Research Article

Investigation of Antileishmanial Activities of Acridines Derivatives against Promastigotes and Amastigotes Form of Parasites Using Quantitative Structure Activity Relationship Analysis

Samir Chtita,¹ Mounir Ghamali,¹ Rachid Hmamouchi,¹ Bouhya Elidrissi,¹ Mohamed Bourass,² Majdouline Larif,³ Mohammed Bouachrine,⁴ and Tahar Lakhlifi¹

¹MCNSL, Faculty of Science, Moulay Ismail University, Meknes, Morocco

²Faculty of Sciences, Sidi Mohamed Ben Abdellah University, Fez, Morocco

³Separation Process Laboratory, Faculty of Science, Ibn Tofail University, Kenitra, Morocco

⁴High School of Technology, Moulay Ismail University, Meknes, Morocco

Correspondence should be addressed to Samir Chtita; samirchtita@gmail.com

Received 13 August 2016; Revised 22 September 2016; Accepted 27 September 2016

Academic Editor: Michael D. Sevilla

Copyright © 2016 Samir Chtita et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In a search of newer and potent antileishmanial (against promastigotes and amastigotes form of parasites) drug, a series of 60 variously substituted acridines derivatives were subjected to a quantitative structure activity relationship (QSAR) analysis for studying, interpreting, and predicting activities and designing new compounds by using multiple linear regression and artificial neural network (ANN) methods. The used descriptors were computed with Gaussian 03, ACD/ChemSketch, Marvin Sketch, and ChemOffice programs. The QSAR models developed were validated according to the principles set up by the Organisation for Economic Co-operation and Development (OECD). The principal component analysis (PCA) has been used to select descriptors that show a high correlation with activities. The univariate partitioning (UP) method was used to divide the dataset into training and test sets. The multiple linear regression (MLR) method showed a correlation coefficient of 0.850 and 0.814 for antileishmanial activities against promastigotes and amastigotes forms of parasites, respectively. Internal and external validations were used to determine the statistical quality of QSAR of the two MLR models. The artificial neural network (ANN) method, considering the relevant descriptors obtained from the MLR, showed a correlation coefficient of 0.933 and 0.918 with 7-3-1 and 6-3-1 ANN models architecture for antileishmanial activities against promastigotes and amastigotes forms of parasites, respectively. The applicability domain of MLR models was investigated using simple and leverage approaches to detect outliers and outsidings compounds. The effects of different descriptors in the activities were described and used to study and design new compounds with higher activities compared to the existing ones.

1. Introduction

For hundreds of years, Leishmaniasis, a disease caused by a number of species of protozoan parasites belonging to the genus *Leishmania*, is recognized as an important public health problem throughout the world [1–3]. Currently, the leishmaniasis are considered to be endemic in 88 countries and, according to World Health Organization (WHO) [4], twelve million people are infected, with about two to three

million new cases each year, and 350 million people are under risk of infection; it is a major public health problem particularly in Latin America, Africa, and Asia [5–9]. To date, no vaccine against any clinical form of Leishmaniasis has been commercialized and treatment relies only on chemotherapy, which has been based on the use of pentavalent antimonial drugs. Other medications, such as pentamidine and amphotericin B, have been used as alternative drugs for resistant parasites. With the emergence of some resistant

strains, the toxicity of current drugs, severe side effects, and high cost and/or restricted therapeutic spectrum, a need for development of new and safer drugs is warranted [2, 3, 10]. A great number of natural and synthetic compounds have been tested in the past years in antileishmanial assays. Their structures are diverse and often contain nitrogen heterocycles such as quinolines, pyrimidines, acridines, phenothiazines, and indoles [11–13].

Many experiments have been performed with the compounds bearing the heterocyclic ring structures to explore their effectiveness against Leishmania. These studies suggested their similar pharmacophoric feature of the heterocyclic scaffold as a potential target for drug discovery of antileishmanial drugs [14].

Leishmania parasites exist in two forms, one is promastigotes and the other is amastigotes. The promastigotes are flagellated and found in sand fly, while the amastigotes are ovoid and nonflagellated form of *Leishmania* [15]. Antileishmanial activity is performed against promastigotes and then amastigotes forms of parasites. Heterocyclic system may also be formed by fusion with other rings, either carbocyclic or heterocyclic.

Since their discovery in the 1880s, acridines family have demonstrated a broad spectrum of pharmacological properties [16]. First employed as antibacterial agents during the beginning of the twentieth century [17]. They have been rapidly revealing interesting antiproliferative activities against both protozoa and tumor cells [18, 19]. Consequently, they have been extensively used in antiparasitic chemotherapy and a wide range of new acridines derivatives have been synthesized and successfully assessed for their antileishmanial activities [20, 21].

In order to open a new way in antileishmanial drug research, a series of sixty acridines derivatives were synthesized [22–24] and studied for their antileishmanial (against promastigotes and amastigotes form of parasites) activities. The aim of this study was to develop a QSAR model able to correlate the structural features of the acridines derivatives with their biological activities.

In general, the QSAR methods are based on the assumption that the activity of a certain chemical compound related to its structure through a certain mathematical algorithm. This relationship can be used in the prediction, interpretation, and assessment of new compounds with desired activities reducing and rationalizing time, efforts, and cost of synthesis and new product development.

The basic assumption to drive a QSAR model is presented due to a mathematical function of the chemical properties which is related to the effect (activity). Therefore, the effect is like the function “ f ” of the chemical properties “ x ”: $y = f(x)$. To find this algorithm, we use a number of chemical compounds with known values of the studied effect (y). For each chemical compound, we calculate a series of parameters (called chemical descriptors). Then, we find an algorithm that provides a quite accurate value, similar to the real experimental value. The final step is to check if the obtained algorithm is able to predict the activity values for other chemicals not used to build up the model (external validation).

Indeed, it is very important to generate a model which worked not only for the chemical substances used within the training set but also for other similar chemicals. Consequently, the challenge is to define the correct statistical properties of the model.

2. Materials and Methods

The current QSAR study investigates prediction and interpretation of the studied compounds and was also used for designing new proposed compounds by using linear and nonlinear methods. It consists of four stages: selection of dataset and generation of molecular descriptors, descriptive analysis, statistical analysis (prediction and evaluation), and suggestion of novel compounds.

A flow chart for the development of the QSAR model along with the various validation methods used in this work is demonstrated in Figure 1.

2.1. Selection of Dataset and Generation of Molecular Descriptors

2.1.1. Selection of Dataset. In this stage, the datasets of the antileishmanial activities (against promastigotes and amastigotes forms of parasites) of various acridine derivatives (4-monosubstituted acridines, 3,6-disubstituted acridines, 4,5-disubstituted acridines, and 7-monosubstituted 9-chloro and 9-amino-2-methoxy acridines) were collected from previous works [22–24]. The molecular structures of the studied molecules with their antileishmanial activities are presented in Table 1. All experimental IC_{50} antileishmanial activity values (μM) were converted to the negative logarithm of IC_{50} ($pIC_{50} = -\log_{10}(IC_{50})$).

2.1.2. Molecular Descriptors Generation. A wide variety of molecular descriptors was calculated using Gaussian 03, ACD/ChemSketch, Marvin Sketch, and ChemOffice programs [25–28] to predict the correlation between these descriptors for the studied molecules with their antileishmanial activities and to develop linear (multiple linear regression (MLR)) and nonlinear (artificial neural network (ANN)) models. Tables 3 and 4 show the selected descriptors (using the PCA method; see more in descriptive analysis results) to be used in this study.

2.2. Descriptive Analysis. In this stage, the principal component analysis (PCA) was used to determine the nonlinearity and nonmulticollinearity among variables (descriptors) and to select descriptors that correlate with the activity. After that, the univariate partitioning (UP) method was used to form dissimilar clusters of compounds, to which the query compounds would be compared for determining the degree of similarity and dividing the dataset into training and test sets.

2.3. Statistical Analysis (Models Development and Evaluation). In this stage, linear and nonlinear QSAR models were developed and evaluated to predict the activities of the test compounds. The study we conducted consists of the multiple

FIGURE 1: Flow chart of the methodology used in this work.

linear regression (MLR) available in the XLSTAT software and the artificial neural network (ANN) available in the Matlab software.

In order to propose models and to evaluate quantitatively the physicochemical effects of the substituents on the activities of molecules, we submitted the data matrix constituted obviously from the used variables (descriptors) corresponding to the dataset molecules to a descendant MLR analysis and to an ANN. We use the coefficients R , R^2 , R^2_{adj} , MSE, and P value to select the best regression performance [29], where R is the correlation coefficient; R^2 is the coefficient of determination; R^2_{adj} is the coefficient adjusted for degrees of freedom. Mean squared error (MSE) is the standard error of the coefficient of each descriptor and of the global model, which gives an indication of the valid inclusion of a descriptor in a QSAR model. P value is the probability (P) of Fisher statistics (F), which gives an indication of the probability that a QSAR is a chance correlation.

In order to assess the significance of the models and accurate prediction ability for new compounds:

- we use an internal validation procedure (leave-one-out cross validation), whereby one compound is removed and the rebuilt model with the remaining molecules is used to predict the response of the eliminated compound. This one is then returned and a second is removed, and the cycle is repeated, and

so on, until all compounds have been removed one by one, and an overall correlation coefficient R_{cv} is computed;

- after the model is built, an external prediction is necessary. In this one, the obtained model was used to predict the activities of a test set comprising compounds that are similar to those not used in the training set. This is usually performed by splitting a dataset into a training set and a test set, typically in a 1/5 ratio. Further, before performing the external validation of a model, it is very important to check for the presence of systematic error that violates the basic assumptions of the least squares regression model. If high systematic error (bias) is present in the model, then such model should be discarded and performing any external validation test is of no use on such biased model. Xternal Validation Plus is a tool that checks the presence of systematic errors in the model and further computes all the required external validation parameters, while judging the performance of actual prediction quality of a QSAR model based on recently proposed MAE-based criteria [30];
- a model is valid only within its training domain and new molecules must be considered as belonging to the domain before the model is applied (OECD Principle 3 [31]). Without applicability domain (AD),

TABLE I: Chemical structure and antileishmanial activities of studied compounds (see Figure 8).

Number	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	PRO ^c	pIC ₅₀ ^b	AMA ^d
1	H	NHCOCH ₃	H	H	H	NHCOCH ₃	H	-2.533	-0.653	
2	H	NHCOC ₂ H ₅	H	H	H	NHCOC ₂ H ₅	H	-2.204	-0.886	
3	H	NHCOC ₃ H ₇	H	H	H	NHCOC ₃ H ₇	H	-1.669	-1.107	
4	H	NHCOPh	H	H	H	NHCOPh	H	-0.634	-0.041	
5	H	NHCO-p-PhCl	H	H	H	NHCO-p-PhCl	H	—	0.699	
6	H	NHCO-p-PhF	H	H	H	NHCO-p-PhF	H	-0.230	1.523	
7	H	NHCO-p-PhOMe	H	H	H	NHCO-p-PhOMe	H	-0.681	-0.041	
8	H	NHCO-m,p-Ph(OMe) ₂	H	H	H	NHCO-m,p-Ph(OMe) ₂	H	-1.207	0.097	
9	H	NHCOMe	H	H	H	NHCOMe	H	-2.270	Tox ^a	
10	H	NHCOMe	H	H	H	NHCO-p-PhCl	H	-1.061	-0.462	
11	H	NHCOMe	H	H	H	NHCO-p-PhF	H	-2.123	0.174	
12	H	NHCOMe	H	H	H	NHCO-p-PhOMe	H	-1.939	Tox ^a	
13	H	NHCOMe	H	H	H	NHCO-m,p-Ph(OMe) ₂	H	—	-0.114	
14	H	H	CH ₃	NH ₂	OMe	H	H	0.301	0.398	
15	H	H	CH ₂ OH	NH ₂	OMe	H	H	-0.732	-0.613	
16	H	H	CH ₂ Br	NH ₂	OMe	H	H	-0.556	-0.663	
17	H	H	(CH ₂) ₂ OCOOMe	NH ₂	OMe	H	H	-0.380	-0.114	
18	H	H	(CH ₂) ₂ OCO(CH ₂) ₂ CH ₃	NH ₂	OMe	H	H	-0.491	-0.398	
19	H	H	(CH ₂) ₂ OCOCH ₂ CH(CH ₃) ₂	NH ₂	OMe	H	H	-0.342	Tox ^a	
20	H	H	(CH ₂) ₂ OCOPh	NH ₂	OMe	H	H	0.398	0.699	
21	H	H	(CH ₂) ₂ OCOPhF	NH ₂	OMe	H	H	-0.114	0.222	
23	H	H	(CH ₂) ₂ OCOPhOMe	NH ₂	OMe	H	H	-0.279	Tox ^a	
24	H	H	CH ₃	Cl	OMe	H	H	-0.041	-0.362	
25	H	H	CH ₂ OH	Cl	OMe	H	H	-2.262	-0.959	
26	H	H	CH ₂ Br	Cl	OMe	H	H	-1.703	-1.170	
27	H	H	(CH ₂) ₂ OCOOMe	Cl	OMe	H	H	-2.178	-1.877	
28	H	H	(CH ₂) ₂ OCO(CH ₂) ₂ CH ₃	Cl	OMe	H	H	-1.707	-1.628	
29	H	H	(CH ₂) ₂ OCOCH ₂ CH(CH ₃) ₂	Cl	OMe	H	H	-1.446	-1.561	
30	H	H	(CH ₂) ₂ OCOPh	Cl	OMe	H	H	-2.135	Tox ^a	
31	H	H	(CH ₂) ₂ OCOPhF	Cl	OMe	H	H	-2.295	-1.645	
32	H	H	(CH ₂) ₂ OCOPhCl	Cl	OMe	H	H	-2.194	-2.191	

TABLE I: Continued.

Number	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	R ₇	PRO ^c	pIC ₅₀ ^b	AMA ^d
33	H	H	(CH ₂) ₂ OCOPhOMe	Cl	OMe	H	H	-2.098	-2.088	
34	CH ₂ NH ₂	H	H	H	H	H	H	-0.230	-0.531	
35	CH ₂ OH	H	H	H	H	H	H	Tox ^a	-1.515	
36	CH ₂ NHCO(CH ₂) ₃ Cl	H	H	H	H	H	H	-1.655	-1.427	
37	CH ₂ OCO(CH ₂) ₃ Cl	H	H	H	H	H	H	-2.200	-2.185	
38	CH ₂ NHCOCH=CH ₂	H	H	H	H	H	H	-2.245	-0.833	
39	CH ₂ OCOCH=CH ₂	H	H	H	H	H	H	-1.464	—	
40	CH ₂ NHCOPh	H	H	H	H	H	H	-1.654	-1.121	
41	CH ₂ OCOPh	H	H	H	H	H	H	-1.885	Tox ^a	
42	CH ₂ NHCO-p-PhF	H	H	H	H	H	H	-1.815	—	
43	CH ₂ OCO-p-PhF	H	H	H	H	H	H	-1.741	Tox ^a	
44	CH ₂ NHCO-p-PhCl	H	H	H	H	H	H	Tox ^a	-1.004	
45	CH ₂ OCO-p-PhCl	H	H	H	H	H	H	-1.790	-1.433	
46	CH ₂ NHCO-p-PhOMe	H	H	H	H	H	H	-1.513	-0.973	
47	CH ₂ OCO-p-PhOMe	H	H	H	H	H	H	-1.471	-0.869	
48	CH ₂ NHCO-p-PhNNMe ₂	H	H	H	H	H	H	-1.539	-0.672	
49	CH ₂ OCO-p-PhNNMe ₂	H	H	H	H	H	H	-1.819	Tox ^a	
50	CH ₂ NH ₂	H	H	H	H	H	CH ₂ NH ₂	-0.820	-0.531	
51	CH ₂ OH	H	H	H	H	H	CH ₂ OH	Tox ^a	0.222	
52	CH ₂ NHCO(CH ₂) ₃ Cl	H	H	H	H	H	CH ₂ NHCO(CH ₂) ₃ Cl	-0.663	-0.813	
54	CH ₂ NHCOCH=CH ₂	H	H	H	H	H	CH ₂ NHCOCH=CH ₂	-1.061	Tox ^a	
56	CH ₂ NHCOPh	H	H	H	H	H	CH ₂ NHCOPh	-0.556	-1.236	
57	CH ₂ OCOPh	H	H	H	H	H	CH ₂ OCOPh	Tox ^a	-0.716	
58	CH ₂ NHCO-p-PhF	H	H	H	H	H	CH ₂ NHCO-p-PhF	-0.756	-0.255	
60	CH ₂ NHCO-p-PhCl	H	H	H	H	H	CH ₂ NHCO-p-PhCl	—	-1.562	
62	CH ₂ NHCO-p-PhOMe	H	H	H	H	H	CH ₂ NHCO-p-PhOMe	-1.797	Tox ^a	
63	CH ₂ OCO-p-PhOMe	H	H	H	H	H	CH ₂ OCO-p-PhOMe	Tox ^a	-1.825	
64	CH ₂ NHCO-p-PhNNMe ₂	H	H	H	H	H	CH ₂ NHCO-p-PhNNMe ₂	-0.940	-0.724	
65	CH ₂ OCO-p-PhNNMe ₂	H	H	H	H	H	CH ₂ OCO-p-PhNNMe ₂	Tox ^a	-1.667	

^aToxic: toxicity observed on human macrophages at concentrations that did not display antileishmanial activity; -p-: para; -m-: meta; ^bpIC₅₀ = -log (IC₅₀); ^cPRO: antileishmanial activity against promastigotes parasite form; ^dpIC₅₀: antileishmanial activity against promastigotes parasite form.

FIGURE 2: The architecture used in our study of the artificial neural network.

each model can predict the activity of any compound, even with a completely different structure from those included in the study. Therefore, the AD is a tool to find out compounds that are outside the applicability domain of the built QSAR model and it detects outliers present in the training set compounds. There are several methods for defining the applicability domain (AD) of QSAR models [32], but the most common one is determining the leverage values h_i ($h_i = x_i^T (X^T X)^{-1} x_i$ ($i = 1, 2, \dots, n$)) for each compound, where x_i is the descriptor row-vector of query compound, X is $n \times k-1$ matrix of k model descriptor values for n training set compounds, and the superscript T refers to the transpose of matrix/vector [32, 33]. In this study, we use Williams plot; in this plot, the applicability domain is established inside a squared area within standard deviation $\pm x$ (in this study $x = 3$; "three-sigma rule" [34]) and a leverage threshold h^* ($h^* = 3 * (k + 1)/n$) [35], where n is the number of training set compounds; k is the number of model descriptors. The leverage (h) greater than the warning leverage (h^*) suggested that the compound was very influential on the model [36]. The results of the leverage approach were compared with that of the simple approach introduced by Roy et al. [37].

2.3.1. Multiple Linear Regression (MLR). The descendent multiple linear regression (MLR) analysis, based on the elimination of aberrant descriptors (one by one) until a valid model (including the critical probability: P value < 0.05 for all descriptors and the model complete), was employed to find a linear model of the activity of interest, which takes the following form:

$$Y = a_0 + \sum_{i=1}^n a_i x_i, \quad (1)$$

where Y is the studied activity, which is the dependent variable. a_0 is the intercept of the equation. x_i are the molecular descriptors. a_i are the coefficients of those descriptors.

This method is one of the most popular methods of QSAR/QSPR thanks to its simplicity in operation, reproducibility, and ability to allow easy interpretation of the features used. The important advantage of the linear regression analysis is that it is highly transparent; therefore, the algorithm is available and predictions can be made easily [38]. It has served also to select the descriptors used as the input parameters in the artificial neural network (ANN).

TABLE 2: Software packages used in this work.

Drawing chemical structures	Marvin Sketch, ACD/ChemSketch, and ChemBioDraw
Generating 3D structures	Gauss View 3.0 and ChemBio3D
Calculating chemical descriptors	Gaussian 03, Marvin Sketch 6.2, ChemSketch, and ChemBio3D
Developing QSAR models	XLSTAT 2009 and Matlab 7.9.0 (version 2011)

2.3.2. Artificial Neural Networks (ANN). The artificial neural networks (ANN) are used in order to increase the probability of characterizing the compounds and to generate a predictive QSAR model between the set of molecular descriptors obtained from the MLR models and the observed activities values. The ANN model is done on the MATLAB R2009b software. It consists of three layers of neurons, called input layer, hidden layer, and output layer (Figure 2). The input layer formed by a number of neurons equal to the number of descriptors obtained in the multiple linear regression models and the output layer represents the calculated activity values. For determination of the number of hidden neurons in the hidden layer, where all calculations of parameter optimization of neural networks are made, a parameter ρ has been proposed. The parameter ρ plays a major role in determining the best artificial neural network architecture [39, 40]; ρ is defined as follows:

$$\rho = \frac{\text{number of data points in the training set}}{\text{sum of the number of connections in the ANN}}. \quad (2)$$

In order to avoid overfitting or underfitting, it is recommended that the value of ρ should be between 1.00 and 3.00; if $\rho < 1$, the network simply memorizes the data, whereas if $\rho > 3$, the network is not able to generalize [41].

2.4. Software Packages Used in Our QSAR Development Study.

There are various free and commercial software available for QSAR development. These include specialized software for drawing chemical structures, generating 3D structures, calculating chemical descriptors, and developing QSAR models. The software packages used in our works are represented in Table 2.

3. Results and Discussions

3.1. Principal Component Analysis (PCA). PCA is a useful statistical technique for summarizing all the information encoded in the structures of the compounds. It is very helpful

TABLE 3: The absolute values of correlation coefficients with descriptors and activities.

Descriptors	<i>r</i>
Promastigotes	
Henry's law constant (K_H)	0.130
Ideal gas thermal capacity (IGTC)	0.108
Number of H-Bond acceptors (NHA)	0.299
Number of H-Bond donors (NHD)	0.250
Balaban index (<i>J</i>)	0.133
Shape coefficient (<i>I</i>)	0.135
Total valence connectivity (TVC)	0.292
N%	0.238
H%	0.260
Gibbs free energy (<i>G</i>)	0.119
Dipole moment (μ)	0.213
Electronegativity (χ)	0.599
Electrophilicity index (ω)	0.630
Molecular topological index (MTI)	0.116
Polar surface area (PSA)	0.282
Total connectivity (TC)	0.269
Wiener index (<i>W</i>)	0.124
E_{HOMO}	0.554
E_{LUMO}	0.632
E_{Gap}	0.227
Amastigotes	
Total energy, <i>E</i>	0.160
E_{HOMO}	0.286
Electronegativity (χ)	0.316
Polar surface area (PSA)	0.391
O%	0.105
Index of refraction (<i>n</i>)	0.548
Density (<i>d</i>)	0.242
Critical pressure (CP)	0.176
Gibbs free energy (<i>G</i>)	0.195
Henry's law constant (K_H)	0.387
Number of H-Bond acceptors (NHA)	0.198
Number of H-Bond donors (NHD)	0.587
Number of rotatable bonds (NRB)	0.139
Sum of valence degrees (SVD)	0.146
Dipole moment (μ)	0.224
E_{LUMO}	0.338
Electrophilicity index (ω)	0.335
N%	0.436
C%	0.155
Surface tension (γ)	0.553
Boiling point (TB)	0.254
Critical temperature (CT)	0.276
Heat of formation (H°)	0.152
log <i>P</i>	0.248
Melting point (<i>T</i>)	0.314
Partition coefficient (PC)	0.163
Shape coefficient (<i>I</i>)	0.176
Winner index (<i>W</i>)	0.103

for understanding the distribution of the compounds. This is an essentially descriptive statistical method, which aims to extract the maximum of information contained in the dataset compounds [42, 43]. In this work, the PCA is used to overview the examined compounds for similarities and dissimilarities and to select descriptors that show a high

TABLE 4: Descriptors selected by the principal component analysis (PCA) and software packages used in the calculation of descriptors.

Software	Descriptors
Gaussian 03	Highest occupied molecular orbital energy E_{HOMO} (eV); lowest unoccupied molecular orbital energy E_{LUMO} (eV); hardness η (eV) $= (E_{LUMO} - E_{HOMO})/2$; electronegativity χ (eV) $= -(E_{LUMO} + E_{HOMO})/2$; electrophilicity index ω (eV) $= \chi^2/2\eta$; total energy <i>E</i> (eV); dipole moment μ (Debye); energy gap between E_{HOMO} and E_{LUMO} values E_{Gap} (eV)
ChemOffice	Heat of formation H° (kJ mol ⁻¹); Gibbs free energy <i>G</i> (kJ mol ⁻¹); ideal gas thermal capacity (IGTC) (J mol ⁻¹ K ⁻¹); melting point <i>T</i> (Kelvin); critical temperature (CT) (Kelvin); boiling point (TB) (Kelvin); critical pressure (CP) (Bar); Henry's law constant K_H ; total valence connectivity (TVC); partition coefficient (PC); molecular topological index (MTI); number of rotatable bonds (NRB); shape coefficient <i>I</i> ; sum of valence degrees (SVD); total connectivity (TC)
ChemSketch	Percent ratios of nitrogen, hydrogen, oxygen, and carbon atoms (N%; H%; O%; and C%); surface tension γ (dyne/cm); index of refraction (<i>n</i>); density (<i>d</i>)
Marvin Sketch	log <i>P</i> ; Winner index (<i>W</i>); number of H-Bond acceptors (NHA); number of H-Bond donors (NHD); Balaban index (<i>J</i>); polar surface area (PSA) ($Å^2$)

correlation with the response activity; this one gives extra weight because it will be more effective at prediction [44]. Tables 3 and 4 present the descriptors with a correlation coefficient with the activity higher, in absolute value, than 0.1. The absence of any serious multicollinearity between the descriptors present in the model was confirmed by the correlation matrix.

3.2. Univariate Partitioning (UP). The aim of the UP was the recognition of groups of objects based on their similarity; this method is based on the criterion of partitioning proposed by Fisher [45]. In this study, the division of the dataset into training and test sets has been performed. In this one, from each obtained cluster, one compound for the training set was selected randomly to be used as test set compound. The partitioning results are given in Table 5.

3.3. Multiple Linear Regression (MLR). The QSAR analysis was performed using the values of the chemical descriptors selected by the PCA method and, on the other hand, the experimental values of the antileishmanial activities for 60 of the acridines derivatives (effect). Table S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2016/5137289> shows the value of each molecular descriptor that is configured in established MLR models and the QSAR model

TABLE 5: The partitioning results of compounds into training and test sets.

Classes	The test set compounds	
	Promastigotes 3, 6, 10, 14, 17, 21, 37, 38, 42, 64	Amastigotes 5, 6, 26, 27, 29, 36, 37, 44, 46, 50
1	1, 8, 37, 40, 56	1, 10, 15, 16, 34, 48, 50, 57, 64
2	2, 14, 29, 34, 48	2, 25, 38, 44, 47, 51, 52
3	3, 18, 19	3, 26, 40, 56
4	4, 24, 25, 27, 30, 42, 58	4, 7, 8, 11, 13, 17, 21, 46
5	6, 26	5, 14, 20
6	7, 9, 12, 21, 33, 39, 41, 47	6
7	10, 11, 16, 31, 32, 43, 45	18, 29, 58
8	15, 17, 23, 28, 36, 49, 52, 54	27, 63
9	20, 38, 46, 62	24, 28, 31, 35, 36, 45, 60, 65
10	50, 64	32, 33, 37

built is represented by the following equations and the values of the statistical parameters.

MLR Model for pIC₅₀ Promastigotes

$$\begin{aligned} \text{pIC}_{50}(\text{PRO}) = & 3.030 - 8.214 \cdot 10^{-02} \text{N\%} + 0.239\mu \\ & - 1.264\omega - 0.233\text{NHA} + 0.732\text{NHD} \\ & + 3.311 \cdot 10^{-05} \text{MTI} + 3.184 \cdot 10^4 \text{TVC}. \end{aligned} \quad (3)$$

Statistical Parameters. $R^2 = 0.723$; $R = 0.850$; $R^2_{\text{adj}} = 0.664$; $\text{MSE} = 0.189$; P value $< 10^{-4}$.

MLR Model for pIC₅₀ Amastigotes

$$\begin{aligned} \text{pIC}_{50}(\text{AMA}) = & -7.314 + 1.191E_{\text{HOMO}} \\ & + 2.543 \cdot 10^{-02} \text{CT} - 0.350K_{\text{H}} \\ & - 1.632 \cdot 10^{-02} T + 0.652\text{NHA} \\ & + 1.262\text{NHD}. \end{aligned} \quad (4)$$

Statistical Parameters. $R^2 = 0.663$; $R = 0.814$; $R^2_{\text{adj}} = 0.598$; $\text{MSE} = 0.208$; P value $< 10^{-4}$.

- (i) For the two models, P value is lower than 0.0001; it means that we would be taking a lower than 0.01% risk in assuming that the null hypothesis (no effect of the explanatory variables) is wrong and that the regressions equations have statistical significance. Therefore, we can conclude with confidence that the selected variables do bring a significant amount of information.
- (ii) Higher value of R^2 and R^2_{adj} and lower mean squared error (MSE) indicate that the two proposed models are predictive and reliable.
- (iii) Our obtained models were validated internally by the leave-one-out cross validation technique; the cross validation coefficient R^2_{cv} for the two models was

determined based on the predictive ability of the model. The value of R^2_{cv} is higher than 0.5 ($R^2_{\text{cv}} = 0.536$ for the MLR pIC₅₀ (PRO) model and $R^2_{\text{cv}} = 0.525$ for the MLR pIC₅₀ (AMA) model); indicating good predictability of the model.

- (iv) True predictive power of these models is to test their ability to predict perfectly pIC₅₀ of compounds from an external test set (compounds that were not used for the developed model); pIC₅₀ of the remaining set of 10 compounds are deduced from the quantitative models proposed with the compounds used in training set by MLR. These models will be able to predict the activities of test set molecules in agreement with the experimentally determined value. The observed and calculated pIC₅₀ values are given in Table 6. The predictive capacity of the models was judged, the higher value of R^2_{test} ($R^2_{\text{test}} = 0.660$ for the MLR pIC₅₀ (PRO) model and $R^2_{\text{test}} = 0.718$ for the MLR pIC₅₀ (AMA) model) indicates the improved predictability of the model.
- (v) Xternal Validation Plus indicates the absence of systematic errors in the model and a moderate performance of prediction quality of a QSAR model based on proposed MAE-based criteria (Table S2).
- (vi) The values of calculated activities from (3) and (4) are given in Table S1 and the correlations of calculated and observed activities values are illustrated in Figure 3.

In the first model (pIC₅₀ PRO), the descriptors influencing negatively the activities are the percent ratio of nitrogen (N%), the number of H-Bond acceptors (NHA), and the electrophilicity index (ω), and the parameters influencing positively the activities are the dipole moment (μ), the number of H-Bond donors (NHD), the molecular topological index (MTI), and the total valence connectivity (TVC).

- (i) The number of H-Bond acceptors (NHA) has a negative sign in the model, which suggests that increased activity can be achieved by decreasing the number

TABLE 6: The values of chemical descriptors and the observed and predicted activities using the MLR models for the test set.

Number	N%	μ	Test set (promastigotes)					pIC ₅₀ PRO	
			ω	NHA	NHD	MTI	TVC	Obs.	MLR
3	12.030	4.259	5.024	3	2	14075	$3.891 \cdot 10^{-07}$	-1.669	-2.047
6	9.270	5.640	4.654	5	2	28763	$1.716 \cdot 10^{-10}$	-0.230	-1.014
10	10.780	5.730	5.267	3	2	16459	$4.902 \cdot 10^{-08}$	-1.061	-1.832
14	11.760	2.698	3.062	3	1	4387	$3.050 \cdot 10^{-05}$	0.301	-0.012
17	8.580	4.969	3.431	4	1	10329	$5.188 \cdot 10^{-07}$	-0.380	-0.664
21	7.180	5.214	3.498	5	1	18600	$1.334 \cdot 10^{-08}$	-0.114	-0.551
37	4.460	6.939	4.779	2	0	8245	$4.076 \cdot 10^{-06}$	-2.200	-1.781
38	10.680	5.122	4.374	2	1	6388	$5.083 \cdot 10^{-06}$	-2.245	-1.512
42	8.480	5.260	4.457	3	1	12005	$1.307 \cdot 10^{-07}$	-1.815	-1.608
64	13.170	5.297	4.388	5	2	47653	$1.201 \cdot 10^{-10}$	-0.940	-0.455

Number	E_{HOMO}	K_{H}	Test set (amastigotes)				pIC ₅₀ AMA	
			T	NHA	NHD	CT	Obs.	MLR
5	-5.850	1064.61	1300.866	3	2	19.110	0.699	-0.787
6	-5.785	1005.95	1247.456	5	2	18.715	1.523	0.332
26	-5.665	583.700	926.651	2	0	7.815	-1.170	-1.456
27	-5.603	622.270	930.050	3	0	8.827	-1.877	-1.628
29	-5.640	618.850	938.380	3	0	9.007	-1.561	-1.467
36	-5.966	676.650	1037.977	2	1	12.201	-1.427	-0.774
37	-6.025	552.770	930.967	2	0	8.348	-2.185	-1.457
44	-5.919	749.400	1083.064	2	1	12.828	-1.004	-0.978
46	-5.790	752.980	1078.419	3	1	13.925	-0.973	-0.733
50	-5.461	625.750	913.563	3	2	13.449	-0.531	-1.028

FIGURE 3: Correlations of observed and predicted activities (training set in black and test set in red) values calculated using MLR models.

of heteroatoms (nitrogen or oxygen atoms), mostly the nitrogen ones for decreasing the ratio of nitrogen (N%) having also a negative sign in the model.

- (ii) The electrophilicity index ω has a negative sign in the model, which suggests that increased activity can be achieved by decreasing the electrophilicity of the

acridine derivatives (a high value of electrophilicity describes a good electrophile, while a small value of electrophilicity describes a good nucleophile).

- (iii) The dipole moment μ has a positive sign in the model, which suggests that increased activity can be achieved by increasing the polarity of the acridine derivatives.

- (iv) The number of H-Bond donors (NHD) has a positive sign in the model, which suggests that increased activity can be achieved by increasing of the heteroatom with one or more hydrogen atoms.
- (v) The molecular topological index (MTI) coefficient (topological parameter) has a negative sign in the model, which suggests that increased activity can be achieved by increasing the flexibility of the substituent side chain.
- (vi) The total valence connectivity (TVC) (topological parameter) has a positive sign in the model, which suggests that increased activity can be achieved by increasing in branching of the acridine derivatives, because the overall connectivity increases with both molecule size and complexity, as expressed in branching and cyclicity of molecular skeleton [46].

In the second model (pIC_{50} AMA), the descriptors influencing negatively the activity are Henry's law constant (K_H) and the melting point (T), and the parameters influencing positively the activities are E_{HOMO} energy, the critical temperature (CT), the number of H-Bond acceptors (NHA), and the number of H-Bond donors (NHD).

- (i) Henry's law constant (K_H) has a negative sign in the model, which suggests that increased activity can be achieved by decreasing solubility; consequently, by decreasing the polarity, $K_H = c/p$ (c : solubility; p : pressure).
- (ii) The melting point (T) has a negative sign in the model, which suggests that increased activity can be achieved by decreasing the polarity and the branching of molecule (increasing branching makes the molecule less compact. As the surface area of the molecule decreases, it will become more compact and thus easier to pack).
- (iii) The highest occupied molecular orbital energy E_{HOMO} (negative values) has a positive sign in the model, which suggests that the higher of E_{HOMO} , the weaker donating electron ability, is showing the fact that the nucleophilic reaction occurs more easily and the activity of the compound is higher.
- (iv) The critical temperature (CT) has a positive sign in the model, which suggests that increased activity can be achieved by increasing the critical temperature.
- (v) The number of H-Bond acceptors (NHA) has a positive sign in the model, which suggests that increased activity can be achieved by increasing the number of heteroatoms (nitrogen or oxygen atoms).
- (vi) The number of H-Bond donors (NHD) has a positive sign in the model, which suggests that increased activity can be achieved by increasing of the heteroatoms with one or more hydrogen atoms.

Comparing the importance of each descriptor on pIC_{50} of acridines, we must know the standardized coefficient and the t -test values of them in the MLR equations. The bigger the

absolute value of the t -test value is, the greater the influence of the descriptor is.

In (3), the t -test values are 1.773, -4.035 , 7.408, 1.720, -3.361 , -2.429 , and -4.570 for $N\%$, μ , ω , NHA, NHD, MTI, and TVC, respectively. Moreover, in (4), the t -test values are 3.196, 3.556, -3.764 , -3.802 , -3.838 , and -6.012 for T , K_H , CT, E_{HOMO} , NHA, and NHD, respectively.

This meant that the t -test values of ω , μ , NHD, and TVC are both larger than those of the other descriptors, which indicates that, in this model, the influence of these descriptors on activity is stronger than that of the others. It shows also the importance of electrophilicity index in the prediction of pIC_{50} PRO. Moreover, the t -test value of NHD is larger than that of the other descriptors, which indicates that the influence of this descriptor, in this model, on activity is stronger than that of the others.

In the conclusion, these results illustrate that, to increase the antileishmanial activity against promastigotes parasites, we will decrease the electrophilicity and increase the branching, the polarity, and the number of hydrogen atoms attached in the heteroatom of the acridine derivatives. Moreover, to increase the antileishmanial activity against amastigotes parasites, we will decrease the solubility, the polarity, and the branching and increase the electrophilicity and the number of heteroatoms mostly attached in the hydrogen atoms and the critical temperature of the acridine derivatives.

(i) *Detection of Outliners: Applicability Domain (AD)*. The applicability domain (AD) of these models was evaluated by leverage analysis expressed as Williams plot (Figures 4 and 5), in which the standardized residuals (r) and the leverage threshold values ($h^* = 0.585$ and 0.553 for pIC_{50} (PRO) and pIC_{50} (AMA), resp.) were plotted. Any new value of predicted pIC_{50} data must be considered reliable only for those compounds that fall within this AD on which the model was constructed.

From these figures, it is obvious that there is no response outlier in training set and no response outside in test set. Only one chemical is identified as outside for MLR model for pIC_{50} PRO and two chemicals for MLR model for pIC_{50} AMA; these outsidings compounds are given as follows.

For pIC_{50} PRO. Compound number **64** in test set has higher leverage which is greater than h^* value of 0.585 and all the compounds have a standard deviation into $\pm x$ interval ($x = 3$).

For pIC_{50} AMA. Compounds numbers **5** and **6** in test set have a standard deviation value greater than the $\pm x$ interval ($x = 3$).

These results are confirmed using the simple approach (Tables S3 and S4), the chemical number **8** is identified as outlier and the chemical number **64** is identified as outside for MLR model for pIC_{50} PRO, and any chemical is identified as outside or outlier for MLR model for pIC_{50} AMA.

These erroneous predictions could probably be attributed to wrong experimental data or to the structure of these outsidings; Cl, F, or NMe_2 substitutes the tree compounds at the para position of the phenyl ring (Figure 6); maybe

FIGURE 4: Williams plot of standardized residual versus leverage for pIC_{50} (PRO) MLR model (with $h^* = 0.585$ and residual limits = ± 3).

the selected descriptors do not pay much attention to these special substructures. The predictions of tree compounds are extrapolations of the model, but fortunately they are all “good leverage” chemicals.

3.4. Artificial Neural Network (ANN). In this study, we used a feed forward network with two layers, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer (Figure 2). The architecture of the artificial neural network used in this work was 7-3-1 and $\rho = 1.82$ for pIC_{50} (PRO) and 6-3-1 and $\rho = 2$ for pIC_{50} (AMA).

The output layer represents the calculated activity values (predicted pIC_{50}). The calculation result and the performance of established models are recorded in the output layer.

To justify the predictive quality of models, total data are distributed randomly into three groups. The first group (70% of the total data) used to drive the system. The second group (15% of the total data) will be used to validate the network; and the remaining 15% that did not participate in the learning models will be used as an independent test of network generalization. The distribution of the total data into training, validation, and test sets is shown in Table 7.

The correlation between the experimental and calculated values using the artificial neural network models is highly significant, as illustrated in Figure 7 and as indicated by better R and R^2 and the small MSE values for all three phases: training, validation, and test (Table 7). The predicted activities calculated with the artificial neural network and the observed values are given in Table S5.

The results obtained by MLR and ANN are very sufficient to conclude the performance of the models. A comparison

FIGURE 5: Williams plot of standardized residual versus leverage for pIC_{50} (AMA) MLR model (with $h^* = 0.553$ and residual limits = ± 3).

of the quality of the statistical terms of these models shows that the ANN has substantially better predictive capability. ANN was able to establish a more satisfactory relationship between the molecular descriptors and the activity of the studied compounds compared to MLR; but the most negative side of this method is the fact that it is poorly transparent, whereas transparency of MLR approach gives the most interpretable results and gives a good explanation of activities with descriptors. Consequently, we can design new compounds with improved values of activity compared to the studied compounds using the MLR models. Taking into account the above results, we added suitable substitutions and then we moved to calculate their activities using the proposed models equations ((3) and (4)). Therefore, the suggested models will reduce the time and cost of synthesis as well as the determination of the antileishmanial activities against promastigotes and amastigotes forms of parasites for the acridine derivatives.

3.5. Design New Compounds with Higher Antileishmanial Activities. According to the above discussions, the MLR models could be applied to other acridines derivatives according to Table 1 and could add further knowledge in the improvement of new way in antileishmanial drug research. If we develop a new compound with better values than the existing ones, it may give rise to the development of more active compounds than those currently in use.

In this way, we carried out structural modification starting from compounds having the highest pIC_{50} values as template (number 14 and number 20). The structures of the designed compounds and their parameter values calculated

FIGURE 6: Chemical structures of the outside compounds.

FIGURE 7: Correlations between observed and predicted activities values calculated using ANN models (training set in black, validation set in red, and test set in blue).

TABLE 7: The coefficient of determination, the coefficient of correlation, and the MSE obtained by the model established by the ANN for the three phases: Training, validation, and test.

$\text{pIC}_{50} \text{ PRO}$		
Model	$n = 51$; $\text{MSE} = 0.079$; $R^2 = 0.870$; $R = 0.933$	
Training set	[1; 3; 4; 5; 6; 7; 11; 12; 13; 14; 16; 19; 21; 25; 26; 27; 28; 30; 32; 34; 35; 36; 37; 38; 39; 40; 41; 43; 44; 46; 47; 48; 49; 50; 51]	$\text{MSE} = 0.044$; $R^2 = 0.914$; $R = 0.956$
Validation set	[8; 17; 18; 22; 23; 24; 29; 31]	$\text{MSE} = 0.125$; $R^2 = 0.922$; $R = 0.960$
Test set	[2; 9; 10; 15; 20; 33; 42; 45]	$\text{MSE} = 0.186$; $R^2 = 0.725$; $R = 0.851$
$\text{pIC}_{50} \text{ AMA}$		
Model	$n = 48$; $\text{MSE} = 0.104$; $R^2 = 0.918$; $R = 0.843$	
Training set	[2; 3; 4; 5; 6; 7; 10; 11; 13; 16; 18; 19; 20; 21; 23; 27; 28; 29; 30; 31; 32; 33; 34; 35; 36; 37; 38; 39; 41; 43; 44; 45; 46; 47]	$\text{MSE} = 0.060$; $R^2 = 0.897$; $R = 0.947$
Validation set	[1; 8; 14; 15; 17; 24; 25]	$\text{MSE} = 0.198$; $R^2 = 0.710$; $R = 0.843$
Test set	[9; 12; 22; 26; 40; 42; 48]	$\text{MSE} = 0.228$; $R^2 = 0.791$; $R = 0.890$

TABLE 8: Values of descriptors, antileishmanial activity pIC₅₀, and leverages (h) for the new designed compounds.

Designed compounds	Number	Descriptors values	pIC ₅₀ (PRO)	Number	Descriptors values	pIC ₅₀ (AMA)
	X ₁	R ₁ = OH; R ₂ = CH ₃ N% = 12.490; μ = 2.232; ω = 3.099 NHA = 3; NHD = 2; MTI = 3651 TVC = 3.34 10 ⁻⁰⁵	0.571	Y ₁	R ₁ = OCH ₃ ; R ₂ = CF ₃ E _{HOMO} = -5.235; CT = 885.365; K _H = 9.302; T = 581.430; NHA = 6; NHD = 1	1.395
				Y ₂	R ₁ = OH; R ₂ = CF ₃ E _{HOMO} = -5.312; CT = 903.801; K _H = 12.056; T = 647.130; NHA = 6; NHD = 2	0.998
				Y ₃	R ₁ = OCH ₃ ; R ₂ = F E _{HOMO} = -5.034; CT = 893.897; K _H = 10.174; T = 566.560; NHA = 4; NHD = 1	0.485
				Y ₄	R ₁ = R ₂ = OC ₂ H ₅ E _{HOMO} = -4.696; CT = 927.241; K _H = 11.222; T = 622.010; NHA = 4; NHD = 1	0.464
	X ₂	R ₁ = OH; R ₂ = OH N% = 12.380; μ = 1.080; ω = 3.131 NHA = 4; NHD = 3; MTI = 3497 TVC = 1.49 10 ⁻⁰⁵	0.170	Y ₅	R ₁ = R ₂ = OCH ₃ E _{HOMO} = -4.731; CT = 916.967; K _H = 11.469; T = 599.470; NHA = 4; NHD = 1	0.442
				Y ₆	R ₁ = OCH ₃ ; R ₂ = OC(CH ₃) ₃ E _{HOMO} = -4.872; CT = 928.656; K _H = 11.099; T = 635.700; NHA = 4; NHD = 1	0.109
				Y ₇	R ₁ = OCH ₃ ; R ₂ = CH ₂ NH ₂ E _{HOMO} = -4.801; CT = 934.561; K _H = 14.185; T = 660.500; NHA = 4; NHD = 2	0.122
	X ₃	R ₁ = OH; R ₂ = meta-NH ₂ N% = 11.250; μ = 5.944; ω = 3.392 NHA = 5; NHD = 3; MTI = 17128 TVC = 2.23 10 ⁻⁰⁸	0.838	Y ₈	R ₁ = OCH ₃ ; R ₂ = para, meta -F; E _{HOMO} = -5.081; CT = 1006.097; K _H = 13.267; T = 762.370; NHA = 6; NHD = 1	0.308
	X ₄	R ₁ = OCH ₃ ; R ₂ = para-NH ₂ N% = 10.850; μ = 7.114; ω = 3.295 NHA = 5; NHD = 2; MTI = 19138 TVC = 2.04 10 ⁻⁰⁸	0.607	Y ₉	R ₁ = OCH ₃ ; R ₂ = para NH ₂ ; E _{HOMO} = -4.966; CT = 1066.417; K _H = 16.853; T = 831.93; NHA = 5; NHD = 2	0.199
	X ₅	R ₁ = OCH ₃ ; R ₂ = meta-NH ₂ N% = 10.850; μ = 6.193; ω = 3.353 NHA = 5; NHD = 2; MTI = 18992 TVC = 2.04 10 ⁻⁰⁸	0.309	Y ₁₀	R ₁ = OCH ₃ ; R ₂ = meta NH ₂ ; E _{HOMO} = -4.993; CT = 1066.417; K _H = 16.853; T = 831.930; NHA = 5; NHD = 2	0.166
	X ₆	R ₁ = OH; R ₂ = H N% = 7.820; μ = 5.631; ω = 3.483 NHA = 4; NHD = 2; MTI = 15585 TVC = 4.47 10 ⁻⁰⁸	0.381			

FIGURE 8

by the same methods as well as pIC_{50} values theoretically predicted by the MLR models are listed in Table 8.

From the predicted activities, it has been observed that the designed compounds have higher pIC_{50} values than the existing compounds in the case of the 60 studied compounds (Table 1).

The leverage values (h) for the new designed compounds are 0.268, 0.602, 0.775, 0.527, 0.386, and 0.369 for $X_1, X_2, \dots, X_6$, respectively, and 0.649, 0.797, 0.201, 0.251, 0.235, 0.166, 0.176, 0.339, 0.169, and 0.166 for $Y_1, Y_2, \dots, Y_{10}$, respectively (Table 8). Only compounds X_2, X_3, Y_1 , and Y_2 are defined as outliers and consequently they are not be considered, because they have higher leverage which is greater than h^* (0.585 for pIC_{50} (PRO) and 0.553 for pIC_{50} (AMA)); we suggest all other compounds as candidates that will be synthesized and evaluated as antileishmanial drugs.

4. Conclusion

Following the five principles of the Organisation for Economic Co-operation and Development (OECD) for the validation of QSAR models, two different modelling methods, multiple linear regression (MLR) and artificial neural network (ANN), were used in the construction of QSAR models for the antileishmanial activities against two forms of parasites (promastigotes and amastigotes) of acridine derivatives. The accuracy and predictability of the proposed models were proven by the comparison of key statistical terms of models. The good results obtained with the internal and external validations show that the proposed models in this paper are able to predict activities with a great performance and that the selected descriptors are pertinent. The applicability domain (AD) of the MLR model was defined. The resulting models have shown that we have established a relationship between some descriptors and the activities in satisfactory manners. The ANN results have substantially better predictive capability than the MLR, but the latter gives the most important interpretable results.

The obtained results show that, to increase antileishmanial activity against promastigotes parasites, we will increase electrophilicity and decrease branching, polarity, and number of hydrogen atoms attached in the heteroatom of the acridine derivatives. Moreover, to increase antileishmanial activity against amastigotes parasites, we will increase solubility, polarity, and branching and decrease electrophilicity and number of heteroatoms mostly attached in the hydrogen atoms and critical temperature of acridine derivatives. Comparing t -test and standardized coefficient values of descriptors indicates that the influences of the electrophilicity index

(ω) on pIC_{50} (PRO) and of the number of H-Bond donors (NHD) on pIC_{50} (AMA) are stronger than those of the others.

The most important finding from this research is that we have designed and proposed new compounds with higher values of activities compared to existing ones by adding suitable substituents and calculating their activity using the regression equations. Consequently, the proposed models will reduce the time and cost of synthesis as well as the determination of the antileishmanial activities against promastigotes and amastigotes forms of parasites of acridine derivatives.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors are grateful to the "Association Marocaine des Chimistes Théoriciens" (AMCT) for its pertinent help concerning the programs.

References

- [1] M. R. Moein, R. S. Pawar, S. I. Khan, B. L. Tekwani, and I. A. Khan, "Antileishmanial, antiplasmodial & cytotoxic activities of 12, 16-dideoxyaegyptinone B from *Zhumeria majdae* Rech. f. & Wendelbo," *Phytotherapy Research*, vol. 22, pp. 283–285, 2008.
- [2] L. G. Rocha, J. R. G. S. Almeida, R. O. Macêdo, and J. M. Barbosa-Filho, "A review of natural products with antileishmanial activity," *Phytomedicine*, vol. 12, no. 6-7, pp. 514–535, 2005.
- [3] O. Kayser, A. F. Kiderlen, and S. L. Croft, "Natural products as potential antiparasitic drugs," *Studies in Natural Products Chemistry*, vol. 26, pp. 779–848, 2002.
- [4] WHO—World Health Organization, 2011, <http://www.who.int/leishmaniasis/disease.epidemiology/en/index.html>.
- [5] S. M. B. Jeronimo, P. Duggal, R. F. S. Braz et al., "An emerging peri-urban pattern of infection with *Leishmania chagasi*, the protozoan causing visceral leishmaniasis in northeast Brazil," *Scandinavian Journal of Infectious Diseases*, vol. 36, no. 6-7, pp. 443–449, 2004.
- [6] M. V. L. Marlet, D. K. Sang, K. Ritmeijer, R. O. Muga, J. Onsongo, and R. N. Davidson, "Emergence or re-emergence of visceral leishmaniasis in areas of Somalia, northeastern Kenya, and south-eastern Ethiopia in 2000–2001," *Transactions of the Royal Society of Tropical Medicine & Hygiene*, vol. 97, no. 5, pp. 515–518, 2003.
- [7] J. Querido, "Emergency initiative to reduce leishmaniasis in Afghanistan," *The Lancet Infectious Diseases*, vol. 4, no. 10, p. 599, 2004.
- [8] B. L. Herwaldt, "Leishmaniasis," *The Lancet*, vol. 354, no. 9185, pp. 1191–1199, 1999.
- [9] Control of the leishmaniases, Control of the leishmaniases, 2010, Geneva, Switzerland, http://apps.who.int/iris/bitstream/10665/44412/1/WHO_TRS_949_eng.pdf.
- [10] P. Desjeux, "The increase in risk factors for Leishmaniasis worldwide," *Transactions of the Royal Society of Tropical Medicine & Hygiene*, vol. 95, no. 3, pp. 239–243, 2001.

- [11] G. Chakrabarti, A. Basu, P. P. Manna, S. B. Mahato, N. B. Mandal, and S. Bandyopadhyay, "Indolylquinoline derivatives are cytotoxic to *Leishmania donovani* promastigotes and amastigotes in vitro and are effective in treating murine visceral leishmaniasis," *Journal of Antimicrobial Chemotherapy*, vol. 43, no. 3, pp. 359–366, 1999.
- [12] S. A. Gamage, D. P. Figgitt, S. J. Wojcik et al., "Structure-activity relationships for the antileishmanial and antitrypanosomal activities of 1'-substituted 9-anilinoacridines," *Journal of Medicinal Chemistry*, vol. 40, no. 16, pp. 2634–2642, 1997.
- [13] M. O. F. Khan, S. E. Austin, C. Chan et al., "Use of an additional hydrophobic binding site, the Z site, in the rational drug design of a new class of stronger trypanothione reductase inhibitor, quaternary alkylammonium phenothiazines," *Journal of Medicinal Chemistry*, vol. 43, no. 16, pp. 3148–3156, 2000.
- [14] D. Pathak, M. Yadav, N. Siddiqui, and S. Kushawah, "Antileishmanial agents: an updated review," *Der Pharma Chemica*, vol. 3, no. 1, pp. 39–249, 2011.
- [15] L. Gupta, A. Talwar, Nishi, S. Palne, S. Gupta, and P. M. S. Chauhan, "Synthesis of marine alkaloid: 8,9-Dihydrococcinamide B and its analogues as novel class of antileishmanial agents," *Bioorganic and Medicinal Chemistry Letters*, vol. 17, no. 14, pp. 4075–4079, 2007.
- [16] W. A. Denny, "Acridine derivatives as chemotherapeutic agents," *Current Medicinal Chemistry*, vol. 9, no. 18, pp. 1655–1665, 2002.
- [17] M. Wainwright, "Acridine—a neglected antibacterial chromophore," *Journal of Antimicrobial Chemotherapy*, vol. 47, no. 1, pp. 1–13, 2001.
- [18] I. Antonini, "DNA-binding antitumor agents: From pyrimido[5,6,1-de]acridines to other intriguing classes of acridine derivatives," *Current Medicinal Chemistry*, vol. 9, no. 18, pp. 1701–1716, 2002.
- [19] M. Demeunynck, F. Charmantray, and A. Martelli, "Interest of acridine derivatives in the anticancer chemotherapy," *Current Pharmaceutical Design*, vol. 7, no. 17, pp. 1703–1724, 2001.
- [20] S. A. Gamage, D. P. Figgitt, S. J. Wojcik et al., "Structure-activity relationships for the antileishmanial and antitrypanosomal activities of 1'-substituted 9-anilinoacridines," *Journal of Medicinal Chemistry*, vol. 40, no. 16, pp. 2634–2642, 1997.
- [21] C. M. Mesa-Valle, J. Castilla-Calvente, M. Sanchez-Moreno, V. Moraleda-Lindez, J. Barbe, and A. Osuna, "Activity and mode of action of acridine compounds against *Leishmania donovani*," *Antimicrobial Agents and Chemotherapy*, vol. 40, no. 3, pp. 684–690, 1996.
- [22] C. Di Giorgio, K. Shimi, G. Boyer, F. Delmas, and J.-P. Galy, "Synthesis and antileishmanial activity of 6-mono-substituted and 3,6-di-substituted acridines obtained by acylation of proflavine," *European Journal of Medicinal Chemistry*, vol. 42, no. 10, pp. 1277–1284, 2007.
- [23] C. Di Giorgio, F. Delmas, N. Filloux et al., "In vitro activities of 7-substituted 9-chloro and 9-amino-2-methoxyacridines and their bis- and tetra-acridine complexes against *Leishmania infantum*," *Antimicrobial Agents and Chemotherapy*, vol. 47, no. 1, pp. 174–180, 2003.
- [24] G. Carole, D. M. Michel, C. Julien et al., "Synthesis and antileishmanial activities of 4,5-di-substituted acridines as compared to their 4-mono-substituted homologues," *Bioorganic & Medicinal Chemistry*, vol. 13, no. 19, pp. 5560–5568, 2005.
- [25] Adamo and Baron (2000), Parac & Grimme (2003) Gaussian 03.
- [26] ACDLABS 10, "Advanced Chemistry Development, Inc., Toronto, ON, Canada," 2015, <http://www.acdlabs.com/>.
- [27] MarvinSketch 5.11.4, "Chem Axon," 2012, <http://www.chemaxon.com>.
- [28] ChemBioOffice, *PerkinElmer Informatics*, 2010, <http://www.cambridgesoft.com>.
- [29] J. C. Dearden, M. T. D. Cronin, and K. L. E. Kaiser, "How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR)," *SAR and QSAR in Environmental Research*, vol. 20, no. 3–4, pp. 241–266, 2009.
- [30] K. Roy, R. N. Das, P. Ambure, and R. B. Aher, "Be aware of error measures. Further studies on validation of predictive QSAR models," *Chemometrics and Intelligent Laboratory Systems*, vol. 152, pp. 18–33, 2016.
- [31] OECD, *Guidance Document on the Validation of QSAR Models*, Organization for Economic Co-Operation & Development, Paris, France, 2007.
- [32] L. Eriksson, J. Jaworska, A. P. Worth, M. T. D. Cronin, R. M. McDowell, and P. Gramatica, "Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs," *Environmental Health Perspectives*, vol. 111, no. 10, pp. 1361–1375, 2003.
- [33] P. Gramatica, "Principles of QSAR models validation: internal and external," *QSAR and Combinatorial Science*, vol. 26, no. 5, pp. 694–701, 2007.
- [34] G. E. Batista and D. F. Silva, "How k-nearest neighbor parameters affect its performance," in *Proceedings of the Argentine Symposium on Artificial Intelligence*, pp. 1–12, Instituto de Ciencias Matemáticas de Computação, São Carlos, Brazil, 2009.
- [35] T. I. Netzeva, A. P. Worth, T. Aldenberg et al., "Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships," *Alternatives to Laboratory Animals*, vol. 33, no. 2, pp. 155–173, 2005.
- [36] J. C. Dearden, "The history and development of quantitative structure-activity relationships (QSARs)," *International Journal of Quantitative Structure-Property Relationships*, vol. 1, no. 1, pp. 1–44, 2016.
- [37] K. Roy, S. Kar, and P. Ambure, "On a simple approach for determining applicability domain of QSAR models," *Chemometrics and Intelligent Laboratory Systems*, vol. 145, pp. 22–29, 2015.
- [38] K. Roy, S. Kar, and R. Das Narayan, "Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences & Risk Assessment," in *Chapter 6-Selected Statistical Methods in QSAR*, pp. 191–229, Academic Press, Boston, Mass, USA, 2015.
- [39] S.-S. So and W. Graham Richards, "Application of neural networks: quantitative structure-activity relationships of the derivatives of 2,4-diamino-5-(substituted-benzyl)pyrimidines as DHFR inhibitors," *Journal of Medicinal Chemistry*, vol. 35, no. 17, pp. 3201–3207, 1992.
- [40] T. A. Andrea and H. Kalayeh, "Applications of neural networks in quantitative structure-activity relationships of dihydrofolate reductase inhibitors," *Journal of Medicinal Chemistry*, vol. 34, no. 9, pp. 2824–2836, 1991.
- [41] A. Golbraikh and A. Tropsha, "Beware of q^2 !", *Journal of Molecular Graphics and Modelling*, vol. 20, no. 4, pp. 269–276, 2002.
- [42] S. Chtita, R. Hmamouchi, M. Larif, M. Ghamali, M. Bouachrine, and T. Lakhli, "QSPR studies of 9-anilinoacridine derivatives for their DNA drug binding properties based on density functional theory using statistical methods: model, validation and influencing factors," *Journal of Taibah University for Science*, vol. 10, no. 6, pp. 868–876, 2016.

- [43] S. Chtita, M. Larif, M. Ghamali, M. Bouachrine, and T. Lakhlifi, "Quantitative structure–activity relationship studies of dibenzo[a,d]cycloalkenimine derivatives for non-competitive antagonists of N-methyl-d-aspartate based on density functional theory with electronic and topological descriptors," *Journal of Taibah University for Science*, vol. 9, no. 2, pp. 143–154, 2015.
- [44] R. G. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*, John Wiley & Sons, Chichester, UK, 2003.
- [45] W. D. Fisher, "On grouping for maximum homogeneity," *Journal of the American Statistical Association*, vol. 53, no. 284, pp. 789–798, 1958.
- [46] A. N. Choudhary, A. Kumar, and V. Juyal, "Quantitative structure activity relationship (QSAR) analysis of substituted 4-oxothiazolidines and 5-arylidines as lipoxygenase inhibitors," *Mini-Reviews in Medicinal Chemistry*, vol. 10, no. 8, pp. 705–714, 2010.