

Research Article

Virtual Screening Based on QSAR and Molecular Docking of Possible Inhibitors Targeting Chagas CYP51

J. Cotuá ¹, H. Llinás ², and S. Cotes ³

¹Grupo de Investigación Max Planck, Facultad de Química y Farmacia, Universidad del Atlántico, Barranquilla, Colombia

²Departamento de Matemáticas y Estadística, División de Ciencias Básicas, Universidad del Norte, Barranquilla, Colombia

³Departamento de Química y Biología, División de Ciencias Básicas, Universidad del Norte, Barranquilla, Colombia

Correspondence should be addressed to S. Cotes; scotes@uninorte.edu.co

Received 6 October 2020; Revised 2 February 2021; Accepted 12 February 2021; Published 28 February 2021

Academic Editor: Ester Chiessi

Copyright © 2021 J. Cotuá 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.

Chagas is a neglected tropical disease caused by the parasite *Trypanosoma cruzi* with no effective treatment in all its forms. There is a need to find more effective therapeutic alternatives with reduced toxicity. In this contribution, multiple linear regression models were used to identify the molecular descriptors that best describe the inhibitory activity of 52 fenarimol analogues against *Trypanosoma cruzi*. The topological, physicochemical, thermodynamic, electronic, and charge descriptors were evaluated to cover a wide range of properties that frequently encode biological activity. A model with high predictive value was obtained based on geometrical descriptors and descriptors encoding hydrophobicity and London dispersion forces as necessary for the inhibition of *Trypanosoma cruzi*-CYP51. Docking methodology was implemented to evaluate molecular interactions *in silico*. The virtual screening results in this study can be used for rational design of new analogues with improved activity against Chagas disease.

1. Introduction

Chagas, a tropical disease endemic to Latin America, is caused by the parasite *Trypanosoma cruzi* (*T. cruzi*), first described in 1909 [1–3]. The most affected population is very low-income people, who do not have the economic resources to pay for expensive treatments and live in conditions of high vulnerability [4–6]. The pharmacological therapy currently used is aimed at preventing the chronicity of the infection since the efficacy of antiparasitic drugs decreases in the chronic phases of the disease. The effectiveness of the treatment in the acute phase has been demonstrated [7], while the chronic stage can take up to 20 years resulting in multiple-organ damage [6–8]. Studies show that treatment effectiveness in the acute phase is 50–70%, while in the chronic phase, the effectiveness is reduced to 8–30% [9, 10]. Additionally, treatment depends on age of the patient and long-term side effects.

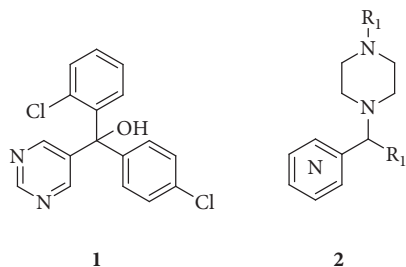
There is no effective treatment for Chagas disease in all its forms; treatment with nifurtimox and benznidazole is highly toxic [6, 7, 11]. Both compounds are heterocyclic with

a furan and imidazole-nitrogenated ring, respectively, which act by inducing cytotoxicity in the parasite by interacting with the nitroreductase enzymes of the parasites [12–15]. Mutagenic and carcinogenic activities, hepatotoxicity, and bone marrow suppression have been shown to be significant with the use of these medications. Buschini et al. [16] reassessed the genotoxic and mutagenic activity for nifurtimox and benznidazole, respectively. They found that both drugs damage the infectious agent and DNA of host cells at concentrations in the range of plasma concentrations of patients treated by chemotherapy. Benznidazole and nifurtimox contain nitro groups that produce nitrogenated metabolic radicals that affect the parasite and are responsible for the toxicity, mutagenicity, genotoxicity, and carcinogenicity attributed to them [17]. Owing to the high toxicity, few therapeutic alternatives go to clinical trials [18]; a greater selectivity of the drug and reduced adverse effects on the host are required. Thus, there is a need to develop therapeutic alternatives with improved efficacy and reduced toxicity [18–20]. Sterol 14 α -demethylase cytochrome P450 (CYP51) inhibitors with trypanocidal activity have been identified

since 1990 [21]. CYP51 is an important target in the treatment of Chagas disease because the inhibition of sterol synthesis is lethal to the parasite. Urbina et al. reported the use of triazoles as inhibitory compounds of CYP51 in *T. cruzi* in 2002 [22–25]. The binding site of the ligand in CYP51 contains a hydrophilic region, a hydrophobic region, and a channel formed by hydrophobic residues between the β -helix and the N-terminal helix [26]. Azoles, such as D0870, posaconazole, and ravuconazole, act through the coordination of nitrogen with heme iron of the CYP51 enzyme. The crystallographic structure of *T. cruzi*-CYP51 linked to posaconazole and fluconazole (Figure 1) has been reported [27] which favors the design of therapeutic alternatives based on methods such as structure-based drug design. The active site on CYP51 takes the form of Y, one end being longer than the other two. The longer end is a pocket at an angle to the heme group plane, which can interact with up to four fused aromatic rings in this highly hydrophobic pocket.

Other types of substances, such as naphthoquinones, diamines, nitroimidazoles, and their derivatives have been evaluated, albeit with high toxicity. In addition to CYP51 inhibitors, other therapeutic targets for the development of new drugs include inhibitors of cruzipain, inhibitors of pyrophosphate enzymes, HGPRT (hypoxanthine-guanine phosphoribosyltransferase), and trypanothione reductase, the principal enzymes of *T. cruzi* metabolism [23, 28].

In the last 25 years, significant advances have been made in the research of *T. cruzi* and the complete elucidation of its genome [30, 31]; however, the role of many of the encoded proteins is unknown. In addition, significant advances have not been reported in the field of natural product research owing to the lack of knowledge about human toxicity [17]. Thus, the only treatment available for Chagas disease is the use of nifurtimox and benznidazole. In two decades of research, it has not been possible to find more effective therapeutic alternatives with reduced toxicity, and the therapeutic combinations of the drugs already available are under clinical study. Cheminformatics and molecular modeling approaches represent an important initiative considering the high costs of traditional research [32, 33].



Fenarimol 1 is a fungicide discovered with moderate activity *in vitro* (IC_{50} , 350 nM) against *T. cruzi* [34]. Structure-activity relationship studies to improve the inhibitory capacity of CYP51 led to pharmacophore 2, in which an aromatic ring is replaced by piperazine with substitutions at the N-4 atom.

The new derivatives with amide, sulfonamide, aromatic, carbamate, and carbonate substituents were evaluated for

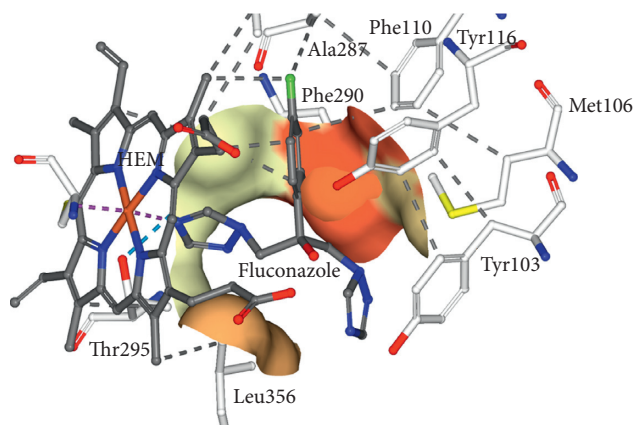


FIGURE 1: Fluconazole bound to *T. cruzi*-CYP51 [29].

their ability to inhibit *T. cruzi* [35] *in vitro* and showed very promising results. Additionally, Keenan et al. [35] demonstrated that piperazine analogues of fenarimol do not exhibit cytotoxicity. These findings make the piperazine analogues of fenarimol excellent candidates for more advanced optimization studies.

Quantitative structure-activity relationship (QSAR) is an approach that uses statistical models to correlate the molecular properties encoded in molecular descriptors with the biological activity of a group of molecules that share a pharmacophore. To the best of our knowledge, there is a lack of QSAR studies of piperazine analogues of fenarimol reported.

Several regression methods are available for establishing the correlation between the activity and the molecular structure descriptors [33]. Linear methods are efficient in relating the structure to a given biological activity, but the accuracy of nonlinear methods is considered superior [33]. However, nonlinear methods are harder to interpret and suffer from overfitting when the number of descriptors is greater than the number of samples in the dataset [36]. Consequently, we used the multiple linear regression (MLR) model because of its reliability, accuracy, and ease of interpretation, which is supported by an extensive number of studies conducted in the field of medicinal chemistry. MLR is one of the most widely applicable methods to the conduct of QSAR studies [33, 37–39].

A large number of descriptors of molecular structures are available; their selection is one of the most important initial tasks of QSAR studies. Descriptors based on topological indexes contain information related to molecular shape, size, flexibility, and the degree of branching. Their correlation with biological activity has been widely demonstrated. The physicochemical descriptors encode properties of hydrophobicity and electronic and steric effects, properties that are attributed to different forms of chemical interaction with biological targets, because of which they are most frequently used for QSAR studies. The partial charge descriptors contain chemical information resulting from the electronic distribution and molecular geometry combined in the same descriptor. The charge descriptors allow prediction of electronic density regions favorable for interaction with

biological targets. A total of 53 descriptors were selected as widely as possible to evaluate the structural diversity of the analogues under study.

Thermodynamic and electronic molecular descriptors have been calculated with Gaussian 09 [40]; other descriptors including topostructural, topochemical, and topological indexes; physicochemical descriptors; and charge descriptors were calculated with MOE [41] to determine the structural factors with greater relevance as determinants of the biological activity against *T. cruzi* in the series of piperazine analogues of fenarimol used in this article. Additionally, this study evaluated piperazine analogues of fenarimol's interaction with the CYP51 receptor through docking methodology.

2. Materials and Methods

2.1. Descriptor Selection and Model Generation. Molecular geometries of the compounds were optimized using the MM2 molecular mechanics method; the generated structures were further optimized using DFT employing B3LYP and 6-311+G(d,p) basis sets. The results were used for descriptor calculations. DFT calculations with B3LYP functional have been successful in the prediction of several molecular properties [42]. No imaginary frequencies were found. Therefore, the calculated geometries were local minima in a potential energy surface. Molecular-geometry optimizations were performed using Gaussian 09 Revision-A.02-SMP software [40]. For all calculations and comparisons in this study, only the most stable conformations were considered. Several types of molecular descriptors were calculated using MOE [41]. Microsoft Excel Data Analysis add-in and several validation analyses were used to evaluate the statistical significance of the employed models.

A molecular descriptor is a numerical representation of a molecular property derived from a standardized experiment or mathematical procedure. Any molecular property may be used, and descriptors can be determined experimentally from physical properties or computed from structural features such as van der Waals forces, atomic and bond counts, distances, pharmacophoric features, partial charges, volume, and shape. To reduce the number of descriptors, the correlation coefficients between each descriptor and pIC₅₀ were calculated. The selected descriptors for the creation of the model have a coefficient of determination (R^2) ≥ 0.5 [43]; in total, 12 descriptors were selected, which are defined in Table 1. The values for the selected descriptors are found in Table 2.

The descriptors of Table 2 were selected using the successive step mode [44], and the following parameters were taken into account: use of 1 descriptor for every 5 compounds [45] and selection of descriptors of different nature to avoid collinearity between the variables [46]. Subsequently, multiple linear regression (MLR) was used to study the linear relationship between pIC₅₀ and the remaining descriptors, but only those models with R^2 higher than 0.568 were considered valid [47]. The linear relationship between pIC₅₀ and descriptors was determined using the standard equation:

$$Y = BX + \epsilon, \quad (1)$$

where Y is an $nx1$ vector of values of the dependent variable pIC₅₀, X is an nxk matrix of explanatory variables (known as molecular descriptors), B is a $kx1$ vector of the estimated parameters, and ϵ is an $nx1$ residual vector whose components are assumed to be independent, normal, random variables with mean zero and known variance σ^2 .

Stepwise regression fitting was performed with the successive addition of descriptors that significantly improved the fit; this process was repeated until no further improvement in the model was possible. Next, the descriptors whose loss gives the most statistically insignificant deterioration of the model fit were removed. This process was repeated until no further descriptors could be deleted without a statistically significant loss of fit. Using these conditions, 21 models were obtained; these models can be accessed from supplementary Table S1 with their corresponding statistics in supplementary Table S2. The best predictive model with the lowest number of descriptors, a high determination coefficient (R^2), and no collinearity among the selected descriptors was selected for further improvement.

The "leave-one-out" (LOO) cross-validation scheme was used to assess the validity of the models. During the validation process, the model function is trained on data from all molecules excluding one. pIC₅₀ for the molecule that was left out was then predicted using the descriptors for the model. Statistical parameters such as Q^2 (cross-validated correlation coefficient), R^2 (determination coefficient), standard deviation (SD), and SE (standard error) were taken into account to evaluate the quality of the proposed QSAR models:

$$Q^2 = 1 - \frac{\sum_{i=1}^{\text{training}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{\text{training}} (y_i - \bar{y}_i)^2},$$

$$R^2 = \frac{\sigma_{XY}^2}{\sigma_x^2 \sigma_y^2}, \quad (2)$$

$$SD = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}},$$

$$SE = \frac{SD}{\sqrt{n}}$$

where y_i , \hat{y}_i , and \bar{y}_i are the measured, predicted, and averaged (over the whole dataset) values of the dependent variable, respectively; σ_{XY}^2 is the covariance of X and Y ; σ_x^2 is the variance of X ; σ_y^2 is the variance of Y ; and n is the number of samples used for model building.

2.2. Enzyme Collection and Docking Preparation. The structure of the receptor was obtained from the Protein Data Bank (<https://www.rcsb.org/>) database, identified in the repository as a crystal structure of sterol 14 α -demethylase (CYP51) from *Trypanosoma cruzi* in complex with the inhibitor fluconazole. PDB ID: 3KHM. The initial crystal

TABLE 1: Symbols, types, and definitions of selected descriptors.

Code	Descriptor type	Definition
Length	Geometry	Length to reach the binding pocket
Pmi	Shape	Principal moment of inertia
Std_dim1	Shape	Standard dimension 1
Vsa	Surface area	van der Waals surface area
Vol	Volume	van der Waals volume
Weight	Physical properties	Molecular weight
apol	Physical properties	Sum of the atomic polarizabilities
SMR	Physical properties	Molecular refractivity
Log P(o/w)	Physical properties	Log of the octanol/water partition coefficient
ASA	Surface area	Water accessible surface area
Q_VSA_FPPOS	Physical properties	Partial charge and molecular surface area
Q_VSA_HYD	Physical properties	Partial charge and molecular surface area

structure was preprocessed by adding hydrogens and deleting water molecules and fluconazole. The docking simulations were performed using the AutoDock Vina code (version 1.1.2) employing 3-way multithreading, Lamarckian genetic algorithm [48] for flexible-ligand docking in a docking box center_x = 5.514, center_y = -26.332, center_z = 21.612, size_x = 20.883, size_y = 20.600, and size_z = 22.377, and exhaustiveness = 8. The most favorable interactions are indicated by the lowest free-bond energy (ΔG). Figures were drafted using the Discovery Studio Visualizer [49] and UCSF Chimera viewers [50].

3. Results

Pharmacological data of 52 fenarimol analogues (molecules 3–54 in Table 3) were taken from the literature [35]. pIC_{50} ($-\log\text{IC}_{50}$) was used as a measure of biological activity, IC_{50} values corresponding to concentration of the compound required to inhibit 50% of *T. cruzi* measured under the same experimental conditions.

Figure 2 shows the pIC_{50} distribution of the data, spanning three orders of magnitude. We utilized the principal component regression (PCR) and partial least square regression (PLS) for examining the linear relationship between IC_{50} and the corresponding descriptors. However, MLR afforded the best predictive model with the lowest number of predictive variables and a high determination coefficient (R^2) of the most significant variables.

Model 01 was obtained using MLR with 57% of the variation in biological activity explained by the descriptors and standard error (SE) 0.374 (Table 4).

Model 01:

$$\begin{aligned} \text{pIC}_{50} = & 0.160 + 0.150(\text{Length}) + 0.186(\text{LogP}(o/w)) \\ & - 2.809(\text{Q_VSA_FPPOS}). \end{aligned} \quad (3)$$

The jackknife technique was implemented to eliminate the outliers [51] to improve the statistical quality of model 01. After removing molecules 10, 11, 24, 25, 27, 45, 52, 53, and 54 as outliers, model 02 was achieved. Occurrence of outliers depends on three main factors [46]: errors in the

reported biological activity or calculated descriptors' values; different mechanisms of action for the dataset used to build the model; and sampling design errors. When applied to the information used to build our model, none of the above three factors were able to sufficiently explain why molecules 10, 11, 24, 25, 27, 45, 52, 53, and 54 were classified as outliers.

Model 02:

$$\begin{aligned} \text{pIC}_{50} = & -0.168 + 0.168(\text{Length}) + 0.269(\text{LogP}(o/w)) \\ & - 3.318(\text{Q_VSA_FPPOS}). \end{aligned} \quad (4)$$

The coefficient of determination R^2 increases from 0.568 in model 01 to 0.817 in model 2, and the SE decreases from 0.374 to 0.256 (Table 5). The improved statistical quality of model 02 is reflected since 82% of the variation in biological activity is explained by the selected descriptors. Table 6 shows no collinearity in the selected descriptors, and therefore, the resulting model 02 has good stability.

From the LOO cross-validation procedure in Table 7, the square of the cross-validation coefficient LOO (Q^2) is obtained to evaluate the robustness and the predictive capacity of the model. Complete LOO cross-validation statistics for model 02 are available in supplementary Table S3. Good internal prediction power was achieved because the correlation coefficient of cross-validation LOO (Q^2) was greater than 0.5 [47]. Cross-validated pIC_{50} predicted values are shown in Table 8.

On the contrary, the R^2 value of the original model should not be significantly greater than the Q^2 value, and the difference between R^2 and Q^2 should not be more than 0.3 [52]. In our case that the difference is 0.040 ($R^2 = 0.820$; $Q^2 = 0.780$) indicates that the model is not overfitted, and $R^2 = 0.820$ indicates a good correlation between the observed and predicted pIC_{50} (Table 7).

Model 02 was used to calculate pIC_{50} of all fenarimol analogues (Table 8).

With these new values, the observed pIC_{50} vs. predicted pIC_{50} graph is shown in Figure 3. Figure 3 demonstrates the reproducibility of the predicted data with respect to that observed in the pIC_{50} data taken to develop the model. The results showed a Pearson correlation coefficient of 0.904,

TABLE 2: Calculated values of selected descriptors.

Molecule	Length (Å)	Pmi	std_dim1	VSA (Å ²)	Vol (Å ³)	Weight (amu)	apol	SMR	Log P(o/w)
ASA (Å ²)								Q_VSA_PPPOS	Q_VSA_HYD
3	7160.268	407.902	379.266	60.469	10.521	3.710	636.217	0.170	317.092
4	8701.990	424.794	392.766	61.720	10.532	4.053	657.953	0.186	285.199
5	6469.489	380.648	348.891	56.042	9.848	2.860	611.650	0.197	261.190
6	9048.030	428.252	399.094	65.322	11.233	4.186	673.269	0.172	312.887
7	8180.707	420.234	390.234	61.830	10.536	3.865	660.841	0.188	280.790
8	7816.521	439.055	404.156	66.055	11.139	3.333	694.238	0.175	327.115
9	9231.642	422.431	388.969	61.135	11.193	3.077	670.930	0.193	298.961
10	7522.210	410.565	380.109	61.431	11.156	3.983	652.137	0.176	302.862
11	7810.817	415.805	391.500	61.322	11.152	4.171	647.558	0.174	307.271
12	7345.892	370.828	390.890	59.142	10.300	2.565	630.753	0.223	288.317
13	6601.999	379.482	345.516	54.715	9.627	1.715	607.200	0.259	232.415
14	5188.416	354.749	324.000	51.298	9.147	2.321	569.209	0.204	249.595
15	6802.074	408.841	377.156	61.259	10.496	2.590	640.178	0.219	280.790
16	6134.051	385.154	352.688	56.831	9.823	1.740	614.968	0.254	224.888
17	8954.287	421.386	391.078	61.925	11.168	1.957	673.543	0.243	262.659
18	9347.768	429.150	400.781	63.264	11.438	3.421	675.089	0.160	309.247
19	8802.489	425.596	391.078	63.373	11.442	3.233	669.767	0.162	304.838
20	9322.357	445.577	407.953	67.374	11.658	3.558	688.026	0.160	309.108
21	10037.038	469.272	429.047	71.141	12.350	3.784	722.632	0.149	340.298
22	8489.916	482.736	449.297	71.218	12.626	3.419	683.786	0.188	333.513
23	11411.812	445.761	407.109	64.887	11.943	3.825	690.003	0.155	322.419
24	8005.313	377.495	345.938	53.054	9.350	0.965	596.130	0.260	195.312
25	7569.777	384.162	351.000	54.381	9.570	2.110	613.082	0.202	224.088
26	5549.176	399.571	369.563	59.427	10.485	2.726	573.457	0.169	286.036
27	6392.631	370.568	337.922	53.130	9.559	1.767	584.210	0.184	255.981
28	7496.365	378.685	338.344	54.491	9.574	1.922	610.296	0.204	219.678
29	7010.333	404.289	359.859	56.087	10.260	2.673	597.120	0.185	251.572
30	5962.690	371.756	335.813	53.920	9.534	0.647	597.193	0.239	219.678
31	7556.569	374.229	338.344	53.164	9.354	0.777	599.876	0.263	190.903
32	6242.631	358.532	326.531	51.913	9.343	0.434	572.658	0.247	222.796
33	8774.911	423.467	389.813	62.124	11.321	4.449	685.639	0.183	313.521
34	9607.058	425.205	395.719	62.422	11.643	3.850	677.247	0.179	311.762
35	14085.754	476.652	442.547	70.376	12.800	3.594	751.421	0.205	336.028
36	8451.066	424.727	394.031	62.900	11.351	4.621	668.822	0.208	309.112
37	9429.794	415.121	384.750	62.531	11.647	3.662	674.789	0.181	307.353
38	11750.636	444.547	408.375	63.275	11.927	4.871	704.128	0.244	300.568
39	9900.265	431.756	394.031	62.632	10.713	4.160	679.777	0.199	280.790
40	8665.371	425.438	387.703	61.271	10.698	4.005	677.022	0.182	317.092
41	12822.549	430.800	399.516	62.718	11.906	3.188	694.422	0.225	296.159
42	8578.309	419.047	378.844	59.944	10.478	2.860	665.675	0.234	288.317
43	8227.207	417.924	382.641	60.909	10.795	3.051	663.322	0.251	288.317
44	6473.907	384.750	353.531	56.158	9.987	2.162	611.725	0.313	243.644

TABLE 2: Continued.

	Molecule ASA (\AA^2)	Length (\AA)	Pmi	std_dim1	VSA (\AA^2)	Vol (\AA^3)	Weight (amu)	apol	SMR	Log P(o/w)
									Q_VSA_FPPOS	Q_VSA_HYD
45	7.600	7090.789	378.032	407.371	53.306	9.579	2.680	588.064	0.286	200.735
46	7.658	7073.802	395.085	381.882	58.760	10.783	4.590	625.417	0.150	311.762
47	7.938	7865.937	418.056	393.918	62.765	11.442	4.356	663.144	0.202	312.087
48	8.228	7550.978	385.138	383.858	56.106	10.342	2.581	629.003	0.194	281.004
49	7.939	9375.027	472.376	417.553	71.335	12.518	4.523	737.190	0.191	353.797
50	7.701	7737.324	443.129	379.935	63.297	11.460	4.402	675.272	0.135	353.208
51	8.060	7050.790	382.503	361.473	56.269	10.325	2.451	622.075	0.246	245.427
52	7.916	7835.126	414.886	412.896	61.328	11.218	3.399	661.976	0.253	287.721
53	7.595	7282.442	389.167	382.870	57.433	10.562	3.445	618.053	0.208	282.987
54	7.920	8965.375	430.363	428.458	62.689	11.233	3.623	683.755	0.218	272.656

TABLE 3: Structural features of selected fenarimol analogues.

Molecule	HET ^a	R ₂	R ₃	<i>T. cruzi</i> IC ₅₀ (μM)
<div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;"> <p>R₁=</p> <p>X = O, NH</p> </div> <div style="text-align: center;"> <p>R₂=</p> <p>A X = H B X = F</p> <p>C X = H D X = F</p> <p>E</p> <p>F</p> </div> </div>				
R ₁ = alkyl and aryl amides				
3	3-Pyr	B	tBu	0.019
4	3-Pyr	D	tBu	0.024
5	3-Pyr	B	cPr	0.026
6	3-Pyr	B	cHex	0.028
7	3-Pyr	C	tBu	0.030
8	3-Pyr	F	iPr	0.035
9	3-Pyr	B	Het ^b	0.036
10	3-Pyr	A	Ph	0.068
11	3-Pyr	B	Ph	0.070
12	5-Pyrim	B	tBu	0.081
13	5-Pyrim	B	cPr	0.171
14	3-Pyr	A	Me	0.240
15	3-Pyr	E	tBu	0.254
16	3-Pyr	E	cPr	0.281
17	3-Pyr	E	Het ^b	0.513
R ₁ = alkyl and aryl sulfonamides				
18	3-Pyr	B	Ph	0.008
19	3-Pyr	A	Ph	0.009
20	3-Pyr	A	cHex	0.010
21	3-Pyr	A	Ph(4-iPr)	0.012
22	5-Pyrim	B	Ph(4-iPr)	0.018
23	3-Pyr	A	Ph(4-Cl)	0.024
24	5-Pyrim	D	Me	0.025
25	3-Pyr	D	Me	0.028
26	3-Pyr	A	iPr	0.028
27	3-Pyr	B	Me	0.035
28	3-Pyr	C	Me	0.077
29	3-Pyr	A	Me	0.110
30	3-Pyr	E	Me	0.365
31	5-Pyrim	C	Me	0.769
32	5-Pyrim	A	Me	0.880
R ₁ = alkyl and aryl carbonates/carbamates				
33	3-Pyr	B	OPh	0.005
34	5-Pyr	B	NHPh	0.006
35	5-Pyrim	B	NHPh(4-iPr)	0.007
36	3-Pyr	A	OPh	0.008
37	3-Pyr	A	NHPh	0.009
38	5-Pyrim	B	NHPh(4-Cl)	0.010
39	3-Pyr	C	OtBu	0.017
40	3-Pyr	B	OtBu	0.020
41	5-Pyrim	A	NHPh(4-Cl)	0.022
42	5-Pyrim	B	OtBu	0.024
43	5-Pyrim	B	NHtBu	0.054
44	3-Pyr	A	NMe ₂	0.120
R ₁ = Ar				
45	3-Pyr	D	Het ^b	0.017
46	3-Pyr	B	Ph	0.018
47	3-Pyr	A	Ph(2-OMe)	0.020
48	3-Pyr	B	2-Pyrimidine	0.026
49	3-Pyr	F	Ph(2-OMe)	0.028
50	3-Pyr	A	Ph	0.034
51	3-Pyr	E	2-Thiazole	0.050
52	5-Pyrim	B	Ph(2-OMe)	0.065
53	5-Pyrim	B	Ph	0.078
54	2-Pyra	C	Ph(2-OMe)	0.240

^aHET: 3-pyr (3-pyridine); 5-pyrim (5-pyrimidine). ^bHet = 2-(4-methyl-1,3-thiazole).

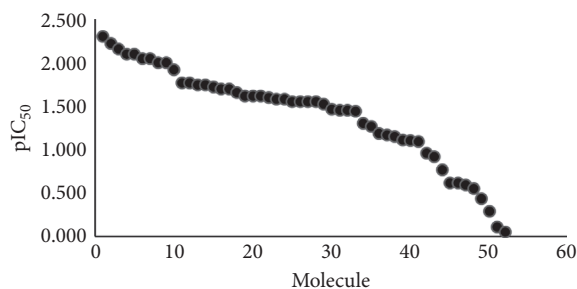


FIGURE 2: Distribution of pIC_{50} values showing a wide range of the inhibitory potency of the molecules examined in this study.

TABLE 4: Statistics for regression model 01.

Multiple correlation coefficient	0.754
Coefficient R^2	0.568
Adjusted R^2	0.541
SE	0.374

TABLE 5: Statistics for regression model 02.

Multiple correlation coefficient	0.904
Coefficient R^2	0.817
Adjusted R^2	0.803
SE	0.256

TABLE 6: Correlation matrix of descriptors.

	Log P(o/w)	Q_VSA_FPPOS	Length
Log P(o/w)	1		
Q_VSA_FPPOS	-0.589	1	
Length	0.509	-0.255	1

TABLE 7: LOO cross-validation statistics for model 02.

Q^2	0.780
Multiple correlation coefficient	0.900
Coefficient R^2	0.820
Adjusted R^2	0.810
SE	0.250

which demonstrates a good linear correlation between observed pIC_{50} and predicted pIC_{50} , and an R^2 correlation coefficient of 0.817 (Table 9). In accordance with the statistical parameters discussed, model 02 meets the statistical quality to predict the activity of new piperazine analogues of fenarimol.

4. Discussion

4.1. Model Interpretation. A QSAR model with statistical validity was obtained. The coefficients of the descriptors in a QSAR model can be used to evaluate the contribution of each descriptor toward biological activity; this information can be utilized to develop effective strategies in designing compounds with improved specificity and biological activity.

Figure 4 shows the estimates of the regression coefficients obtained from model 02. There are descriptors that

TABLE 8: Observed and predicted values of pIC_{50} .

Molecule	pIC_{50}	Predicted pIC_{50}	Residual	pIC_{50} cross-validated
3	1.721	1.542	0.179	1.53
4	1.620	1.637	-0.018	1.64
5	1.585	1.283	0.302	1.27
6	1.553	1.941	-0.388	1.96
7	1.523	1.528	-0.005	1.53
8	1.456	1.436	0.020	1.44
9	1.444	1.464	-0.020	1.46
12	1.092	1.113	-0.021	1.11
13	0.767	0.707	0.060	0.70
14	0.620	0.825	-0.205	0.84
15	0.595	1.129	-0.534	1.15
16	0.551	0.788	-0.237	0.81
17	0.290	0.940	-0.650	0.98
18	2.097	1.694	0.403	1.57
19	2.046	1.600	0.446	1.76
20	2.000	1.784	0.216	1.77
21	1.921	2.181	-0.260	2.23
22	1.745	1.472	0.272	1.46
23	1.620	1.846	-0.227	1.86
26	1.553	1.389	0.164	1.38
28	1.114	0.798	0.315	0.77
29	0.959	1.063	-0.104	1.07
30	0.438	0.337	0.100	0.32
31	0.114	0.293	-0.179	0.32
32	0.056	0.254	-0.198	0.30
33	2.301	2.152	0.149	2.14
34	2.222	2.058	0.163	2.04
35	2.155	2.293	-0.138	2.37
36	2.097	2.120	-0.023	2.12
37	2.046	1.997	0.049	1.99
38	2.000	1.926	0.074	1.90
39	1.770	1.719	0.051	1.72
40	1.699	1.734	-0.035	1.74
41	1.658	1.721	-0.063	1.73
42	1.620	1.253	0.367	1.24
43	1.268	1.265	0.003	1.26
44	0.921	0.583	0.338	0.47
46	1.745	1.856	-0.111	1.87
47	1.699	1.668	0.031	1.64
48	1.585	1.265	0.320	1.25
49	1.553	1.750	-0.197	1.77
50	1.469	1.862	-0.394	1.92
51	1.301	1.030	0.271	1.02

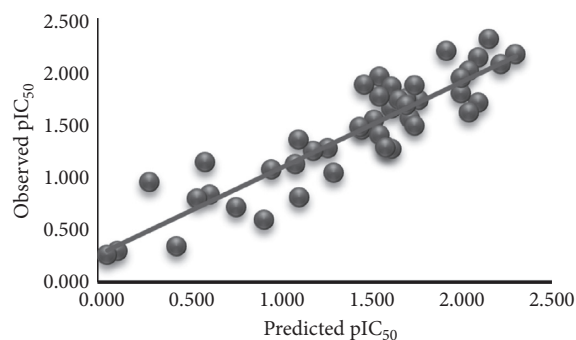


FIGURE 3: Correlation of observed vs. predicted pIC_{50} values using model 02.

TABLE 9: Statistics for the regression model between observed pIC_{50} and predicted pIC_{50} .

Multiple correlation coefficient	0.904
Coefficient R^2	0.817
Adjusted R^2	0.803
SE	0.256

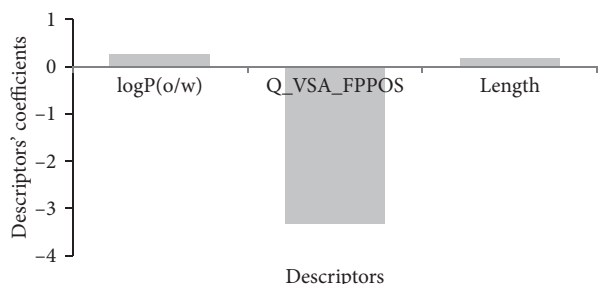


FIGURE 4: Contribution of each descriptor from model 02 to biological activity.

have a positive contribution to the inhibition (Log P(o/w), length) and Q_VSA_FPPOS with a negative contribution. High values of descriptors Log P(o/w) and length and low values of descriptor Q_VSA_FPPOS are favorable for *T. cruzi* inhibition.

Remarkably, the biological activity in our proposed model is explained by geometrical descriptors and descriptors encoding hydrophobicity and London dispersion forces. Descriptors such as length and Log P(o/w) describe the molecular dimensions and hydrophobicity, respectively, for successful binding with the substrate. Log P(o/w) is the log value of the octanol/water partition coefficient. This property was calculated in MOE from a linear atom-type model using 1,847 molecules [53].

Q_VSA_FPPOS is the code for fractional positive polar van der Waals surface area. Q_VSA_FPPOS is the sum of v_i (van der Waals surface area of atom i) such that q_i (partial charge of atom i) is greater than 0.2, divided by the total surface area [54]. According to model 02, low positive polar surface areas favor biological activity against *T. cruzi*. This highlights the importance of descriptors Log P(o/w), length, and Q_VSA_FPPOS in achieving *T. cruzi* inhibition over a wide range of pIC_{50} from 0.056 to 2.301, which in turn makes this model significant [52].

CYP51 is an enzyme with a rigid substrate-binding cavity; fenarimol analogues interact via a coordination bond with iron in the enzyme's heme group, and the other two rings are directed to substrate cavities via van der Waals contact with residues Val213, Ile105, Tyr103, Met106, Tyr116, Ala115, Phe110, Leu127, Met284, Leu130, Ala287, Phe290, Ala291, Thr295, Leu356, Met460, Met360, and Met358 [26]. Two aromatic planar rings (pyridine/pyrimidine and substituted benzene) and a heterocyclic nonplanar ring (piperazine) are the pharmacophoric features of fenarimol analogues. Clearly, the nature of the van der Waals interactions in the binding pocket is due to multiple aromatic stacking and hydrogen bonding. A decrease in the molecular surface area bearing a polar positive charge would favor intermolecular interactions and coordination with iron in the enzyme's heme group.

Low polarity contributes to our model because 5 out of 7 residues (Val102, Ile105, Met106, Phe110, and Ala115) at the active site in CYP51 are nonpolar [55]. Additionally, a 42 residue-long hydrophobic tunnel connects the heme group with the protein surface [55], so high partition coefficients such as Log P(o/w) values are required to enter the mostly hydrophobic CYP51 active site. The longer groups substituted at R1 of the piperazine ring (length descriptor) favor *T. cruzi* inhibition due to a deeper entrance in the binding pocket allowing further contact with the residues.

Consequently, if we want to increase the value of the activity, we will start with the most active group of fenarimol analogues, carbonates, or carbamates 33–44 and then

- (1) Increase Log P: substituents should favor hydrophobicity
- (2) Increase length: the substitution pattern in the piperazine moiety should contain carbonates or carbamates substituted with a large aromatic ring
- (3) Decrease Q_VSA_FPPOS: avoid polar positive groups

4.2. Correlation of Biological Activity with Descriptor Values.

For the biological activity of amides 3–17, it appears that the bulkier alkyl amides 3–8 are more potent than benzamides 10 and 11 as well as being more potent than amides 13, 14, and 16 possessing smaller alkyl groups (Table 3). Amides generally have the shortest lengths (7.596–7.951 Å) at the piperazine end of the pharmacophore (length descriptor) (Table 2). This shorter length of amides at the piperazine end might limit the access to active sites and interactions with residues that are more important for activity. Pyrimidine amides 12 and 13 are less potent than pyridine amides (Table 3) and have Log P(o/w) values of 2.565 and 1.715, respectively (Table 2); these structures have the lowest values of hydrophobicity in the amide group. Therefore, the low hydrophobicity of pyrimidine derivatives might determine their low activity compared to pyridine derivatives.

In the sulfonamide group (18–32), aryl sulfonamides 18–23 showed higher activity than alkyls 24–32. The length of aryl sulfonamides ranged 8.006–10.866 Å, while that of alkyls ranged 6.398–6.708 Å except for 26 (8.253 Å). Alkyl sulfonamides 31 and 32 are pyrimidine derivatives that showed lower potency than the pyridine derivatives. In addition, low Log P(o/w) values (0.434–0.777) were observed for alkyl sulfonamides 31 and 32 and aryl sulfonamide 30 (Table 2). In general, pyrimidinyl analogues were less potent than the corresponding pyridyl compounds (e.g., amides 3, 12 and 5, 13; sulfonamides 24, 25 and 28, 31, as well as 29, 32; and carbamates 42, 40 and 46, 53). Lower basicity of pyrimidinyl rings could be accounted for the reduced coordination of nitrogen with heme iron of the CYP51 enzyme and hence lower activity against *T. cruzi*.

Carbonate and carbamate analogues 33–44 contained piperazines with the longest chain length (10.305–12.952 Å (Table 2)) among the compounds examined in this study and were found to be the most potent against *T. cruzi*. In general,

TABLE 10: Average values of descriptors.

	Length	Q_VSA_FPPOS	Log P(o/w)
Most active molecules	10.155	0.183	3.904
Less active molecules	7.309	0.246	1.797

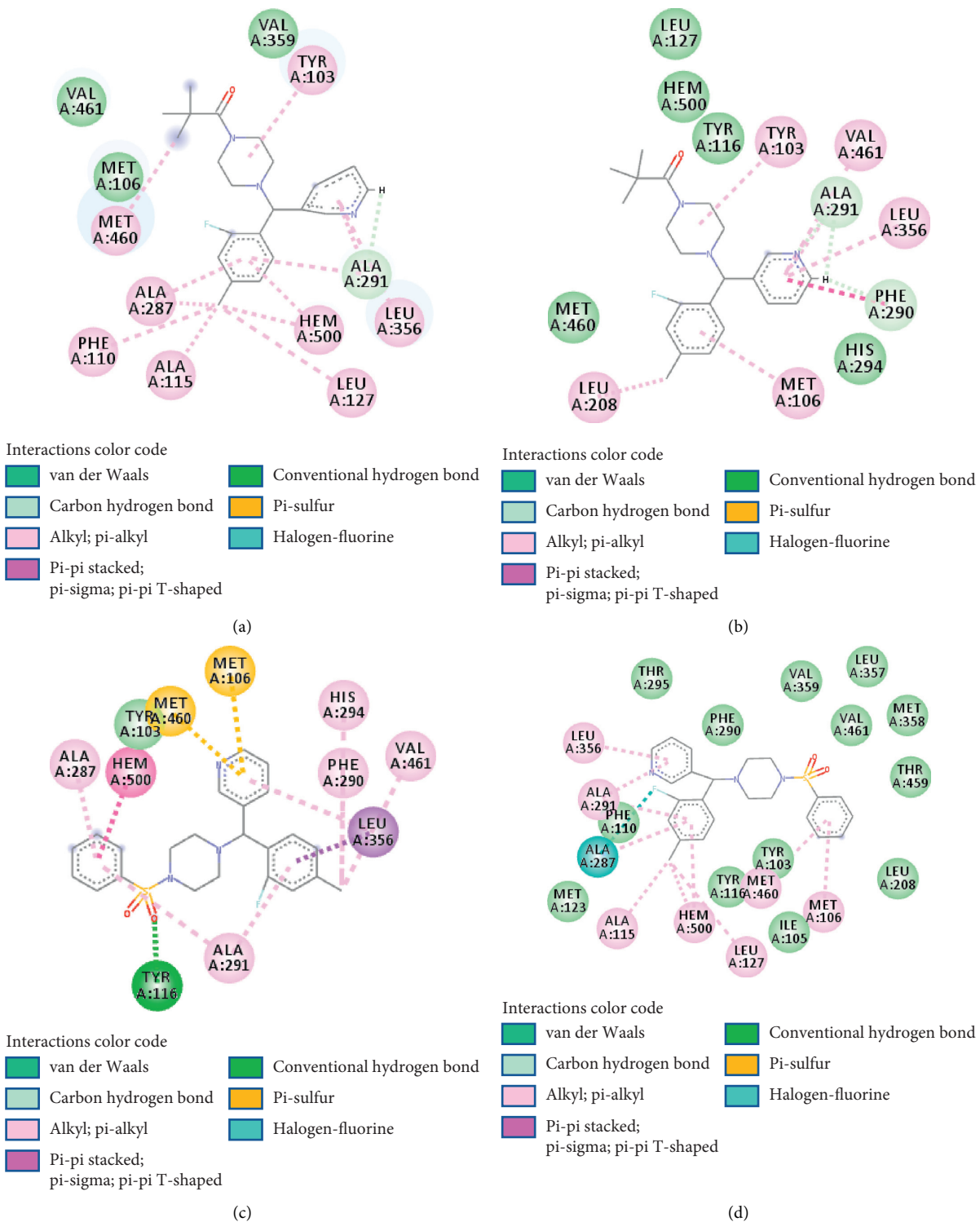


FIGURE 5: Continued.

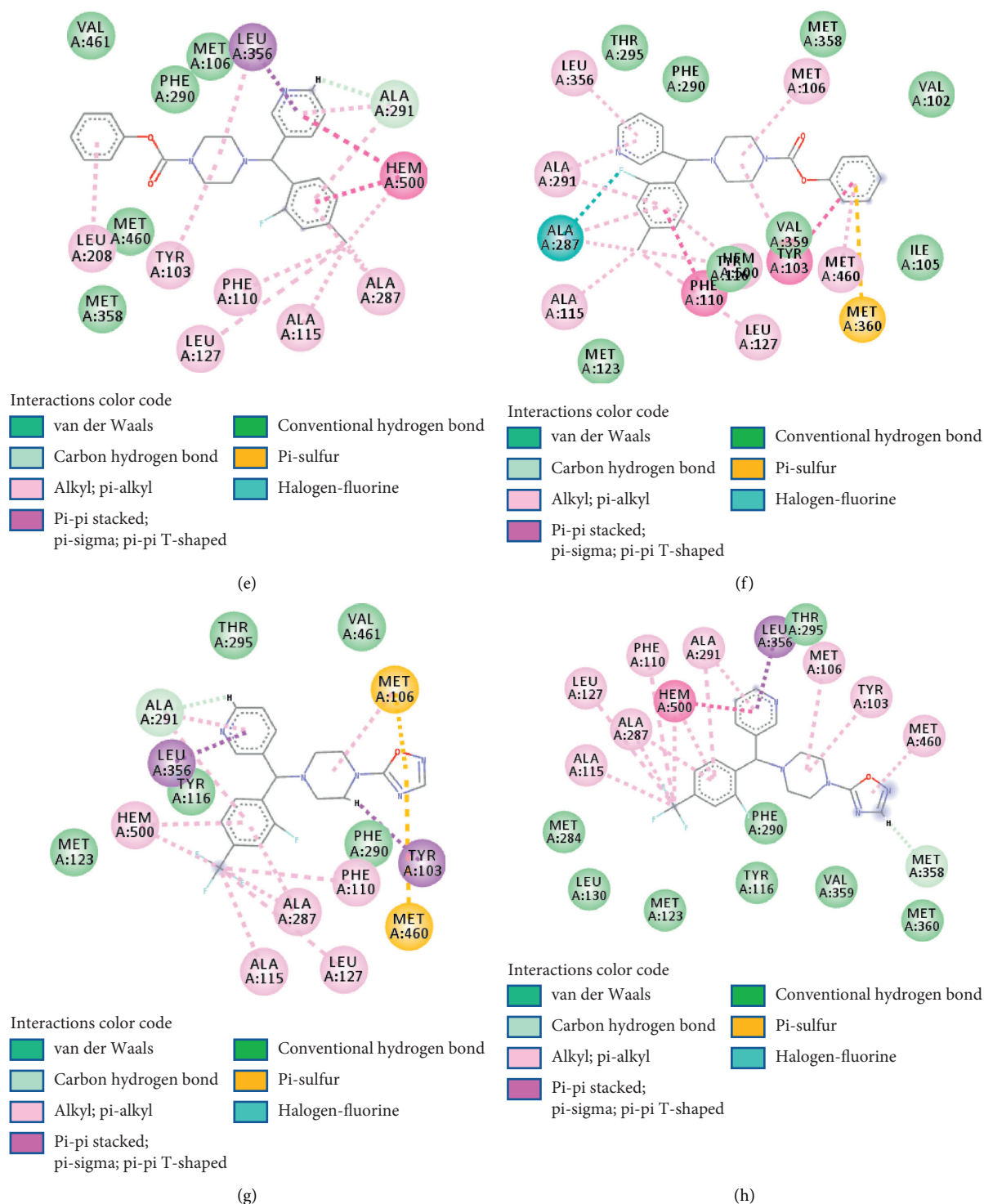


FIGURE 5: Predicted binding of ligand stereoisomers 3R, 3S, 18R, 18S, 33R, 33S, 45R, and 45S to the active site of sterol 14 α -demethylase (CYP51) from *Trypanosoma cruzi*. (a) Molecule 3(R). (b) Molecule 3(S). (c) Molecule 18(R). (d) Molecule 18(S). (e) Molecule 33(R). (f) Molecule 33(S). (g) Molecule 45(R). (h) Molecule 45(S).

it was observed that short substituent lengths, low Log P(o/w), and high Q_VSA_FPPOS decrease the activity of aryl molecules 45–54 to a level comparable to that of amides 3–17. The average values of the descriptor relating to activity are shown in Table 10.

Accessibility to the binding pocket, hydrophobicity, and low positive polar surface area make carbamates the most potent against *T. cruzi*, followed by sulfonamides. The structures of carbamates and sulfonamides contain high electron density areas because of the presence of multiple

TABLE 11: Distance to Fe and scoring function values.

Molecule	Distance to Fe (Å)	Scoring function (Kcal/mole)
3(<i>R</i>)	2.68	-8.2
3(<i>S</i>)	6.02	-7.9
18(<i>R</i>)	10.79	-9.4
18(<i>S</i>)	2.80	-10.3
33(<i>R</i>)	5.53	-9.7
33(<i>S</i>)	2.78	-10.3
45(<i>R</i>)	2.62	-9.5
45(<i>S</i>)	4.74	-10.1

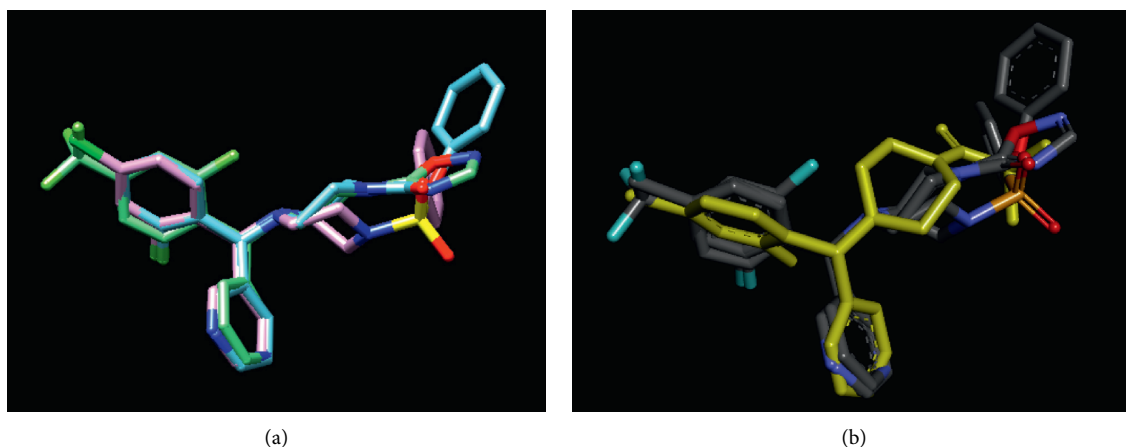


FIGURE 6: (a) Superposed conformations of molecules 18(*S*) (violet), 33(*S*) (pale green), and 45(*S*) (blue). (b) Superposed conformations of molecules 3(*R*) (yellow), 18(*S*), 33(*S*), and 45(*S*) in conventional atom color codes of the Discovery Data program.

electronegative heteroatoms that facilitate the formation of van der Waals contacts with the polar residues present at the active site.

4.3. Docking. Stereochemistry plays a major role in biological activity; there is no scientific literature about the activity of isolated stereoisomers of piperazine analogues of fenarimol used in this paper. Docking is an *in silico* methodology to evaluate interactions at the active site and therefore three-dimensional structural properties such as chirality. Docking was performed using the *R* and *S* stereoisomers of the molecules with the best biological activity for each substituent subtype, that is, amide 03, sulfonamide 18, carbamate 33, and aryl-substituted molecule 45. All stereoisomers act on the active site reported in the literature. Multiple interactions of the aromatic stacking type, π - π , π -alkyl, and π -sulfur type were found for all tested molecules (Figure 5). Hydrogen bonding was found only in 18(*R*) stereoisomers (Figure 5).

Among the stereoisomers, those with the highest affinity for the receptor are 3(*R*), 18(*S*), 33(*S*), and 45(*S*) (Table 11).

Stereoisomers with coordination distances to Fe between 2.62 and 2.80 Å display the strongest interacting conformations. Stereoisomers with greater distances to Fe demonstrated no possible coordination and were, by comparison, weaker interacting conformations. Molecule 45 was an exception because the strongest interacting conformation was not coordinated to Fe (Table 11).

Experimental evaluations of the biological activities of isolated stereoisomers in the series of piperazine analogues of fenarimol are necessary to correlate the findings in this paper. The biological activity expressed as IC_{50} based on currently available data was determined using racemic mixtures of the molecules under study. The spatial arrangement and different coordination distances to Fe can lead to different IC_{50} values depending on the type of *R/S* stereoisomerism exhibited by the fenarimol analogues. Stereoselectivity is one of the most important factors to consider to improve drug safety and efficacy; therefore, efforts to elucidate the relevance of stereoselectivity in the inhibition of sterol 14 α -demethylase (CYP51) from *Trypanosoma cruzi* would benefit this research area.

Piperazine rings in stereoisomers 3(*R*)/(*S*) and 18(*R*)/(*S*) are in a twisted boat conformation, whereas in stereoisomers 33(*S*)/(*R*) and 45(*S*)/(*R*), the adopted conformation is chair. The twisted boat and chair conformations orient toward different residues in the active site. The piperazine ring in stereoisomers 3(*R*)/(*S*) interacts with Tyr103, whereas there are no interactions of piperazine with any residues in stereoisomers 18(*R*)/(*S*). Stereoisomers 33(*S*)/(*R*) also orient towards Met106 and Leu356, respectively, with stereoisomers 45(*S*)/(*R*) also oriented toward Met106.

By superimposing the more stable conformations resulting from docking of molecules 18(*S*), 33(*S*), and 45(*S*), an almost perfect match is observed between piperidines, piperazine, and lateral aromatic substituents except for

sulfonamide 18(S), which suffers a considerable deviation in the piperazine ring (Figure 6(a)). On the contrary, molecule 3(R) does not coincide with the superimposed conformations due to its chiral arrangement (Figure 6(b)).

5. Conclusions

In this study, multiple linear regression (MLR) models were developed to determine the molecular properties relevant to the inhibitory activities of 54 fenarimol analogues against Chagas disease. Twelve descriptors were identified and statistically validated to build a QSAR model. The length, Q-VSA_FPPOS, and Log P(o/w) are determinant descriptors in the prediction of biological activity against Chagas disease. The MLR model has demonstrated that the most active compounds have low values of Q-VSA_FPPOS and high values of length and Log P(o/w). The stability and robustness of the model are supported by the statistical parameters and validation tests applied, indicating that the strategy implemented in this study can be used to predict the biological activity of piperazine analogues of fenarimol and to adopt a rational design for developing new analogues with improved activity against Chagas disease. Stereoisomers 3(R), 18(S), 33(S), and 45(S) showed the highest affinity to the receptor, with coordination distances of approximately 2.62–2.80 Å that were found to have the strongest interactions with the receptor. The most frequently found interactions with CYP51 active sites were aromatic stacking, π - π , π -alkyl, and π -sulfur interactions. Experimental determination of the biological activity of isolated stereoisomers has not been reported in the literature for piperazine analogues of fenarimol; therefore, this paper makes an important contribution to the research needed in this area. For virtual screening and computational design of drugs against Chagas disease, it will be essential to define stereoselectivity to improve drug safety and efficacy. Knowledge of different pharmacokinetic and pharmacodynamic profiles of stereoisomers could lead to substantial development in this research area. Chirality in drug design is an issue for both pharmaceutical industry and regulatory authorities. Factors such as industrial scale, production of stereoisomers, stereochemical stability, toxicological profile, and differential clinical results highlight the importance of stereoselectivity in drug design. In the case of Chagas disease, a neglected tropical disease with no effective treatment and high toxicity levels of current therapeutics, it is of vital importance to explore these issues in depth to find the most effective treatment.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Materials

Table S1: regression models. Table S2: statistics of regression models. Table S3: LOO cross-validation statistics for model 02. (*Supplementary Materials*)

References

- [1] C. Chagas, “Nova tripanozomíase humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem,” *Memórias Do Instituto Oswaldo Cruz*, vol. 1, no. 2, pp. 159–218, 1909.
- [2] C. Chagas, “Nouvelle espèce de trypanosomiase humaine,” *Bulletin de la Société de Pathologie Exotique*, vol. 2, no. 6, pp. 304–307, 1909.
- [3] C. Chagas, “Ueber eine neue Trypanosomiasis des Menschen. Zweite vorläufige Mitteilung,” *Archiv für Schiffs- und Tropenhygiene unter besonderer Berücksichtigung der Pathologie und Therapie*, vol. 13, no. 11, pp. 351–353, 1909.
- [4] M. C. Kulik, H. E. Karim-Kos, J. A. Haagsma et al., “Socio-economic inequalities in neglected tropical diseases: a systematic review,” *PLoS Neglected Tropical Diseases*, vol. 10, no. 5, pp. 1–28, 2016.
- [5] J. Bermudez, C. Davies, A. Simonazzi, J. Pablo Real, and S. Palma, “Current drug therapy and pharmaceutical challenges for Chagas disease,” *Acta Tropica*, vol. 156, pp. 1–16, 2016.
- [6] A. F. Yépes, A. Bahsas, P. Escobar et al., “Synthesis, anti-parasitic activity and QSAR study of a new library of poly-substituted tetrahydronaphtho[1,2-b]azepines,” *Medicinal Chemistry Research*, vol. 27, no. 10, pp. 2239–2264, 2018.
- [7] C. Bern, “Antitrypanosomal therapy for chronic Chagas’ disease,” *New England Journal of Medicine*, vol. 364, no. 26, pp. 2527–2534, 2011.
- [8] Y. Moolani, G. Bukhman, and P. J. Hotez, “Neglected tropical diseases as hidden causes of cardiovascular disease,” *PLoS Neglected Tropical Diseases*, vol. 6, no. 6, p. e1499, 2012.
- [9] C. A. F. Aiub, N. Boechat, M. M. Bastos, I. Felzenszwalb, F. V. C. Mello, and A. S. Carvalho, “Evaluation of genotoxic effects of new molecules with possible trypanocidal activity for Chagas disease treatment,” *Scientific World Journal*, vol. 2013, Article ID 287319, 8 pages, 2013.
- [10] M.-J. Pinazo, L. Guerrero, E. Posada, E. Rodríguez, D. Soy, and J. Gascon, “Benznidazole-related adverse drug reactions and their relationship to serum drug concentrations in patients with chronic Chagas disease,” *Antimicrobial Agents and Chemotherapy*, vol. 57, no. 1, pp. 390–395, 2013.
- [11] F. A. Bernal and T. J. Schmidt, “A comprehensive QSAR study on antileishmanial and antitrypanosomal cinnamate ester analogues,” *Molecules*, vol. 24, no. 23, p. 4358, 2019.
- [12] F. Bianchi, Z. Cucunubá, F. Guhl et al., “Follow-up of an asymptomatic Chagas disease population of children after treatment with nifurtimox (lampit) in a sylvatic endemic transmission area of Colombia,” *PLoS Neglected Tropical Diseases*, vol. 9, no. 2, Article ID e0003465, 2015.
- [13] M. L. Belaunzarán, “Enfermedad de Chagas: globalización y nuevas esperanzas para su cura,” *Revista Argentina de Microbiología*, vol. 47, no. 2, pp. 85–87, 2015.
- [14] L. Ventura-Garcia, M. Roura, C. Pell et al., “Socio-cultural aspects of Chagas disease: a systematic review of qualitative research,” *PLoS Neglected Tropical Diseases*, vol. 7, no. 9, p. e2410, 2013.

- [15] A. Trochine, D. J. Creek, P. Faral-Tello, M. P. Barrett, and C. Robello, "Benzimidazole biotransformation and multiple targets in trypanosoma cruzi revealed by metabolomics," *PLoS Neglected Tropical Diseases*, vol. 8, no. 5, p. e2844, 2014.
- [16] A. Buschini, L. Ferrarini, S. Franzoni et al., "Genotoxicity reevaluation of three commercial nitroheterocyclic drugs: nifurtimox, benzimidazole, and metronidazole," *Journal of Parasitology Research*, vol. 2009, pp. 1–11, Article ID 463575, 2009.
- [17] L. Jaramillo, C. Ruiz, L. Martínez, and S. Vera, "Enfermedad de Chagas: una mirada alternativa al tratamiento," *Revista Cubana de Medicina Tropical*, vol. 69, no. 2, pp. 1–13, 2017.
- [18] S. P. S. Rao, M. P. Barrett, G. Dranoff et al., "Drug discovery for kinetoplastid diseases: future directions," *ACS Infectious Diseases*, vol. 5, no. 2, pp. 152–157, 2019.
- [19] M. C. Field, D. Horn, A. H. Fairlamb et al., "Erratum: anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need," *Nature Reviews Microbiology*, vol. 15, no. 7, p. 447, 2017.
- [20] M. Rycker, B. Baragaña, S. L. Duce, and I. H. Gilbert, "Challenges and recent progress in drug discovery for tropical diseases," *Nature*, vol. 559, no. 7715, pp. 498–506, 2019.
- [21] J. A. Urbina, "Ergosterol biosynthesis and drug development for Chagas disease," *Memórias Do Instituto Oswaldo Cruz*, vol. 104, no. 1, pp. 311–318, 2009.
- [22] J. Urbina, "Chemotherapy of Chagas disease," *Current Pharmaceutical Design*, vol. 8, no. 4, pp. 287–295, 2002.
- [23] J. A. Urbina, "Specific chemotherapy of Chagas disease: relevance, current limitations and new approaches," *Acta Tropica*, vol. 115, no. 1–2, pp. 55–68, 2010.
- [24] J. A. Urbina and R. Docampo, "Specific chemotherapy of Chagas disease: controversies and advances," *Trends in Parasitology*, vol. 19, no. 11, pp. 495–501, 2003.
- [25] J. A. Urbina, G. Payares, C. Sanoja, R. Lira, and A. J. Romanha, "In vitro and in vivo activities of ravuconazole on Trypanosoma cruzi, the causative agent of Chagas disease," *International Journal of Antimicrobial Agents*, vol. 21, no. 1, pp. 27–38, 2003.
- [26] J. Y. Choi, L. M. Podust, and W. R. Roush, "Drug strategies targeting CYP51 in neglected tropical diseases," *Chemical Reviews*, vol. 114, no. 22, pp. 11242–11271, 2014.
- [27] G. I. Lepesheva, T. Y. Hargrove, S. Anderson et al., "Structural insights into inhibition of sterol 14 α -demethylase in the human pathogen trypanosoma cruzi," *Journal of Biological Chemistry*, vol. 285, no. 33, pp. 25582–25590, 2010.
- [28] A. S. Nagle, S. Khare, A. B. Kumar et al., "Recent developments in drug discovery for leishmaniasis and human african trypanosomiasis," *Chemical Reviews*, vol. 114, no. 22, pp. 11305–11347, 2014.
- [29] G. I. Lepesheva, T. Y. Hargrove, S. Anderson et al., "Image from the RCSB PDB (Rcsb.Org) of PDB ID 3KHM," , 2010.
- [30] M. Berriman, E. Ghedin, C. Hertz-Fowler et al., "The genome of the African trypanosome," *Science*, no. 5733, pp. 416–422, 2005.
- [31] El-S. Najib, P. J. Myler, D. C. Bartholomeu et al., "The genome sequence of trypanosoma cruzi, etiologic agent of Chagas disease," *Science*, vol. 2005, no. 409, pp. 409–416, 2005.
- [32] P. M. Njogu, E. M. Guantai, E. Pavadai, and K. Chibale, "Computer-aided drug discovery approaches against the tropical infectious diseases malaria, tuberculosis, trypanosomiasis, and leishmaniasis," *ACS Infectious Diseases*, vol. 2, no. 1, pp. 8–31, 2016.
- [33] A. Dudek, T. Arodz, and J. Galvez, "Computational methods in developing quantitative structure-activity relationships (QSAR): a review," *Combinatorial Chemistry & High Throughput Screening*, vol. 9, no. 3, pp. 213–228, 2006.
- [34] M. Keenan, M. J. Abbott, P. W. Alexander et al., "Analogues of fenarimol are potent inhibitors of trypanosoma cruzi and are efficacious in a murine model of Chagas disease," *Journal of Medicinal Chemistry*, vol. 55, no. 9, pp. 4189–4204, 2012.
- [35] M. Keenan, P. W. Alexander, H. Diao et al., "Design, structure-activity relationship and in vivo efficacy of piperazine analogues of fenarimol as inhibitors of Trypanosoma cruzi," *Bioorganic & Medicinal Chemistry*, vol. 21, no. 7, pp. 1756–1763, 2013.
- [36] D. M. Hawkins, "The problem of overfitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 1, pp. 1–12, 2004.
- [37] V. H. Masand, N. N. E. El-Sayed, M. U. Bambole, V. R. Patil, and S. D. Thakur, "Multiple quantitative structure-activity relationships (QSARs) analysis for orally active trypanocidal N-myristoyltransferase inhibitors," *Journal of Molecular Structure*, vol. 1175, pp. 481–487, 2019.
- [38] V. H. Masand, D. T. Mahajan, A. K. Maldhure, and V. Rastija, "Quantitative structure-activity relationships (QSARs) and pharmacophore modeling for human African trypanosomiasis (HAT) activity of pyridyl benzamides and 3-(oxazol[4,5-b]pyridin-2-yl)anilides," *Medicinal Chemistry Research*, vol. 25, no. 10, pp. 2324–2334, 2016.
- [39] S. Chtita, M. Ghamali, A. Ousaa et al., "QSAR study of anti-human African trypanosomiasis activity for 2-phenylimidazopyridines derivatives using DFT and Lipinski's descriptors," *Heliyon*, vol. 5, no. 3, Article ID e01304, 2019.
- [40] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., Gaussian, Inc., Wallingford, UK, 2016.
- [41] Molecular Operating Environment (MOE), Montreal, Canada, 2018, <https://www.chemcomp.com/Research.htm>.
- [42] K. Burke, "Perspective on density functional theory," *The Journal of Chemical Physics*, vol. 136, no. 15, Article ID 150901, 2012.
- [43] T. Scior, J. Medina-Franco, Q.-T. Do, K. Martinez-Mayorga, J. Yunes Rojas, and P. Bernard, "How to recognize and work-around pitfalls in QSAR studies: a critical review," *Current Medicinal Chemistry*, vol. 16, no. 32, pp. 4297–4313, 2009.
- [44] E. C. Ibezim, P. R. Duchowicz, N. E. Ibezim et al., "Computer-aided linear modeling employing QSAR for drug discovery," *Scientific Research and Essays*, vol. 4, no. 13, pp. 1559–1564, 2009.
- [45] J. G. Topliss and R. J. Costello, "Chance correlations in structure-activity studies using multiple regression analysis," *Journal of Medicinal Chemistry*, vol. 15, no. 10, pp. 1066–1068, 1972.
- [46] 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.
- [47] C. J. Guerra, J. M. López, S. F. Figueredo, A. E. Muñoz, and J. R. Robles, "2D-QSAR analysis of derivatives of quinoxaline 1,4-Di-n-oxides with activity against Chagas' disease," *Química Nova*, vol. 39, no. 6, pp. 647–654, 2016.
- [48] O. Trott and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of Computational Chemistry*, vol. 31, no. 2, pp. 455–461, 2019.
- [49] BIOVIA, Dassault Systèmes, version 16. 1.0., BIOVIA, San Diego, CA, USA, 2019.

- [50] E. F. Pettersen, T. D. Goddard, C. C. Huang et al., "UCSF Chimera? A visualization system for exploratory research and analysis," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.
- [51] A. E. Ensuncho, J. R. Robles, and S. F. Figueredo, "Modelado de las Relaciones Cuantitativas Estructura-Actividad (QSAR) de los Derivados 5-(Nitroheteroaril)-1,3,4-tiadiazol con Actividad Leishmanicida," *Información Tecnológica*, vol. 28, no. 2, pp. 191–202, 2017.
- [52] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C. P. Varghese, and R. K. Agrawal, "Validation of QSAR models-strategies and importance," *International Journal of Drug Design & Discovery*, vol. 2, no. 3, pp. 511–519, 2011.
- [53] P. Labute, MOE LogP(octanol/water), Source code in \$MOE/Lib/Svl/Quasar.Svl/Q_logp.Svl.
- [54] MOE 2020 Tutorials, <https://www.chemcomp.com/moe/help/latest/>.
- [55] C.-K. Chen, S. S. F. Leung, C. Guilbert, M. P. Jacobson, J. H. Mckerrow, and L. M. Podust, "Structural characterization of CYP51 from trypanosoma cruzi and trypanosoma brucei bound to the antifungal drugs posaconazole and fluconazole," *PLoS Neglected Tropical Diseases*, vol. 4, no. 4, p. e651, 2010.