

# Molecular docking and quantitative structure-activity relationships for a series of *Trypanosoma cruzi* dihydroorotate dehydrogenase inhibitors

Gabriela Ciffeli de Jesus<sup>1</sup>, Tatiana Santana Ribeiro<sup>2</sup>,  
Adriano D. Andricopulo<sup>3</sup>, Leonardo Luiz Gomes Ferreira<sup>4\*</sup>

<sup>1,2</sup>Department of Natural Science, Mathematics and Education, Federal University of Sao Carlos, Araras, SP, Brazil, <sup>3,4</sup>Laboratory of Medicinal and Computational Chemistry, Physics Institute of Sao Carlos, Sao Carlos, University of Sao Paulo, SP, Brazil

Caused by the protozoan *Trypanosoma cruzi*, Chagas disease affects six to seven million people worldwide, mainly in Latin America. The drugs currently available for treating the disease are ineffective during its chronic phase and have serious adverse effects. Essential for the survival of *T. cruzi*, the enzyme dihydroorotate dehydrogenase (DHODH) has become a key molecular target for drug discovery in Chagas disease. This study investigates the bi-dimensional and three-dimensional quantitative structure-activity relationships (QSAR) for a series of 64 *T. cruzi* DHODH inhibitors. The results indicate a highly predictive 2D Hologram QSAR (HQSAR) model ( $q^2 = 0.65$ ,  $r^2 = 0.88$ , and  $r^2_{\text{pred}} = 0.82$ ) that identified key molecular fragments that correlate with DHODH inhibition. Moreover, 3D Comparative Molecular Field Analysis (CoMFA) models ( $q^2 = 0.75$ ,  $r^2 = 0.99$ , and  $r^2_{\text{pred}} = 0.66$ ) pointed out the 3D molecular features that determine the activity of the inhibitors. Although restricted to a congeneric series and focused solely on 2D and 3D descriptors, these QSAR models and molecular docking analyses identified key properties and intermolecular interactions for designing and optimizing new compounds as potent *T. cruzi* DHODH inhibitors.

**Keywords:** *Trypanosoma cruzi*. Medicinal chemistry. QSAR. Dihydroorotate dehydrogenase. Inhibitors. Chagas disease.

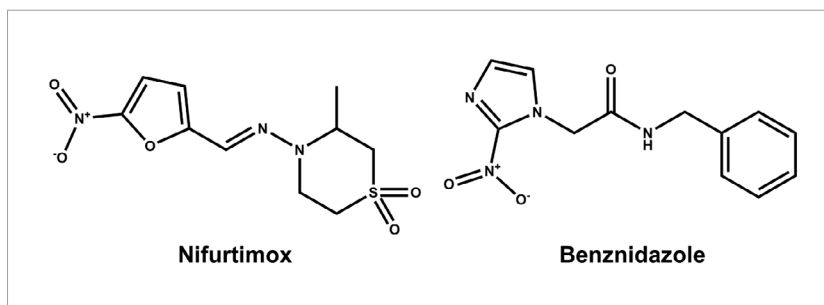
## INTRODUCTION

Neglected tropical diseases (NTDs) comprise a diverse group of 20 conditions affecting over 1.5 billion people in more than 150 countries, primarily in developing nations (Ferreira, Andricopulo, 2019). Among NTDs, Chagas disease is an endemic anthroponosis in Latin America and has become a global health issue affecting 6-7 million people worldwide (WHO, 2024). The highest burden in a nonendemic country occurs in the USA, where approximately 300,000 people are estimated to have the disease (Irish *et al.*, 2022). Chagas disease significantly

affects the economies of endemic countries, causing annual losses exceeding US\$ 7.2 billion and a burden of 138,000 disability-adjusted life years (DALYs) (Arnal *et al.*, 2019) (GBD 2017 DALYs, HALE Collaborators, 2018).

The etiological treatment for Chagas disease has not advanced since the 1970s and continues to rely on nifurtimox and benznidazole (Figure 1), two nitroheterocyclic drugs known for their limited efficacy during the chronic phase of the disease and their severe adverse side effects. These adverse reactions include anorexia, gastrointestinal symptoms, neurological toxicity, irritability, insomnia, weight loss, and seizures (Pérez-Molina *et al.*, 2021). According to the WHO's 2021-2030 roadmap for NTDs, these challenges, compounded with emerging drug resistance, indicate the urgent need for new and innovative treatments for this disease (WHO, 2021).

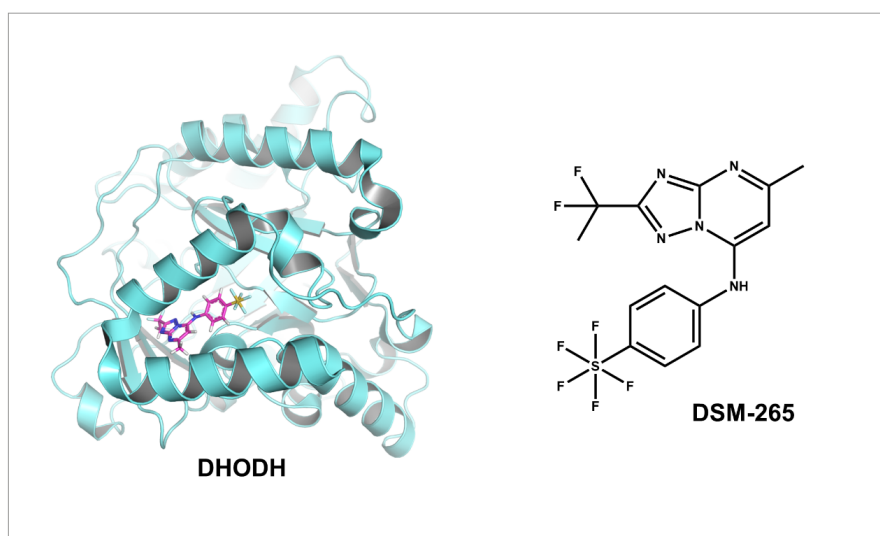
\*Correspondence: L. L. G. Ferreira Laboratório de Química Medicinal e Computacional Instituto de Física de São Carlos Universidade de São Paulo Av. Joao Dagnone, 1100, Santa Angelina, 13563-120, Sao Carlos, SP, Brasil Phone: (16) 3373-9855 Email: leonardo@ifsc.usp.br <https://orcid.org/0000-0002-6947-0639> T. S. Ribeiro, <https://orcid.org/0000-0002-0099-0177> A. D. Andricopulo, <https://orcid.org/0000-0002-0457-818X>



**FIGURE 1** - Structure of nifurtimox and benznidazole, two drugs currently available for the chemotherapy of Chagas disease.

The identification, validation and exploration of novel molecular targets have been a challenging topic in Chagas disease and NTD drug discovery (Ferreira, de Moraes, Andricopulo, 2022). In this situation, the genetic validation of dihydroorotate dehydrogenase (DHODH) as an essential enzyme for the survival of *T. cruzi* is an important advance in the field (Hashimoto *et al.*, 2012). DHODH has additionally been explored as a molecular target in other parasitic diseases such as leishmaniasis, malaria, and sleeping sickness (Pinheiro, Emery, Nonato, 2013; Singh *et al.*, 2017; Chibli *et al.*, 2018; Arakaki *et al.*, 2008). The enzyme catalyzes the oxidation of (*S*)-dihydroorotate to orotate, a key step in the synthesis of

pyrimidine, which plays a central role in the metabolism of DNA, RNA, glycoproteins, and lipids. Following these findings, multiple classes of inhibitors such as natural products, substrate-based, chalcone and flavone hybrids, and pyrimidine derivatives have been designed as DHODH inhibitors to target protozoan diseases (Chibli *et al.*, 2018; Thillainayagam, Malathi, Ramaiah, 2018; Azeredo *et al.*, 2017). The significance of DHODH as a molecular target for protozoan diseases has been demonstrated by DSM-265, a triazole derivative that advanced to phase 2 clinical trials for its potential in treating and preventing malaria (Figure 2) (Llanos-Cuentas *et al.*, 2018).



**FIGURE 2** - X-ray structure of DHODH from *Plasmodium falciparum* in complex with the inhibitor DSM-265 (PDB 5BOO, 2.80 Å). Protein structure is shown as a cartoon and inhibitor is depicted as sticks.

In the present study, we have investigated the molecular features that determine the activity of a series of orotate derivatives as *T. cruzi* DHODH inhibitors. To this end, we have applied for this series of inhibitors (Inaoka *et al.*, 2017), the 2D quantitative structure-activity relationship (QSAR) method Hologram QSAR (HQSAR) and the 3D approach Comparative Molecular Field Analysis (CoMFA) (Salum, Andricopulo, 2010; Cramer, Patterson, Bunce, 1988). HQSAR is a molecular fragmentation method. It generates all possible fragments (linear, cyclic, or branched) and correlates the occurrence of these fragments with the dependent variable, i.e., the biological activity, enzyme inhibition, or any other activity parameter, which was an innovation in QSAR when the method was published (Wang *et al.*, 2017; Salum, Andricopulo, Honório, 2012). The method has the great advantage of not requiring 3D conformations or molecular alignment. However, it may overlook aspects of activity linked to the shape of the compounds under study. CoMFA is a 3D method, which has as its main disadvantage the need for 3D conformation and molecular alignment, which can be challenging to achieve depending on the features of the data set compounds (Ghasemi *et al.*, 2013; Cherkasov *et al.*, 2014). However, CoMFA has the advantage of addressing steric and electrostatic features of the compounds as well as their shape. Importantly, CoMFA was the first method to address these descriptors in QSAR (Cramer, Patterson, Bunce, 1988). This is an important characteristic of the method given that biological activity and binding to enzymes and receptors are directly related to shape and electrostatic features.

Additionally, molecular docking was used for the structural alignment of the compounds and to assess their binding conformations in the binding cavity of DHODH. The combination of these approaches has resulted in solid QSAR models that can predict the activity of newly designed DHODH inhibitors. Furthermore, the integration of the QSAR and molecular docking results provided useful knowledge on the intermolecular recognition between DHODH and the compounds under study.

## MATERIAL AND METHODS

### Computational resources

The HQSAR and CoMFA models were built as previously described in our group (Souza, Ferreira, Andricopulo, 2017; Medeiros *et al.*, 2021) using the molecular modeling package SYBYL-X 2.1 (Certara Inc., Princeton, NJ, USA). The minimum-energy 3D structures for each dataset molecule were generated using the default parameters of the Tripos force field (Clark, Cramer III, Van Opdenbosch, 1989). The conjugate gradient method (Powell, 1977) was used with a 0.005 kcal/mol Å energy step and Gasteiger-Huckel charges with the dielectric constant of water (80) (Gasteiger, Marsili, 1980). The X-ray structure of DHODH deposited in the Protein Data Bank (PDB 2E6A, 1.64 Å) and GOLD 5.3 (Cambridge Crystallographic Data Centre, Cambridge, UK) (Jones *et al.*, 1997) were used for the molecular docking and structural alignment. The enzyme-inhibitor complexes resulting from the molecular docking runs were visualized using Pymol 1.3 (Schrodinger LLC, New York, NY, USA) (Lill, Danielson, 2011).

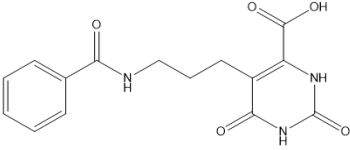
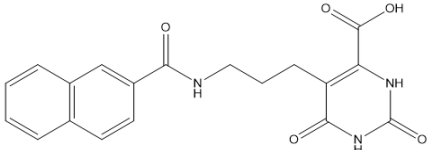
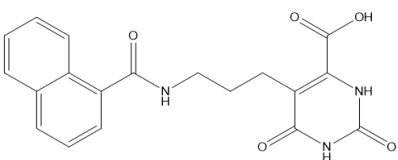
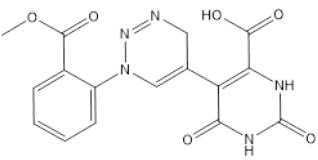
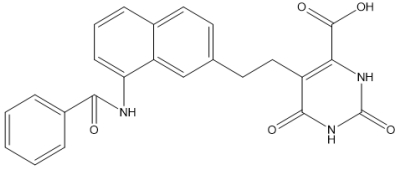
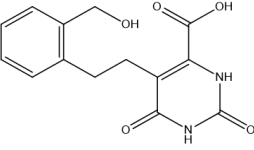
### Dataset compounds and selection of training and test sets

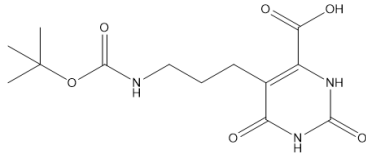
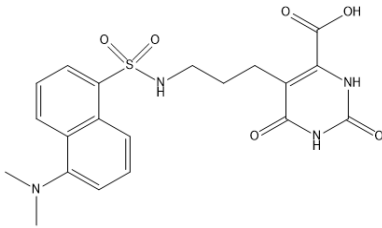
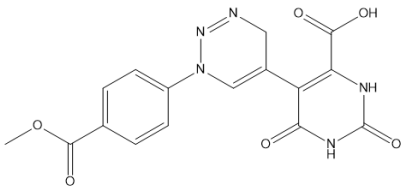
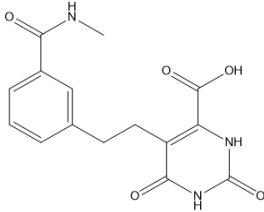
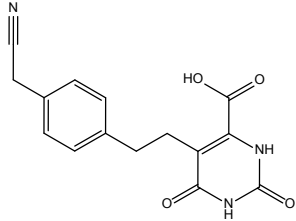
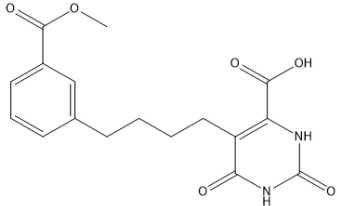
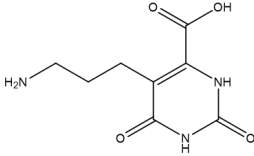
A dataset composed of 64 orotate derivatives selected from the literature was used to generate the 2D and 3D QSAR models (Inaoka *et al.*, 2017). Initially, the values of biological activity (concentration of the compounds required to inhibit 50% of enzyme activity,  $IC_{50}$  values) were converted into  $pIC_{50}$  ( $-\log IC_{50}$ ) to properly scale the data for the QSAR modeling. The dataset  $pIC_{50}$  values range from 5.13 to 7.48, which provides a suitable activity range for the generation of sound QSAR models. Importantly, the dataset compounds inhibit DHODH by the same mechanism, namely competitive inhibition, and the  $IC_{50}$  values were determined under the same experimental conditions. Table I shows the structures and  $pIC_{50}$  values for the complete dataset of DHODH

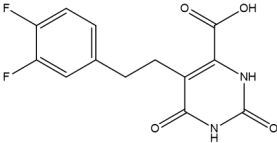
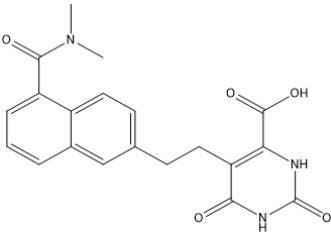
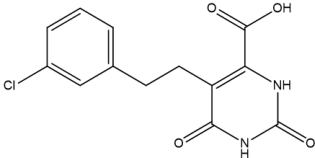
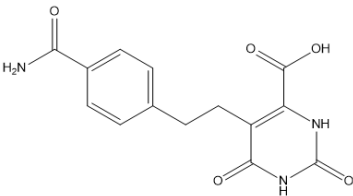
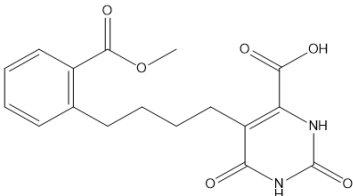
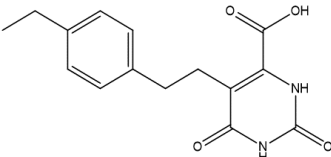
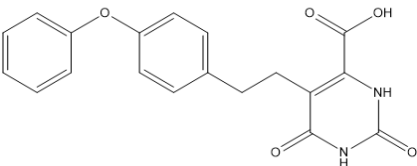
inhibitors used for the QSAR modeling. The dataset was divided into 51 training set molecules for the development of the QSAR models and 13 test set compounds (**1**, **4**, **12**, **20**, **21**, **31**, **34**, **36**, **43**, **50**, **56**, **62** and **63**) for external validation. This selection of both training and test sets was performed by conducting a principal component analysis (PCA) (Bender *et al.*, 2009). The similarity map generated by the PCA routine was built by using UNITY

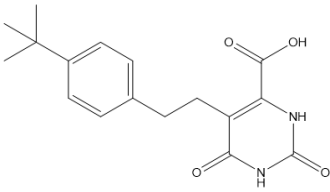
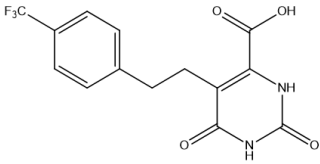
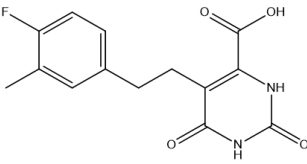
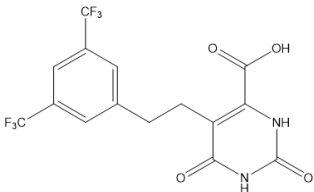
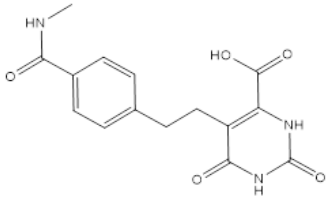
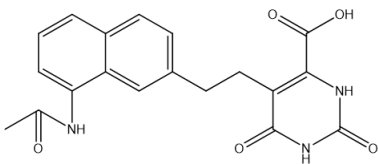
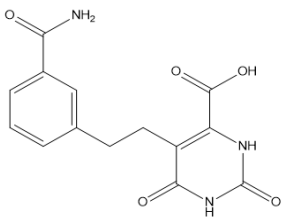
fingerprints as molecular attributes to assess the structural similarity between the compounds (Cereto-Massagué *et al.*, 2015). The PCA algorithm derived two principal components which were used as the initial coordinates to build the similarity map. All points in the map were plotted by calculating the Tanimoto distances between the UNITY fingerprints (Willett, 2006). Both sets included molecules covering a broad spectrum of  $pIC_{50}$  values.

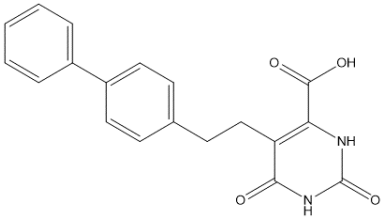
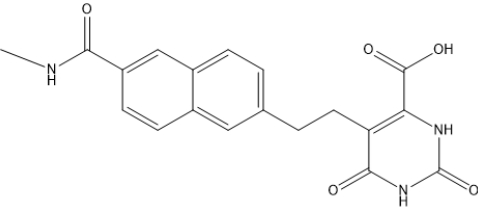
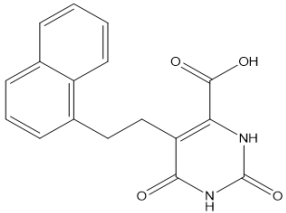
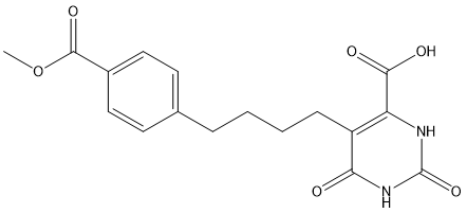
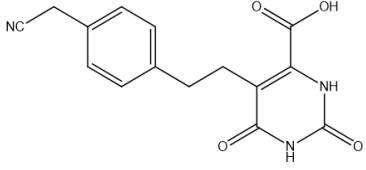
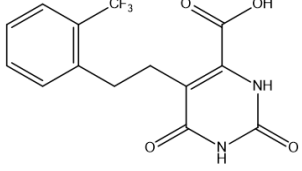
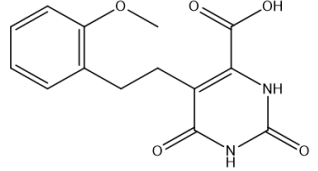
**TABLE I** - Structures and  $pIC_{50}$  values of the dataset compounds

Compound	Structure	$pIC_{50}$
<b>1</b>		5.13
<b>2</b>		5.17
<b>3</b>		5.25
<b>4</b>		5.31
<b>5</b>		5.47
<b>6</b>		5.51

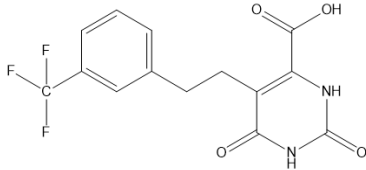
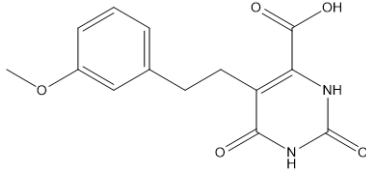
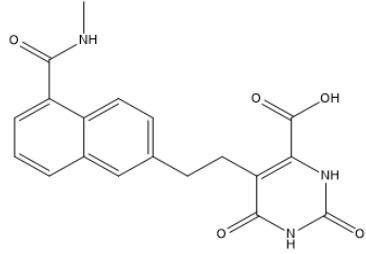
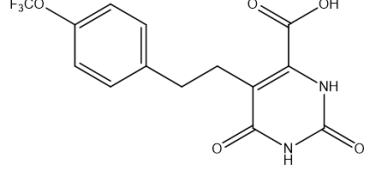
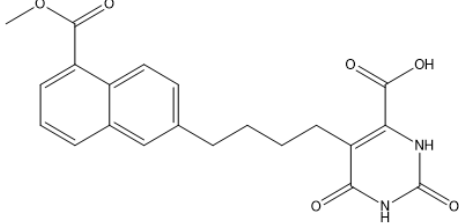
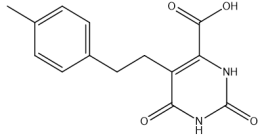
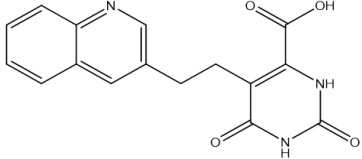
Compound	Structure	pIC <sub>50</sub>
7		5.53
8		5.58
9		5.60
10		5.63
11		6.13
12		5.67
13		5.69

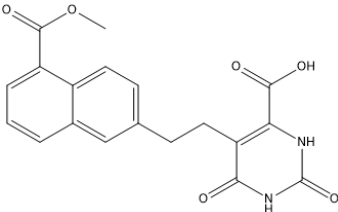
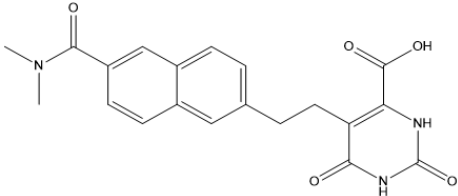
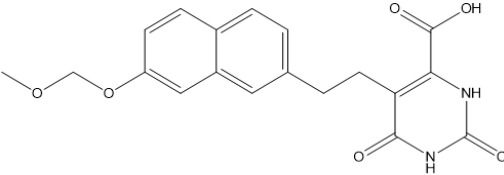
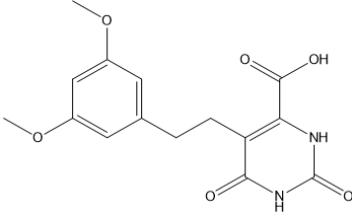
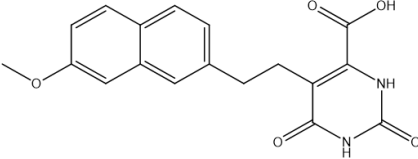
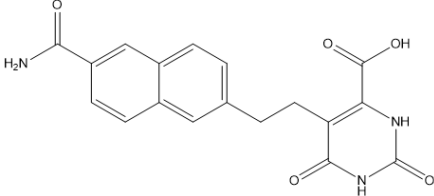
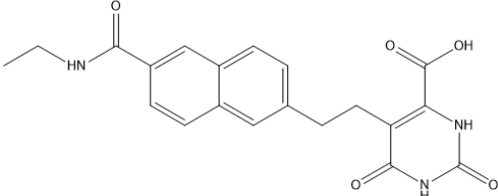
Compound	Structure	pIC <sub>50</sub>
14		5.71
15		5.75
16		5.79
17		5.79
18		5.85
19		5.92
20		5.95

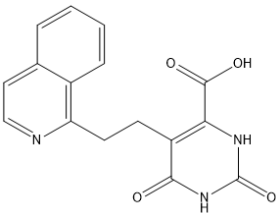
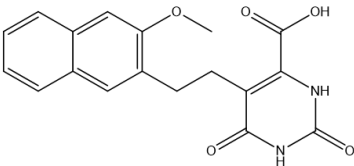
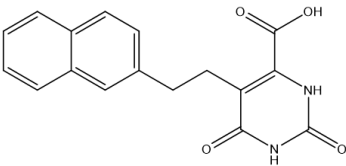
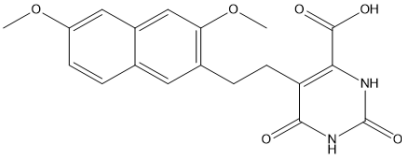
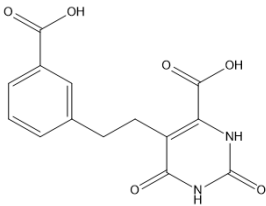
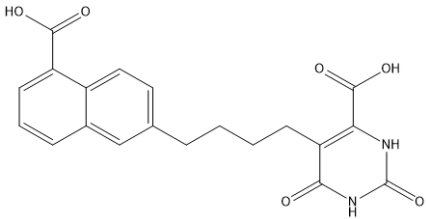
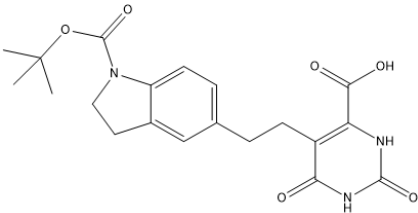
Compound	Structure	pIC <sub>50</sub>
21		5.96
22		5.97
23		5.99
24		6.05
25		6.05
26		6.07
27		6.09

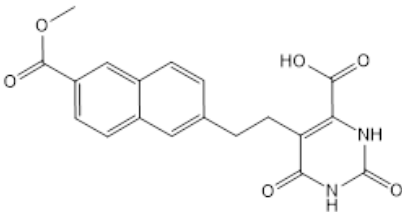
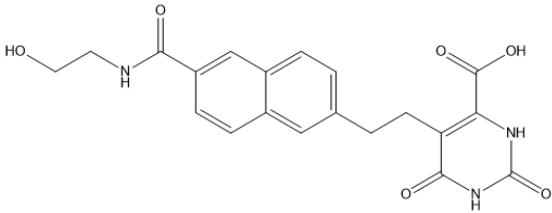
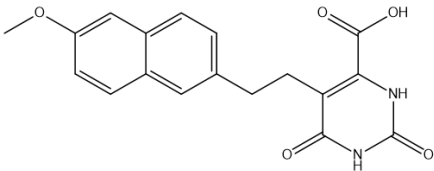
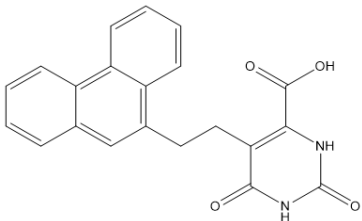
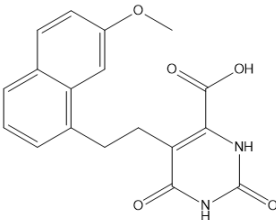
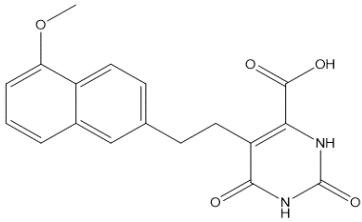
Compound	Structure	pIC <sub>50</sub>
28		6.10
29		6.10
30		6.12
31		6.12
32		6.13
33		6.13
34		6.15

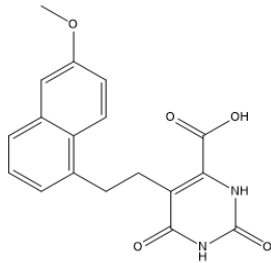
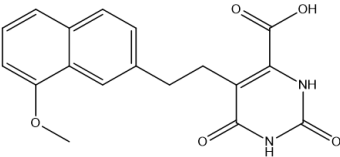
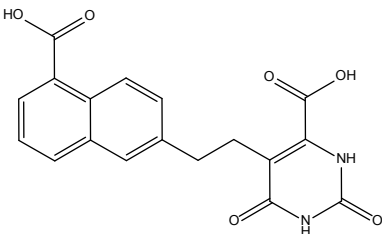


Compound	Structure	pIC <sub>50</sub>
35		6.21
36		6.24
37		6.25
38		6.30
39		6.32
40		6.37
41		6.43

Compound	Structure	pIC <sub>50</sub>
42		6.50
43		6.51
44		6.52
45		6.53
46		6.54
47		6.57
48		6.59

Compound	Structure	pIC <sub>50</sub>
49		6.61
50		6.65
51		6.65
52		6.69
53		6.70
54		6.71
55		6.74

Compound	Structure	pIC <sub>50</sub>
56		6.76
57		6.81
58		6.85
59		6.86
60		6.89
61		6.96

Compound	Structure	pIC <sub>50</sub>
62		6.97
63		7.34
64		7.48

## Molecular docking

The molecular docking studies were performed using the X-ray structure of DHODH (PDB 2E6A, 1.64 Å) and GOLD 5.3 (Inaoka *et al.*, 2008; Jones *et al.*, 1997). The structure of DHODH was prepared by adding hydrogen atoms and excluding solvent and the co-crystallized compound. Only those water molecules that were observed to be conserved across different DHODH X-ray structures were kept for the molecular docking runs (Zhong *et al.*, 2019). The active site was defined as a 12 Å-radius sphere centered on the Cys25 sulfur atom, given the proximity of this residue to orotate. All the binding site residues were checked for likely flipped orientations, tautomerism, and protonation states (ten Brink, Exner, 2009). The active site for molecular docking was restricted to the solvent-accessible surface by applying the cavity detection routine of GOLD 5.3. The molecular docking genetic algorithm was parameterized according to the following setup:

population size = 100; selection pressure = 1.1; number of operations = 100,000; number of islands = 5; niche size = 2; crossover and mutation frequencies = 95; and migration frequency = 10. The scoring function GoldScore was used to rank the predicted binding conformations, and the best-scoring conformation for each inhibitor was used in the 3D QSAR modeling (Medeiros *et al.*, 2021).

## 2D QSAR: hologram QSAR

The HQSAR models were initially developed using fragments containing 4 to 7 atoms and the default hologram lengths: 53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353, and 401 bins (Ferreira, Andricopulo, 2015). Atom type (A), bond type (B), hydrogen atoms (H), connectivity (C), hydrogen bond acceptor/donor (DA), and chirality (C) were used as fragment distinction criteria (Souza, Ferreira, Andricopulo, 2017). For the most statistically significant model, fragment sizes were varied as follows: 2–5, 3–6, 5–8, 6–9, 7–10.

### 3D QSAR: comparative molecular field analysis

The CoMFA steric and electrostatic interaction energies between the aligned compounds and the virtual probe, a sp<sup>3</sup> hybridized carbon atom, were computed using the default Coulomb and Lennard-Jones potentials of SYBYL-X 2.1, respectively (Medeiros *et al.*, 2021). The default CoMFA scaling and a cut-off value of 30 kcal/mol were used to generate the interaction fields (Clark, Cramer III, Van Opdenbosch, 1989). A 3D cubic lattice generated to encompass the dataset compounds were created using a grid spacing of 2 Å. The region-focusing procedure was applied to refine the models and increase the resolution of the CoMFA contour maps. A filtering value of 2 kcal/mol was applied to exclude from the analysis lattice points with negligible energy variations and thus improve the signal-to-noise ratio. The contour maps were built by computing the pairwise products between the standard deviations and the partial least squares (PLS) coefficients of the CoMFA columns (StDev\*Coeff) using the region-focusing method (Cramer, Patterson, Bunce, 1988). The contour maps were generated based on the grid points with a StDev\*Coeff above 80% or below 20% of the maximum computed value.

### QSAR validation procedures

The HQSAR and CoMFA models were generated using a Partial Least Squares (PLS) regression analysis on the training set compounds. The optimum number of components was determined using the leave-one-out (LOO) cross-validation procedure, which derived the cross-validated ( $q^2$ ) correlation coefficient. Then, the optimum number of components was used to determine the full non-cross-validated correlation coefficient ( $r^2$ ) on the entire training set. In addition, leave-many-out (LMO) cross-validation was applied to split the training set into 20 LMO groups, and, therefore, investigate the stability of the most statistically significant model in a context of

slightly greater exclusions of data. Progressive scrambling Y-randomization was used to estimate the susceptibility of the models to chance correlations (Clark, Fox, 2004). Only the training set was used in the production and internal validation of the QSAR models. Ultimately, the best models were evaluated for their external predictive ability by deriving the predictive correlation coefficient ( $r^2_{\text{pred}}$ ) using the test set.

Furthermore, an applicability domain (AD) analysis was conducted. The Organization for Economic Cooperation and Development (OECD) has established that the definition of the AD is required to evaluate the biological and chemical space within which a QSAR model can be applied. Considering the structural similarity of novel, untested compounds to the training set that was used to generate the model, the AD can assess the degree of uncertainty for the prediction of the dependent variable (Weaver, Gleeson, 2008). The AD method used in this study applies the leverage (or influence) versus the biological activity (Student residues) technique. The Chemoface program was used to calculate the AD for the best HQSAR and CoMFA models developed in this study (Nunes *et al.*, 2012).

## RESULTS AND DISCUSSION

### HQSAR models

The preliminary HQSAR models were developed using all available hologram lengths (53 to 401 bins) and fragment sizes varying from 4 to 7 atoms. The following combinations of fragment distinction parameters were used to generate the molecular holograms: A/B, A/B/C, A/B/C/H, A/C, A/H, A/DA, A/B/H, A/B/DA, A/B/H/DA, A/B/C/DA, A/B/C/H/DA, A/C/H e A/C/DA. The results for the LOO cross-validated and full QSAR models are shown in Table II. Additionally, Table II shows the  $r^2_{\text{pred}}$  values for the most statistically significant models, which demonstrate the predictive ability of these models for the test set compounds.

**TABLE II** - HQSAR results obtained for fragment sizes ranging from 4 to 7 atoms

Model	Fragment Distinction	$q^2$	$r^2$	SEE	HL	N	$r^2_{\text{pred}}$
1	A/B	0.558	0.844	0.356	83	7	0.720
2	A/B/C	0.500	0.888	0.375	356	3	-
3	A/B/C/H	0.380	0.658	0.300	307	3	-
4	A/C	0.592	0.878	0.342	71	7	0.801
5	A/H	0.469	0.824	0.395	53	8	-
6	A/DA	0.494	0.807	0.377	53	6	-
7	A/B/H	0.408	0.890	0.417	307	8	-
<b>8</b>	<b>A/B/DA</b>	<b>0.658</b>	<b>0.885</b>	<b>0.310</b>	<b>401</b>	<b>6</b>	<b>0.822</b>
9	A/B/H/DA	0.639	0.875	0.326	61	8	0.700
10	A/B/C/DA	0.542	0.881	0.363	59	7	0.791
11	A/B/C/H/DA	0.415	0.840	0.405	353	6	-
12	A/C/H	0.381	0.624	0.403	353	3	-
13	A/C/DA	0.602	0.883	0.334	353	6	0.772

A = atoms; B = bonds; C = connectivity; H = hydrogen atoms; DA = hydrogen bond donor/acceptor;  $q^2$  = leave-one-out (LOO) cross-validated correlation coefficient;  $r^2$  = non-cross validated correlation coefficient; SEE = standard error of estimate; HL = hologram length; N = optimal number of components;  $r^2_{\text{pred}}$  = predictive correlation coefficient.

Among the 13 HQSAR analyses, model 8 exhibited the most significant results ( $q^2 = 0.658$ ,  $r^2 = 0.885$ , and  $r^2_{\text{pred}} = 0.822$ ). Thus, while keeping the A/B/DA fragment distinction criteria, we evaluated the leverage of the fragment size on the statistical consistency of model 8.

The results in Table III show that varying the fragment size led to a decrease in the key statistical indicators ( $q^2$ ,  $r^2$  and  $r^2_{\text{pred}}$ ). Therefore, model 8 was used in the subsequent HQSAR analyses.

**TABLE III** - HQSAR models generated by using different fragment sizes and the A/B/DA fragment distinction criteria

Model	Fragment Size	$q^2$	$r^2$	SEE	HL	N	$r^2_{\text{pred}}$
14	2–5	0.519	0.892	0.376	356	8	-
15	3–6	0.616	0.861	0.05	401	6	0.786
16	5–8	0.640	0.870	0.05	257	9	0.803
17	6–9	0.630	0.875	0.05	353	6	0.799
18	7–10	0.580	0.870	0.04	199	10	-

A = atoms; B = bonds; C = connectivity; H = hydrogen atoms; DA = hydrogen bond donor/acceptor;  $q^2$  = leave-one-out (LOO) cross-validated correlation coefficient;  $r^2$  = non-cross validated correlation coefficient; SEE = standard error of estimate; HL = hologram length; N = optimal number of components;  $r^2_{\text{pred}}$  = predictive correlation coefficient.

Subsequent leave-many-out (LMO) cross-validation analyses confirmed the stability of model 8 against more significant data withdrawn. Randomly dividing the training set in 20 LMO groups, resulted in a HQSAR model with a  $q^2 = 0.656$  and a  $r^2 = 0.884$ . These results are very close to those found for the LOO-validated model, which shows that slightly greater exclusions of data during the PLS analyses did not impact the quality of the model. Next, to assess the susceptibility of model 8 to chance correlations, a Y-randomization test was conducted. The noise introduced in the model by the systematic perturbations in the dependent variable caused

an expected drop in the  $q^2$  value. The randomized model had a  $Q^2 = 0.55$ , which attests to the stability of the original HQSAR analysis against chance correlations. Stable models should, additionally, yield progressive scrambling effective slopes close to unity (Clark, Fox, 2004), which is the case for model 8 ( $dq^2/dr_{yy}^2 = 1.12$ ). Besides the procedures for internal validation, model 8 was assessed for its external predictive ability for the test set compounds, which is an essential requirement of solid QSAR models. Experimental and predicted  $\text{pIC}_{50}$  values for the complete dataset are listed in Table IV and depicted in Figure 3.

**TABLE IV** - Experimental and predicted  $\text{pIC}_{50}$  values for the final HQSAR and CoMFA models

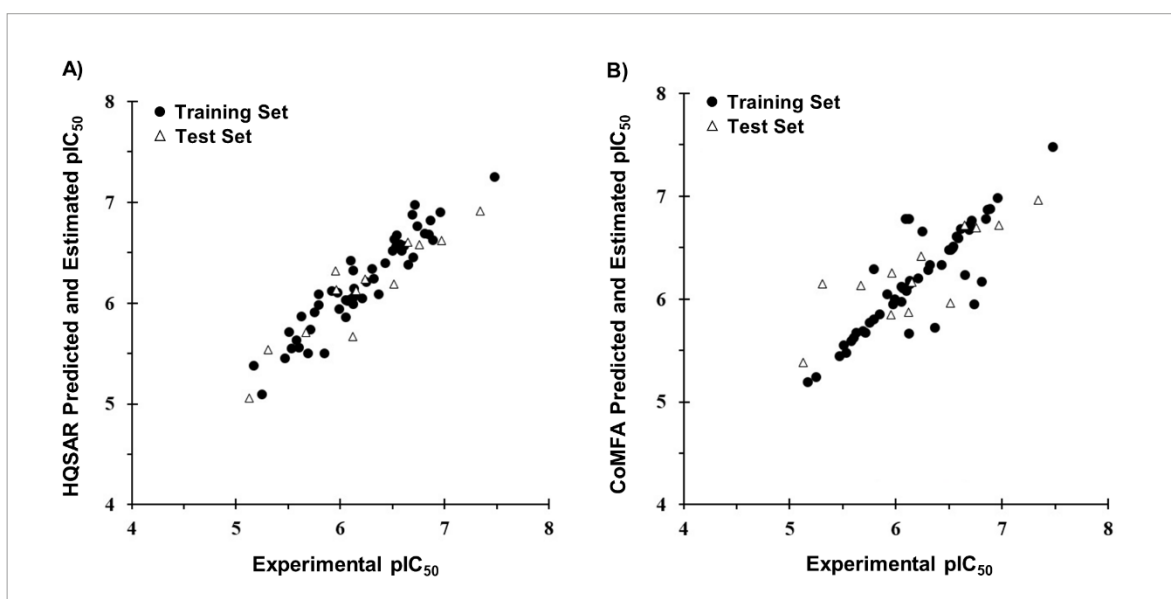
Inhibitor	Experimental	HQSAR		CoMFA	
		Predicted	Residual <sup>1</sup>	Predicted	Residual <sup>1</sup>
1*	5.13	5.06	0.07	5.38	-0.25
2	5.17	5.38	-0.21	5.19	-0.02
3	5.25	5.09	0.16	5.24	0.01
4*	5.31	5.54	-0.23	6.15	-0.84
5	5.47	5.45	0.02	5.44	0.03
6	5.51	5.71	-0.2	5.55	-0.04
7	5.53	5.55	-0.02	5.48	0.05
8	5.58	5.63	-0.05	5.59	-0.01
9	5.60	5.56	0.04	5.62	-0.02
10	5.63	5.87	-0.24	5.67	-0.04
11	6.12	5.99	0.13	5.66	0.46
12*	5.67	5.71	-0.04	6.13	-0.46
13	5.69	5.50	0.19	5.69	0.00
14	5.71	5.74	-0.03	5.67	0.04
15	5.75	5.91	-0.16	5.77	-0.02
16	5.79	5.98	-0.19	5.80	-0.01
17	5.79	6.09	-0.3	6.29	-0.5
18	5.85	5.50	0.35	5.85	0.00
19	5.92	6.12	-0.2	6.05	-0.13
20*	5.95	6.32	-0.37	5.85	0.10



Inhibitor	Experimental	HQ SAR		CoMFA	
		Predicted	Residual <sup>1</sup>	Predicted	Residual <sup>1</sup>
21*	5.96	6.13	-0.17	6.25	-0.29
22	5.97	6.10	-0.13	5.95	0.02
23	5.99	5.94	0.05	6.00	-0.01
24	6.05	5.86	0.19	5.97	0.08
25	6.05	6.03	0.02	6.12	-0.07
26	6.07	6.02	0.05	6.10	-0.03
27	6.09	6.02	0.07	6.78	-0.69
28	6.10	6.04	0.06	6.08	0.02
29	6.10	6.42	-0.32	6.10	0.00
30	6.12	6.32	-0.2	6.78	-0.66
31*	6.12	5.67	0.45	5.87	0.25
32	6.13	6.14	-0.01	6.18	-0.03
33	6.13	6.06	0.07	6.14	-0.01
34*	6.15	6.12	0.03	6.16	-0.01
35	6.21	6.05	0.16	6.20	0.01
36*	6.24	6.24	0	6.42	-0.18
37	6.25	6.21	0.04	6.66	-0.41
38	6.30	6.34	-0.04	6.28	0.02
39	6.32	6.24	0.08	6.33	-0.01
40	6.37	6.09	0.28	5.72	0.65
41	6.43	6.40	0.03	6.33	0.1
42	6.50	6.52	-0.02	6.48	0.02
43*	6.51	6.19	0.32	5.96	0.55
44	6.52	6.63	-0.11	6.48	0.04
45	6.53	6.56	-0.03	6.49	0.04
46	6.54	6.67	-0.13	6.51	0.03
47	6.57	6.58	-0.01	6.61	-0.04
48	6.59	6.52	0.07	6.59	0.00
49	6.61	6.57	0.04	6.68	-0.07
50*	6.65	6.60	0.05	6.72	-0.07
51	6.65	6.38	0.27	6.23	0.42
52	6.69	6.88	-0.19	6.67	0.02
53	6.70	6.45	0.25	6.72	-0.02
54	6.71	6.97	-0.26	6.76	-0.05

Inhibitor	Experimental	HQSAR		CoMFA	
		Predicted	Residual <sup>1</sup>	Predicted	Residual <sup>1</sup>
55	6.74	6.76	-0.02	5.95	0.79
56*	6.76	6.58	0.18	6.69	0.07
57	6.81	6.69	0.12	6.17	0.64
58	6.85	6.68	0.17	6.78	0.07
59	6.86	6.82	0.04	6.87	-0.01
60	6.89	6.62	0.27	6.88	0.01
61	6.96	6.90	0.06	6.98	-0.02
62*	6.97	6.62	0.35	6.72	0.25
63*	7.34	6.91	0.43	6.96	0.38
64	7.48	7.25	0.23	7.48	0.00

<sup>1</sup> Difference between experimental and predicted  $\text{pIC}_{50}$  values. \* Test set compounds.



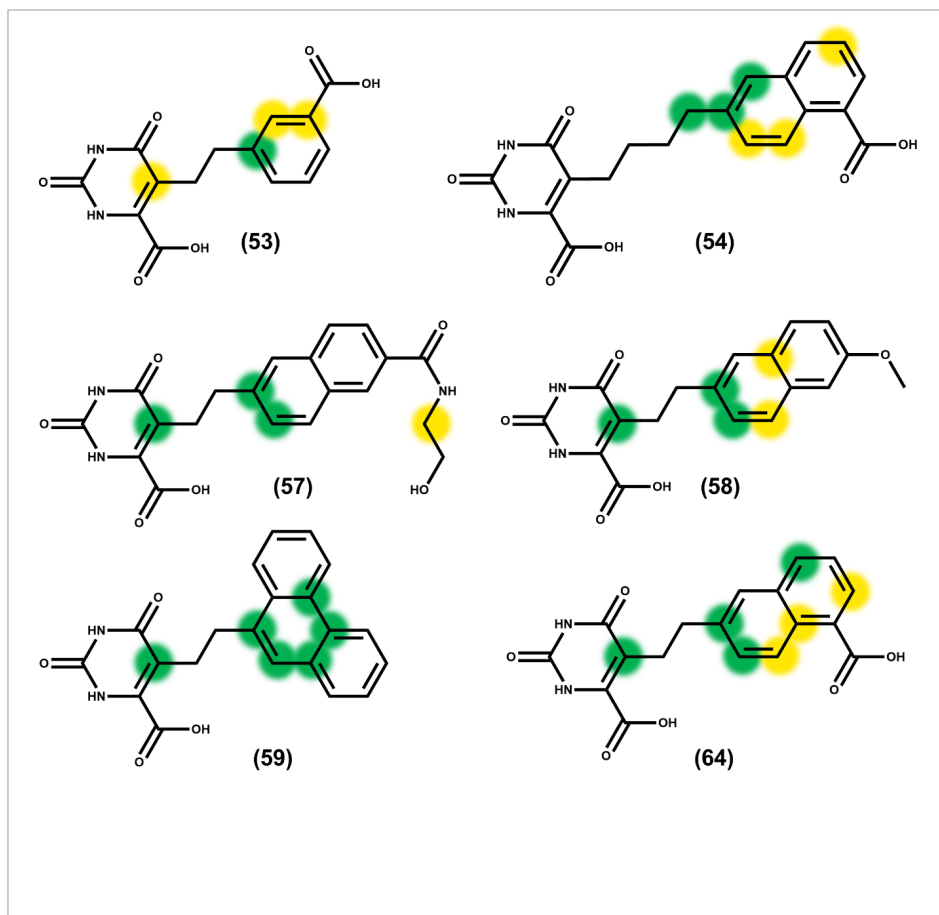
**FIGURE 3** - Experimental *versus* predicted and estimated  $\text{pIC}_{50}$  values for the best HQSAR (A) and CoMFA (B) models. Solid circles: training set; open triangles: test set.

Table IV and Figure 3 indicate a close agreement between the experimental and predicted  $\text{pIC}_{50}$  values for both training and test set compounds. These findings, along with an  $r^2_{\text{pred}}$  value of 0.822, demonstrate the good ability of the best HQSAR model to predict the activity of

new compounds that are structurally related to the dataset of DHODH inhibitors under investigation. Contribution maps generated by the best HQSAR model were used to assign negative and positive contributions of the molecular fragments to the biological activity. Positive contributions

are highlighted as yellow and green fragments. Otherwise, negative contributions to the biological activity are flagged as red, orange and red-orange fragments. White represents

neutral contributions. Figure 4 illustrates the positive contribution maps generated for some compounds.



**FIGURE 4** - HQSAR contribution maps for DHODH inhibitors used in the 2D QSAR modeling studies.

The HQSAR contribution maps indicate that potent inhibitors such as **53** ( $\text{pIC}_{50} = 6.70$ ), **54** ( $\text{pIC}_{50} = 6.71$ ), **57** ( $\text{pIC}_{50} = 6.81$ ), **58** ( $\text{pIC}_{50} = 6.85$ ), **59** ( $\text{pIC}_{50} = 6.86$ ), and **64** ( $\text{pIC}_{50} = 7.48$ ), have aromatic rings such as naphthalene, phenanthrene and benzene rings that stand out as key groups for biological activity. Also important is the flexibility allowed by the aliphatic spacer fragments between the two ring systems. The aliphatic spacer fragment was shown to be important to allow the optimal positioning of the two ring systems in the binding cavity pockets of the enzyme, as shown in the molecular docking

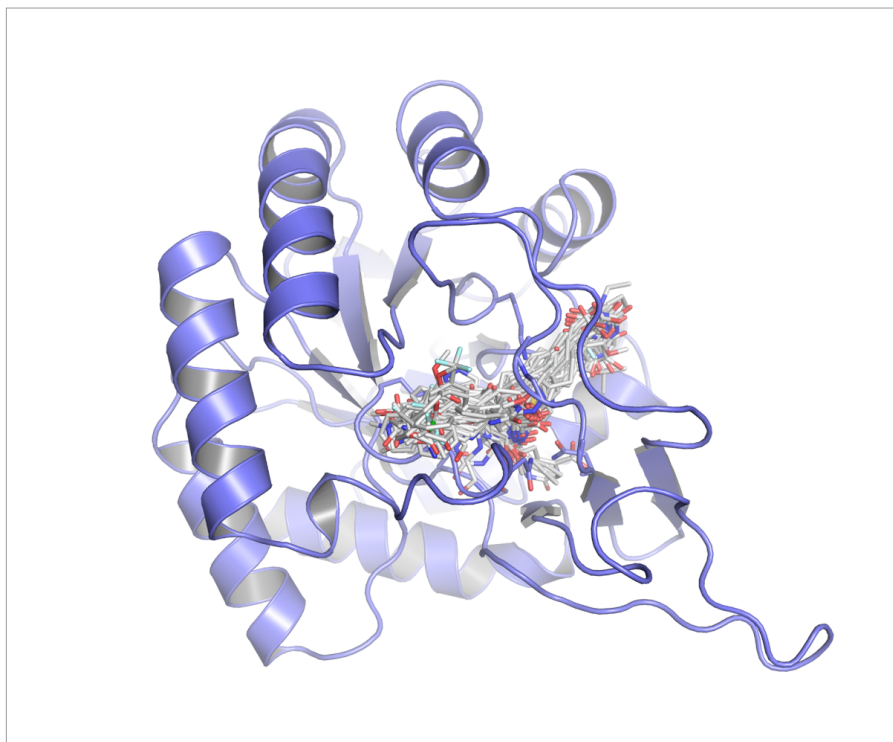
studies. The data reveals that molecules with a spacer made up of 2 carbon atoms yielded the best results in terms of DHODH inhibition.

### Molecular docking and structural alignment

The complete set of compounds was docked into the active site of DHODH to provide the structural alignment for the 3D QSAR CoMFA studies and to investigate the binding mode of the inhibitors. Each molecule was docked using the substrate of the enzyme, orotate, as a reference

for delimiting the binding site. Figure 5 depicts the dataset molecules aligned in the active site of DHODH. This alignment was used to calculate the stereochemical and

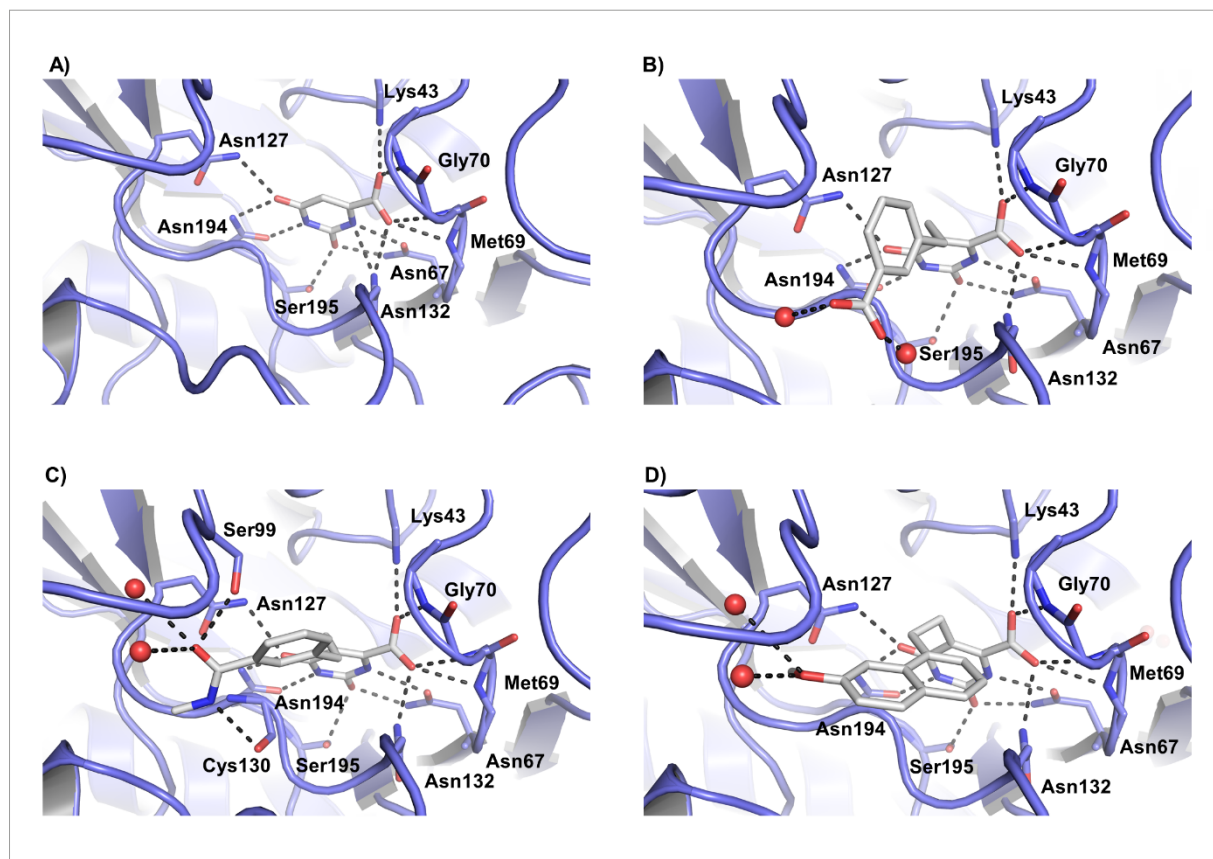
electrostatic fields during the development of the CoMFA models. Also, molecular docking studies were used to analyze the enzyme-inhibitor intermolecular interactions.



**FIGURE 5** - Inhibitors aligned in the active site of DHODH (PDB 2E6A, 1.64 Å). Protein structure is shown as a cartoon and inhibitors are depicted as sticks.

After the molecular docking, we observed that the dataset molecules interact with DHODH in a very similar way in the orotate binding cavity. The orotate ring of the inhibitors interacts with the predominantly polar binding site of the enzyme through an extensive hydrogen-bond network (Figure 6A), as observed in the X-ray structure of the orotate-DHODH complex. The molecular docking results shown herein demonstrate a highly conserved binding mode for the dioxo-tetrahydropyrimidine ring, indicating the key role played by this moiety for DHODH-inhibitor binding. Consistent with these findings, many of the amino acid residues involved in the intermolecular hydrogen-bond interactions are conserved; these include

Asn67, Asn127, Asn132, Asn194, Lys43, Gly70, Met69, and Ser195. It is worth mentioning that water molecules are also relevant for the binding of the inhibitors, mainly by hydrogen bonding with carboxylate, amide and hydroxyl groups, as shown in Figures 6B, 6C and 6D. Considering the highly polar active site of the enzyme due to the nature of the above-mentioned amino acid residues, it is worthwhile to consider the ligands' electron-rich  $\pi$  systems of the aromatic rings attached to the orotate moiety as important players in the interaction with DHODH.



**FIGURE 6** - (A) X-ray structure of DHODH from *Trypanosoma cruzi* in complex with orotate (PDB 2E6A, 1.64 Å). (B, C and D) Binding mode and intermolecular interactions for DHODH-inhibitor complexes obtained through molecular docking. (B) Compound 53. (C) Compound 10. (D) Compound 60. Protein structures are shown as a cartoon and substrate and inhibitors are depicted as sticks. Intermolecular interactions are depicted as dashed lines and water molecules are represented as spheres.

## CoMFA models

3D QSAR CoMFA studies were carried out to investigate the three-dimensional aspects that determine the biological activity of the DHODH inhibitors. The molecules as aligned in Figure 5 were used to calculate

the Lennard-Jones and Coulomb potentials for the CoMFA modeling. The Standard Deviation Coefficient (StDev\*Coeff) method was applied to exclude redundant descriptors and increase the weight of descriptors that contribute most significantly to the activity. The models are presented in Table V.

**TABLE V** - CoMFA models generated using different values of StDev\*Coeff

Model	StDev*Coeff	$q^2$	SEP	$r^2$	SEE	$r^2_{\text{pred}}$	N	S	E
19	-	0.039	0.517	0.683	0.285	-	3	0.495	0.505
20	0.3	0.483	0.341	0.997	0.031	-	10	0.618	0.382
21	0.6	0.75	0.275	0.994	0.042	0.66	10	0.615	0.385
22	0.9	0.705	0.281	0.891	0.171	0.62	5	0.654	0.346

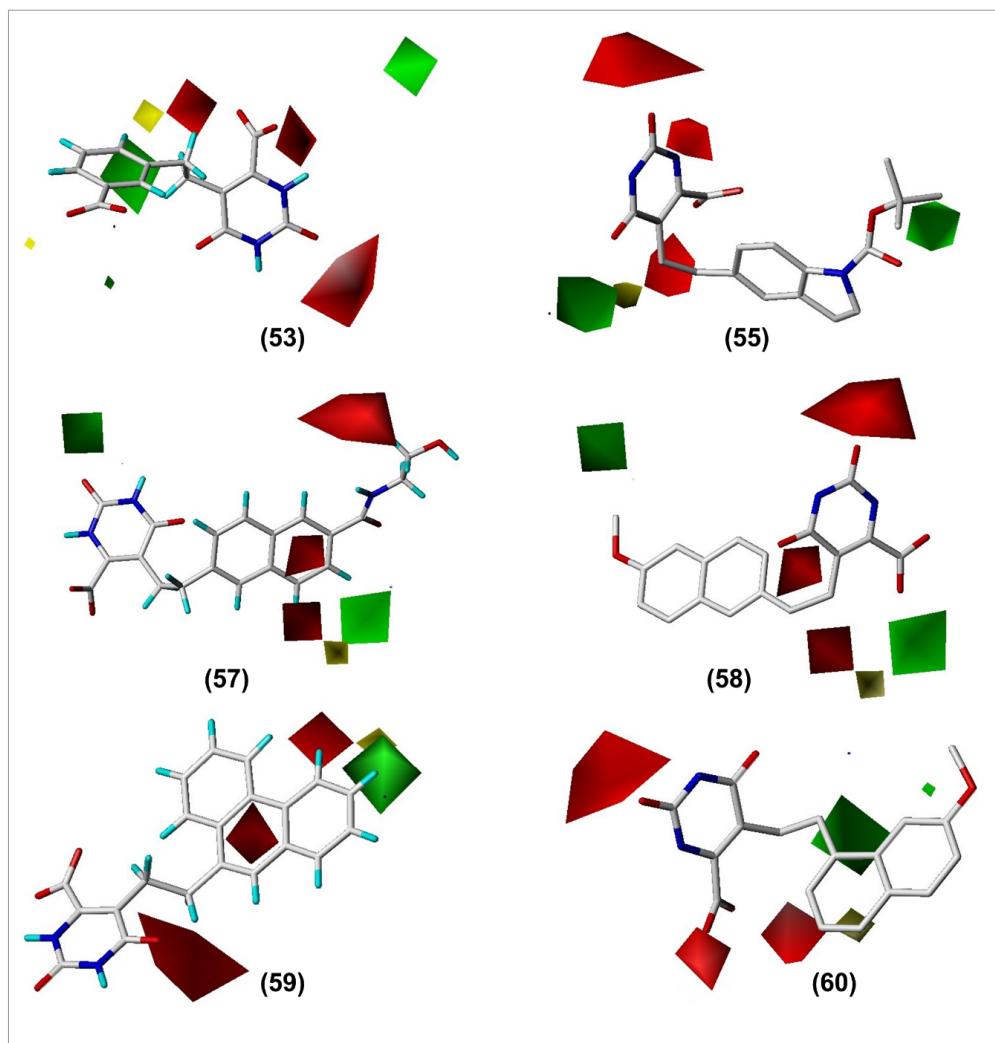
StDev\*Coeff = standard deviation coefficient;  $q^2$  = leave-one-out (LOO) cross-validated correlation coefficient;  $r^2$  = non-cross validated correlation coefficient; SEP = standard error of prediction; SEE = standard error of estimate; N = optimal number of components;  $r^2_{\text{pred}}$  = predictive correlation coefficient; S = fraction of the stereochemical field; E = fraction of the electrostatic field.

The results in Table V show that model 21, which was produced with a StDev\*Coeff value of 0.6, exhibits the highest internal ability as collectively shown by the  $q^2$  and  $r^2$  values. Additionally, the most important indicator, namely the predictive power of the model for the test set compounds ( $r^2_{\text{pred}} = 0.66$ ), is also superior in model 21. Stereochemical descriptors were most relevant than electrostatic variables upon generation of the model by the PLS routine. The experimental and predicted  $\text{pIC}_{50}$  values for model 21 are presented in Table IV and graphically depicted in Figure 3.

The agreement between the experimental and predicted  $\text{pIC}_{50}$  values for both the training and test sets demonstrate the internal statistical consistency of the model as well as its external predictive ability, i.e., for compounds that were not used in the construction of the model. The PLS StDev\*Coeff values were converted into 3D plots known as contour maps, which can be useful for the visual examination of the structure-activity relationships in a 3D environment. The contour maps point out regions in space where the electrostatic and stereochemical descriptors are the most relevant for the variation in the dependent variable (biological activity). Figure 7 depicts the contour maps for potent DHODH inhibitors belonging to the dataset. Green and yellow account for stereochemical fields and represent regions in the 3D environment around the molecules where the

increase and decrease of bulkiness are, respectively, favorable and unfavorable for the increase of the  $\text{pIC}_{50}$  values. Otherwise, red and blue polyhedral account for electrostatic fields, and indicate regions where the presence of electronegative groups, are, respectively, favorable and detrimental to the biological activity of the inhibitors.

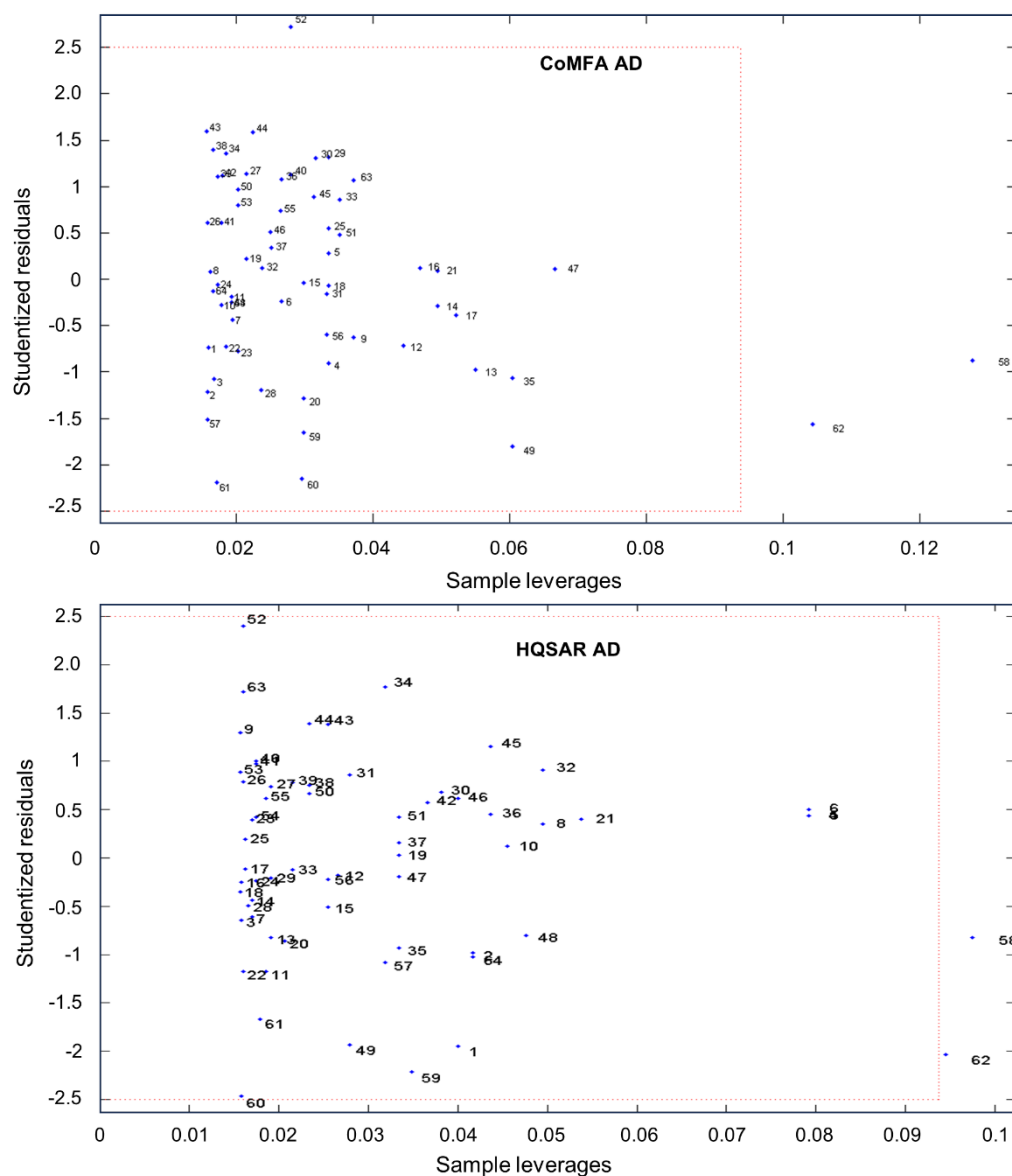
Figure 7 displays CoMFA contour maps for inhibitors **53**, **55**, **57**, **58**, **59** and **60**. The Lennard-Jones and Coulomb descriptors indicate the positive contribution of electronegative and bulky groups around the orotate ring. This is consistent with the binding mode of the substrate, which engages in several hydrogen bonds in the orotate binding pocket in the enzyme active site (Figure 6). Likewise, bulky cyclic groups such as the naphthalene, phenanthrene, benzene and the *N*-substituted indoline ring are highlighted as relevant in terms of the stereochemical variables, which indicate the relevance of these groups for DHODH inhibition. It is worth mentioning that the CoMFA results are consistent with the HQSAR contribution maps which also indicate the positive contributions of these bulky cyclic groups. Electrostatic contours also flag these electron-rich ring systems as relevant for biological activity. These findings agree with the polar nature of the enzyme binding site which contributes to a predominantly electrostatic-driven enzyme-inhibitor intermolecular recognition.



**FIGURE 7** - CoMFA contour maps for DHODH inhibitors.

The AD was calculated for the final HQSAR and CoMFA models. This tool is useful to identify structural and biological activity outlier compounds and determine the domain in which the models are more reliable and applicable (Gramatica, 2007). The technique uses leverage and Student residuals to detect these outlier compounds. The leverage is an indicator of the extent of the influence of a given compound on the developed QSAR model. The leverage of a molecule is the distance between this molecule and the center of the training set in a domain (or space) delimited by the independent variables, i.e., the

molecular descriptors. The other relevant AD indicator, the Student residues, are values defined as units of the standard deviation from the mean activity value. Student residues higher than  $\pm 2.5$  are considered the usual statistical threshold to determine activity outliers. On the other side, molecules that present a high influence on the training stage of model construction are called structural outliers. Figure 8 indicates that compounds 58 and 62 have high leverage values. These molecules, therefore, exerted a strong influence on the generation of both, 2D and 3D QSAR models.



**FIGURE 8** - Applicability domain for the final QSAR models.

QSAR methods are widely used in medicinal chemistry and many different algorithms and studies have been published today in the field of NTDs. A CoMFA and CoMSIA investigation were developed for a set of *T. cruzi* cruzain inhibitors. Correlation coefficients such as an  $r^2_{\text{pred}}$  value of 0.72 in the best CoMFA model and an  $r^2_{\text{pred}}$  value of 0.97 for the best CoMSIA analysis demonstrate the statistical robustness of these models. A study on Chagas disease describes 2D HQSAR and 3D CoMFA

and CoMSIA investigations on potent compounds bearing a benzimidazole core. Statistically sound models were obtained, and the predictive power of the analyses was demonstrated by an  $r^2_{\text{pred}} = 0.65$  for HQSAR, 0.94 for CoMFA, and 0.82 for Comparative Molecular Similarity Index Analysis (CoMSIA) (Pauli *et al.*, 2017). Other methodologies such as multiple linear regression (MLR) and best-first algorithm have been used to develop QSAR models for *T. cruzi* targets. One such an example is an



investigation on semicarbazones that resulted in solid models with the best one yielding an  $r^2_{\text{ext}} = 0.79$  (Scotti *et al.*, 2016). Another worth-mentioning analysis of 62 aryl carboxylic acid amide derivatives as DHODH inhibitors yielded statistically sound QSAR models (Vyas, Ghate, 2013). Methods such as MLR ( $r^2 = 0.85$  and  $q^2 = 0.79$ ), principal component regression (PCR,  $r^2 = 0.71$  and  $q^2 = 0.66$ ), and PLS ( $r^2 = 0.84$  and  $q^2 = 0.80$ ) were used, and comparable results were obtained considering our study.

Literature on QSAR modeling shows that a diversity of QSAR methods and strategies are available for the medicinal chemistry research community. These techniques can produce solid models in terms of both statistical indicators and the chemistry underlying the activity of the investigated compounds. The models described herein have statistical robustness very similar to published data, and more importantly, they serve the purpose of any QSAR model, that is, to provide predictive power and chemical rationale for the activity under investigation.

It should be noted that every QSAR model has its limitations. Nevertheless, the analyses reported here can predict the biological activity of the compounds in the test set and point out characteristics that are key for the inhibition of the target enzyme. Methods that are based on the bi-dimensional structure of the data set molecules, such as HQSAR, are independent of a non-trivial procedure, namely, the structural alignment. This aspect is a major advantage of these methods and can deliver powerful analyses in relatively short timelines (Kleandrova, Speck-Planche, 2020). Although bidimensional QSAR techniques are more straightforward, they have as their main limitation the inability to detect 3D features, such as steric and electrostatic fields and shapes, which are important determinants of biological activity. This is the reason why QSAR methods are more valuable when 2D and 3D approaches are used in a complementary way on the same data set.

In turn, CoMFA is a 3D approach that provides insights related to ligand-receptor interactions because molecular fields and shapes are considered in the analysis. The accounting of Lennard-Jones and Coulomb potentials

allows insights into which intermolecular interactions are more likely to occur when the ligand-receptor complex is forming. This is particularly useful if the three-dimensional atomic coordinates of the molecular target are available. In these situations, the results of the QSAR analysis can contribute to a better understanding of the interaction between the ligand and the molecular target. However, the need for a molecular alignment is a relevant drawback of 3D QSAR methods and this can be a serious obstacle for the construction of meaningful models. However, a major advantage of the models described in this work is that the structure of the molecular target is known, which renders the procedure for molecular alignment more rational in terms of the binding mode of the data set compounds.

In short, this study provided useful insights into the molecular features, both 2D and 3D, that determine the inhibition of the enzyme DHODH by a series of orotate-based inhibitors. Additionally, the developed QSAR models disclosed important factors for the intermolecular interaction between the inhibitors and the enzyme. These findings can enable the design of novel DHODH inhibitors with improved potency, thus contributing to drug research and development (R&D) for Chagas disease.

## CONCLUSION

Predictive and statistically sound 2D and 3D QSAR models for a set of orotate derivatives as inhibitors of *T. cruzi* DHODH were developed. DHODH is essential for the life cycle of the parasite and has been validated as a molecular target for drug discovery studies in Chagas disease. Among the main results of this study, one can highlight the following: the models showed internal consistency for the training set and revealed to be stable against chance correlations and withdrawals of data. Additionally, the final HQSAR and CoMFA models demonstrated to be highly predictive for the test set compounds, which is the most significant indicator for their predictive power. Additionally, the contribution and contour maps flagged the 2D fragments and 3D features most closely related to the biological activity of the dataset inhibitors. Furthermore, the graphical results and

output of the HQSAR and CoMFA models highlighted the aromatic ring systems and the polar substituents attached to these rings, and the orotate moiety as the main drivers for DHODH inhibition. Based on the X-ray structure of orotate, molecular docking studies were conducted and, as a result, the main elements for the enzyme-inhibitor interactions were identified. The findings reported herein indicate a very conserved binding mode of the orotate ring, which is mainly stabilized by an extensive hydrogen-bonding network. A set of amino acid residues were shown to be key for the stabilization not only for orotate but for the investigated inhibitors as well. It is important to note that the requirement for the inhibitors to be structurally aligned is a relevant drawback of 3D QSAR methods such as CoMFA. This can be a relevant hurdle in 3D QSAR modeling. Regarding HQSAR, the method does not consider 3D features such as molecular fields and shapes, which can be addressed as a limitation of the technique. The QSAR models and the investigation of the binding mode of this series of DHODH inhibitors uncovered the main features that underline their biological activity. This provides valuable information for further medicinal chemistry studies in Chagas disease drug discovery.

## ACKNOWLEDGMENTS

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brazil, CIBFar-CEPID (2013/07600-3 and 2013/25658-9).

## REFERENCES

- Arakaki TL, Buckner FS, Gillespie JR, Malmquist NA, Phillips MA, Kalyuzhnyi O, et al. Characterization of trypanosoma brucei dihydroorotate dehydrogenase as a possible drug target; structural, kinetic and RNAi studies. *Mol Microbiol.* 2008;68(1):37-50.
- Arnal A, Waleckx E, Rico-Chávez O, Herrera C, Dumonteil E. Estimating the current burden of chagas disease in Mexico: a systematic review and meta-analysis of epidemiological surveys from 2006 to 2017. *PLoS Negl Trop Dis.* 2019;13(4):e0006859.
- Azeredo LFSP, Coutinho JP, Jabor VAP, Feliciano PR, Nonato MC, Kaiser CR, et al. Evaluation of 7-arylamino pyrazolo[1,5-a]pyrimidines as anti-Plasmodium falciparum, antimalarial, and Pf-dihydroorotate dehydrogenase inhibitors. *Eur J Med Chem.* 2017;126:72-83.
- Bender A, Jenkins JL, Scheiber J, Sukuru SC, Glick M, Davies JW. How similar are similarity searching methods? A principal component analysis of molecular descriptor space. *J Chem Inf Model.* 2009;49(1):108-119.
- ten Brink T, Exner TE. Influence of protonation, tautomeric, and stereoisomeric states on protein-ligand docking results. *J Chem Inf Model.* 2009;49(6):1535-46.
- Cereto-Massagué A, Ojeda MJ, Valls C, Mulero M, Garcia-Vallvé S, Pujadas G. Molecular fingerprint similarity search in virtual screening. *Methods.* 2015;71:58-63.
- Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, et al. QSAR modeling: where have you been? Where are you going to? *J Med Chem.* 2014;57(12):4977-5010.
- Chibli LA, Schmidt TJ, Nonato MC, Calil FA, Da Costa FB. Natural products as inhibitors of Leishmania major dihydroorotate dehydrogenase. *Eur J Med Chem.* 2018;157:852-866.
- Clark M, Cramer III RD, Van Opdenbosch N. Validation of the general purpose tripos 5.2 force field. *J Comput Chem.* 1989;10(8):982-1012.
- Clark RD, Fox PC. Statistical variation in progressive scrambling. *J Comput Aided Mol Des.* 2004;18(7-9):563-76.
- Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc.* 1988;110(18):5959-67.

- Ferreira LLG, Andricopulo AD. Drugs and vaccines in the 21st century for neglected diseases. *Lancet Infect Dis.* 2019;19(2):125-127.
- Ferreira LG, Andricopulo AD. Fragment-Based QSAR and Structural Analysis of a Series of Hydroxyethylamine Derivatives as HIV-1 Protease Inhibitors. *Comb Chem High Throughput Screen.* 2015;18(5):464-75.
- Ferreira LLG, de Moraes J, Andricopulo AD. Approaches to advance drug discovery for neglected tropical diseases. *Drug Discov Today.* 2022;27(8):2278-2287.
- Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. *Tetrahedron.* 1980;36(22):3219-3228.
- GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Lancet.* 2018;392(10159):1859-1922.
- Ghasemi JB, Meftahi N, Pirhadi S, Tavakoli H. Docking and pharmacophore-based alignment comparative molecular field analysis three-dimensional quantitative structure-activity relationship analysis of dihydrofolate reductase inhibitors by linear and nonlinear calibration methods. *J Chemometrics.* 2013;27(10):287-296.
- Gramatica P. Principles of QSAR models validation: Internal and external. *QSAR Comb Sci.* 2007;26:694-701.
- Hashimoto M, Morales J, Fukai Y, Suzuki S, Takamiya S, Tsubouchi A, et al. Critical importance of the de novo pyrimidine biosynthesis pathway for trypanosoma cruzi growth in the mammalian host cell cytoplasm. *Biochem Biophys Res Commun.* 2012;417(3):1002-1006.
- Inaoka DK, Iida M, Hashimoto S, Tabuchi T, Kuranaga T, Balogun EO, et al. Design and synthesis of potent substrate-based inhibitors of the trypanosoma cruzi dihydroorotate dehydrogenase. *Bioorg Med Chem.* 2017;25(4):1465-1470.
- Inaoka DK, Sakamoto K, Shimizu H, Shiba T, Kurisu G, Nara T, et al. Structures of Trypanosoma cruzi dihydroorotate dehydrogenase complexed with substrates and products: atomic resolution insights into mechanisms of dihydroorotate oxidation and fumarate reduction. *Biochemistry.* 2008;47(41):10881-91.
- Irish A, Whitman JD, Clark EH, Marcus R, Bern C. Updated estimates and mapping for prevalence of chagas disease among adults, united states. *Emerg Infect Dis.* 2022;28(7):1313-1320.
- Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol.* 1997;267(3):727-748.
- Kleandrova VV, Speck-Planche A. The QSAR Paradigm in Fragment-Based Drug Discovery: From the Virtual Generation of Target Inhibitors to Multi-Scale Modeling. *Mini Rev Med Chem.* 2020;20(14):1357-1374.
- Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. *J Comput Aided Mol Des.* 2011;25(1):13-9.
- Llanos-Cuentas A, Casapia M, Chuquiyauri R, Hinojosa JC, Kerr N, Rosario M, et al. Antimalarial activity of single-dose DSM265, a novel plasmodium dihydroorotate dehydrogenase inhibitor. in patients with uncomplicated Plasmodium falciparum or Plasmodium vivax malaria infection: a proof-of-concept, open-label, phase 2a study. *Lancet Infect Dis.* 2018;18(8):874-883.
- Medeiros AR, Ferreira LLG, de Souza ML, de Oliveira Rezende Junior C, Espinoza-Chávez RM, Dias LC, et al. Chemoinformatics studies on a series of imidazoles as cruzain inhibitors. *Biomolecules* 2021;11(4):579.
- Nunes CA, Freitas MP, Pinheiro ACM, Bastos SC. Chemoface: A novel free user-friendly interface for chemometrics. *J Braz Chem Soc.* 2012;23:2003-2010.
- Pauli I, Ferreira LG, de Souza ML, Oliva G, Ferreira RS, Dessoy MA, et al. Molecular modeling and structure-activity relationships for a series of benzimidazole

derivatives as cruzain inhibitors. *Future Med Chem.* 2017;9(7):641-657.

Pérez-Molina JA, Crespillo-Andújar C, Bosch-Nicolau P, Molina I, et al. Trypanocidal treatment of chagas disease. *Enferm Infecc Microbiol Clin (Engl Ed).* 2021;39(9):458-470.

Pinheiro MP, Emery Fda S, Nonato MC. Target sites for the design of anti-trypanosomatid drugs based on the structure of dihydroorotate dehydrogenase. *Curr Pharm Des.* 2013;19(14):2615-2627.

Powell MJD. Restart procedures for conjugate gradient method. *Math Program.* 1977;12:241-254.

Salum LB, Andricopulo AD. Fragment-based QSAR strategies in drug design. *Expert Opin Drug Discov.* 2010;5(5):405-12.

Salum LB, Andricopulo AD, Honório KM. A fragment-based approach for ligand binding affinity and selectivity for the liver X receptor beta. *J Mol Graph Model.* 2012;32:19-31.

Scotti MT, Scotti L, Ishiki HM, Peron LM, de Rezende L, do Amaral AT. Variable-selection approaches to generate QSAR models for a set of antichagasic semicarbazones and analogues. *Chemometr Intell Lab Syst.* 2016;154:137-149.

Singh A, Maqbool M, Mobashir M, Hoda N. Dihydroorotate dehydrogenase: a drug target for the development of antimalarials. *Eur J Med Chem.* 2017;125:640-651.

Souza AS, Ferreira LG, Andricopulo AD. 2D and 3D QSAR studies on a series of antichagasic fenarimol derivatives. *IJQSPR.* 2017;2(1):44-63.

Thillainayagam M, Malathi K, Ramaiah S. In-Silico molecular docking and simulation studies on novel

chalcone and flavone hybrid derivatives with 1,2,3-triazole linkage as vital inhibitors of *Plasmodium falciparum* dihydroorotate dehydrogenase. *J Biomol Struct Dyn.* 2018;36(15):3993-4009.

Vyas VK, Ghatge M. QSAR study on a series of aryl carboxylic acid amide derivatives as potential inhibitors of dihydroorotate dehydrogenase (DHODH). *Med Chem.* 2013;9(2):222-39.

Wang T, Yuan XS, Wu MB, Lin JP, Yang LR. The advancement of multidimensional QSAR for novel drug discovery - where are we headed? *Expert Opin Drug Discov.* 2017;12(8):769-784.

Weaver S, Gleeson MP. The importance of the domain of applicability in QSAR modeling. *J Mol Graph Model.* 2008;26:1315-1326.

Willett P. Similarity-based virtual screening using 2D fingerprints. *Drug Discov Today.* 2006;11:1046-1053.

World Health Organization. WHO. Chagas disease (also known as American trypanosomiasis). [cited 2024 Feb 5]. Available from: [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)).

World Health Organization. WHO. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. [cited 2024 Feb 5]. Available from: <https://www.who.int/publications/i/item/9789240010352>.

Zhong H, Wang Z, Wang X, Liu H, Li D, Liu H, et al. Importance of a crystalline water network in docking-based virtual screening: a case study of BRD4. *Phys Chem Chem Phys.* 2019;21(45):25276-25289.

Received for publication on 08<sup>th</sup> February 2024

Accepted for publication on 10<sup>th</sup> May 2024

Associated Editor: Guilherme Martins Gelfuso