J. Med. Chem*. **2003**, *46*, 25.

Submitted: November 22, 2019
Published online: March 11, 2020