ISBN: 978-628-95207-8-1. ISSN: 2414-6390. Digital Object Identifier: 10.18687/LACCEI2024.1.1.1785

2D-QSAR Study of Thiazole derivatives as 5-Lipoxygenase inhibitors

Ariza-Rúa, Danilo L, MSc[1]; Hurtado Marquez, Julio Seferino, MSc[2]; Marbello-Peña, Humberto, MSc[3]; Edisson Chavarro-Mesa, Ph. D[4]; Liliana Carranza-López, Ph. D[5]
[1-4]Universidad Tecnológica de Bolívar, Colombia. dariza@utb.edu.co, jhurtado@utb.edu.co, hmarbello@utb.edu.co, echavarro@utb.edu.co. [5]Universidad Libre - Seccional Barranquilla, lilianap.carranzal@unilibre.edu.co

***Abstract* – Inhibition of 5-lipoxygenase (5-LO) has become a rational approach for the development of anti-inflammatory drugs. The aim of this study was to work on a trendy machine learning approach using an open-source data analysis Python script to discover 5-lipoxygenase inhibitors (5-LOX) by building two-dimensional quantitative structure-activity relationships (2D-QSAR) of a series of 59 thiazole derivatives that act as inhibitors of 5-LOX. The generated 2D QSAR model showed a good correlation coefficient of 0.626 and a good test set prediction coefficient of 0.621. The predictive ability of the 2D QSAR models was evaluated externally (test set with 12 compounds). The proposed model provided significant statistical quality.**

Keywords—2D-QSAR models, Thiazole derivatives, 5-Lipoxygenase, anti-inflammatory drugs.

I. Introduction

Inflammation is a normal response of the body to injury, infection, or irritation, and plays a critical role in the healing process. However, when inflammation becomes chronic it can contribute to the development of various diseases. Inflammation involves a complex interplay of various molecular and cellular components, and in the development of anti-inflammatory drugs, studies focus on several key objectives. Some important targets to inhibit inflammation are: cyclooxygenase (COX), tumor necrosis factor-alpha (TNF-α). Interleukins (IL), nuclear factor-kappa B (NF-κB), Janus kinases (JAK), phosphodiesterase (PDE) and lipoxygenases (LO).

5-Lipoxygenase (5-LOX) is a key enzyme in the regulation of fundamental biological processes and play a crucial role in the synthesis of specialized lipid mediators known as leukotrienes.[1] This enzyme, which belongs to the lipoxygenase family, has become the focus of increasing scientific attention due to its intrinsic association with the inflammatory response and cellular homeostasis. 5-LOX catalyzes the conversion of polyunsaturated fatty acids, particularly arachidonic acid, into a range of bioactive products that trigger diverse cellular responses, from enhancing inflammation to modulating cell proliferation and apoptosis.[2] Overexpression of 5-LOX is associated with various chronic inflammatory diseases such as rheumatoid arthritis, atherosclerosis, asthma, inflammatory bowel disease, lupus and certain cancers. [3-9]. Therefore, the development of specific inhibitors of this enzyme represents an active area in the search for new therapeutic strategies to control inflammation and treat various pathologies. These inhibitors could provide a more precise and specific approach to modulate 5-LOX activity, without other essential cellular functions to negatively influence.[10].

In the field of pharmaceutical research, the constant search for effective bioactive compounds has led to increasing attention to thiazole derivatives. These compounds are characterized by their five-membered heterocyclic structure with a sulfur and nitrogen atom and have proven to be versatile and promising building blocks in the development of biologically active agents. Its presence in various approved drugs supports its potential to modulate various biological pathways and provides a solid platform for the synthesis of new compounds with improved therapeutic properties. Drugs that contain thiazoles in their structure include thiazolidinediones (TZDs), which are used in the treatment of type 2 diabetes [11]; Thiazides used in the treatment of hypertension and heart failure[12]; dasatinib, dabrafenib, ixabepilone, patellamide A, and epothilone applied as anticancer drugs[13]. Some thiazoles and their derivatives exhibit several biologically significant biological activities, including antibacterial, antiprotozoal, antitubercular, as well as antifungal, anthelmintic, anti-inflammatory, and analgesic agents [14-16].

There are several techniques that are part of the computer-aided drug design process, such as quantitative structure-activity relationship (QSAR). A QSAR model is a mathematical model that correlates the chemical structure of a molecule with its biological activity, physicochemical properties, or other relevant endpoints. The purpose of QSAR modeling is to predict or explain the activity of a compound based on its structural features.[17]

In QSAR modeling, various molecular descriptors are calculated from the chemical structure of compounds. These descriptors are numerical representations of the chemical structure of the compounds and can include properties such as molecular weight, size, shape, electronic properties and structural features. Statistical or machine learning techniques are then applied to correlate these descriptors with the observed biological or chemical activities of the compounds. The resulting QSAR model can be used to predict the activity of new or untested compounds based solely on their molecular structure, which is particularly valuable for drug discovery, toxicology, and environmental risk assessment. QSAR models are widely used in pharmaceutical research to prioritize

compounds for further experimental testing, optimize chemical synthesis, and design new compounds with desired properties. In this work, we use the in silico QSAR technique to build 2D QSAR models using linear regression machine learning algorithms implemented in Python.

II. Material and Methods

A. *Data Collection*

To study the 2D-QSAR model, a series of 59 thiazole derivatives with inhibitory activity toward 5-LO was used[18-21]. The structures and biological activity of the thiazole derivatives are listed in Table 1. To eliminate the possibility of errors in data representation and reproducibility, all experimental activity values IC50(µm) were converted to pIC50 (pIC50 = $-\log_{10}$ (IC50)). The structures of the compounds were drawn using the freeware ACD/Chemsketch and saved in mol file format. The structures were then optimized using the energy minimization software Avogadro version 1.2.0. The resulting minimized structures were used as inputs for computing descriptors.

B. *Molecular descriptors calculation*

The molecular descriptors were calculated using the PaDEL software integrated in ChemDes [22]. The resulting descriptors were consolidated into a single CSV file. Subsequently, the CSV file containing both the molecular descriptors and the numerical pIC50 values was uploaded to Weka for attribute selection. In Weka, attribute selection refers to the process of selecting a relevant subset of features (attributes) from a data set to build a machine learning model. This process is important because not all features can contribute significantly to the prediction and some may even introduce noise into the model. Attribute selection in Weka is a crucial process for improving the efficiency and accuracy of machine learning models by removing irrelevant or redundant features. In this work, out of 1875 molecular descriptors, only 10 molecular descriptors remained, and the remaining invariant descriptors were removed using the attribute evaluator CfsSubsetEval and the BestFirst search method implemented in Weka.

C. *Construction of the model*

The 2D-QSAR models were generated using the multiple linear regression (MLR) method. For this purpose, the dataset, consisting of 59 compounds with 10 attributes each, was divided into training and test sets using Weka's filter functions. The filter generated a random subsample of the data set, with the sample size specified as a percentage. In this case, 80% of the data, corresponding to 47 compounds, was assigned to the training set, while the remaining 20%, representing 12 of the total 59 compounds, formed the test set.

D. *Validation of the model*

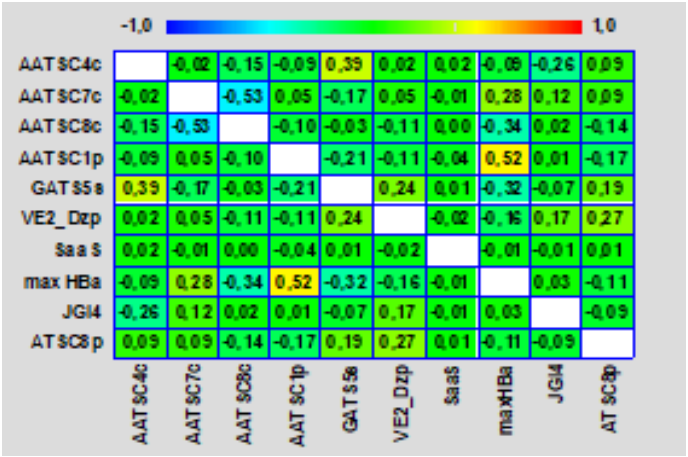
The predictive ability of the 2D QSAR models was evaluated externally (test set with 12 compounds). The quality of the model is guaranteed if the results comply with the global QSAR standard. This means that an R-squared value (R^2) is greater than 0.6, a predicted R-squared value (R^2pred) is greater than 0.5, a cross-validated correlation coefficient (Q^2) is greater than 0.6, and a p-value (95%) is less than 0.05, a high F-test value and low values for both R^2 Random and Q^2 Random.[23]

III. Results

The main molecular descriptors identified were: AATSC4c, AATSC7c, AATSC8c, AATSC1p, GATS5s, VE2_Dzp, SaaS, maxHBa, JGI4 and ATSC8p. Their values are listed in Table 2.

The Pearson product-moment correlation coefficient shows the degree of independence between the selected descriptors. Table 3.

Table 3. Pearson product-moment correlation coefficient of selected descriptors



The QSAR model obtained by multiple linear regression associates the selected molecular descriptors with the biological activity PIc50 of the molecules, as shown in the following mathematical formula:

pIC50 = -0.1207 *AATSC4c -0.2878 *AATSC8c + 0.3065 *AATSC1p - 0.0665 *GATS5s - 0.1048 *VE2_Dzp +0.1411 *SaaS + 0.1325 *maxHBa + 0.2375*JGI4 - 0.061 *ATSC8p - 6.2479.

The obtained model has satisfactory statistical properties. R^2 coefficient of determination of 0.626 and an RMSE of 0.488

IV. Discussion

Linear regression is a statistical method for modeling the relationship between a dependent variable and one or more independent variables. For linear regression, the scikit-learn library (sklearn) and its specific linear regression module (sklearn.linear_model.LinearRegression) are used. This library uses efficient and optimized implementations of algorithms, that can be faster and less error-prone than manual implementation. Sklearn provides built-in functions for

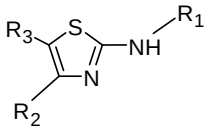
evaluating model performance, such as the mean_squared_error function, which simplifies model evaluation.

The MLR module of sklearn was used to predict the test set. The results showed that the performance of the model can be assessed using an R^2 coefficient of determination of 0.621 and an RMSE of 0.397. These values indicate a high level of accuracy of the model predictions as they closely match the observed values. Furthermore, the explanation derived from the original data set for the model remains consistent.

Figure 1 illustrates the behavior of the model with the prediction and Figure 2 illustrates the linear correlation between the observed pIC50 values and the predicted PIC50 values for the inhibitory activity of 5-LOx. The two data sets, with the test set shown in blue and the training set shown in red. The observed and predicted activity pIC50 values for the training set and test set compounds obtained using the Multiple Linear Regression tool are presented in Table 4.

The Multiple Linear Regression (MLR) model successfully predicted the activity pIC50 values for both the training and test set compounds. The observed and predicted values showed satisfactory agreement, suggesting that the model effectively captured the underlying relationships between the molecular descriptors and biological activity. This suggests the potential utility of the MLR model in predicting the activity of compounds not included in the training set, thereby facilitating the discovery and development of new bioactive molecules. However, further validation and refinement may be required to improve the predictive performance and generalizability of the model across different chemical spaces.

Table 1. Structure, and biological activity of thiazole derivatives.



C.no.	R1	R2	R3	Ic50	PIc50	Set
1	4-hydroxiphenyl	4-ethoxyphenyl	H	0.76	6.11919	Training
2	4-hydroxiphenyl	4-methylphenyl	H	0.9	6.04576	Training
3	benzo[d] [1,3] dioxol-5-yl	4-chlorophenyl	H	3.20	5.49485	Training
4	4-(hydroxymethyl) phenyl	4-chlorophenyl	H	8.41	5.07520	Training
5	2-hydroxyphenyl	4-chlorophenyl	H	0.9	6.04576	Training
6	4-hydroxyphenyl	4-fluorophenyl	methyl	0.53	6.27572	Training
7	4-chlorophenyl	4-hydroxyphenyl	H	3.1	5.50864	Training
8	3-hydroxyphenyl	4-chlorophenyl	H	5.6	5.25181	Training
9	4-hydroxyphenyl	4-bromophenyl	H	0.15	6.82391	Training
10	2-methoxyphenyl	4-chlorophenyl	H	22.5	4.64782	Training
11	4-hydroxyphenyl	cyclohexyl	methyl	0.35	6.45593	Training
12	4-hydroxyphenyl	[1,3]-thiazol-2-yl	H	0.88	6.05552	Training
13	4-hydroxyphenyl	cyclopropyl	H	2.8	5.55284	Training
14	4-hydroxyphenyl	2,4-dichlorophenyl	methyl	0.41	6.38722	Training
15	4-hydroxyphenyl	2,4-dichlorophenyl	H	0.35	6.45593	Training
16	4-hydroxyphenyl	(3-bromo-4-morpholin-1-yl) phenyl	H	0.67	6.17393	Training
17	4-hydroxyphenyl	(4-morpholin-1-yl) phenyl	H	1.9	5.72125	Training
18	3-hydroxyphenyl	adamantyl	methyl	1.99	5.70115	Training
19	4-hydroxyphenyl	adamantyl	methyl	1.00	6.00000	Test
20	3,5-dimethyl-4-hydroxyphenyl	adamantyl	methyl	0.29	6.53760	Training
21	4-methoxyphenyl	adamantyl	methyl	4.10	5.38722	Test
22	4-hydroxiphenyl	phenyl	H	0.11	6.95861	Training
23	4-hydroxiphenyl	4-fluorophenyl	H	0.099	7.00436	Test
24	4-hydroxiphenyl	4-chlorophenyl	H	0.068	7.16749	Training
25	4-hydroxiphenyl	4-nitrophenyl	H	0.025	7.60206	Training
26	4-hydroxiphenyl	2-fluorophenyl	H	0.081	7.09151	Test
27	4-hydroxiphenyl	3-fluorophenyl	H	0.065	7.18709	Training
28	2-hydroxiphenyl	4-chlorophenyl	H	0.14	6.85387	Test
29	4-hydroxiphenyl	2,5-dichlorophenyl	H	0.092	7.03621	Training
30	4-hydroxiphenyl	2,4-difluorophenyl	H	0.085	7.07058	Training
31	4-hydroxi-3,5-dichlorophenyl	4-chlorophenyl	H	0.146	6.83565	Test
32	4-hydroxi-3,5-dimethylphenyl	4-chlorophenyl	H	0.127	6.89620	Training
33	4-hydroxi-3-chloro-5-methylphenyl	4-fluorophenyl	H	0.063	7.20066	Training
34	4-hydroxi-3,5-dichlorophenyl	4-fluorophenyl	H	0.093	7.03152	Test
35	4-hydroxi-2,5-dimethylphenyl	4-fluorophenyl	H	0.104	6.98297	Training

36	4-hydroxi-2,3,5-trimethylphenyl	4-fluorophenyl	H	0.224	6.64975	Training
37	4-aminophenyl	phenyl	H	0.12	6.92082	Training
38	4-aminophenyl	4-fluorophenyl	H	0.105	6.97881	Training
39	4-aminophenyl	4-bromophenyl	H	0.201	6.69680	Training
40	4-aminophenyl	4-chlorophenyl	H	0.16	6.79588	Training
41	4-aminophenyl	4-nitrophenyl	H	0.092	7.03621	Training
42	2-aminophenyl	4-chlorophenyl	H	0.421	6.37572	Test
43	4-aminophenyl	2,5-dichlorophenyl	H	0.06	7.22185	Test
44	4-aminophenyl	2,4-difluorophenyl	H	0.07	7.15490	Test
45	4-hydroxiphenyl	3,4-dichlorophenyl	H	0.9	6.04576	Training
46	3-hydroxyphenyl	phenyl	H	8.3	5.08092	Training
47	4-methoxyphenyl	phenyl	H	4.7	5.32790	Training
48	phenyl	4-chlorophenyl	H	5.6	5.25181	Training
49	3,4-dimethoxyphenyl	4-chlorophenyl	H	5.8	5.23657	Training
50	2,4-dimethoxyphenyl	4-chlorophenyl	H	7.3	5.13668	Training
51	3,5-dimethoxyphenyl	4-chlorophenyl	H	23.5	4.62893	Training
52	4-trifluoromethyl	4-chlorophenyl	H	5,9	5.22915	Test
53	4-trifluoromethoxy	4-chlorophenyl	H	2,9	5.53760	Training
54	3,5-ditrifluoromethyl	4-chlorophenyl	H	3,3	5.48149	Training
55	4-(1,1,1,3,3,3-hexafluoro-2-hydroxypropan-2-yl)	4-chlorophenyl	H	5,0	5.30103	Training
56	phenyl	4-methoxy	H	1.6	5.79588	Test
57	4-fluorophenyl	4-hydroxiphenyl	H	3.5	5.45593	Training
58	4-acetyl	4-fluorophenyl	H	4.9	5.30980	Training
59	2,5-dimethyl-4-hydroxypheyl	4-chlorophenyl	H	0.05	7.30103	Training

Table 2. Values of experimental activity pIC50 and descriptors calculated.

Pic50	ATSC8p	AATSC4c	AATSC7c	AATSC8c	AATSC1p	GATS5s	VE2_Dzp	SaaS	maxHBa	JGI4
7.301	-2.2083	-0.0014	0.0019	-0.0003	0.0315	1.037	0.0048	-1.1971	9.7469	0.0491
5.3098	0.4056	-0.0007	-0.0052	0.0053	0.0198	0.6164	0.0025	-1.286	12.9489	0.0358
5.4559	-1.6866	-0.0017	-0.0013	0.0041	0.0372	0.377	0.0117	-1.2586	12.8408	0.0388
5.7959	-0.5559	-0.0009	-0.0024	0.0022	0.0101	0.9311	0.0021	-1.1281	5.1776	0.0302
5.301	1.2726	0.0015	0.0044	-0.0055	0.0295	1.1059	0.0049	-1.6024	12.9427	0.0328
5.4815	-1.6621	0.0025	0.0062	-0.0006	0.019	1.2118	0.0063	-1.8789	13.1771	0.0401
5.5376	0.1651	-0.0005	0.0076	-0.0059	0.0098	1.1572	0.0237	-1.3435	12.1458	0.0405
5.2291	-3.3099	-0.0001	-0.0074	0.0067	0.0196	1.0389	0.0134	-1.3527	12.5527	0.0346
4.6289	-1.8765	0.003	-0.0023	0.0006	0.018	1.3897	0.0048	-1.0877	3.5866	0.0338
5.1367	1.6745	-0.0014	-0.0002	0.0029	-0.0002	0.9063	0.0115	-1.2257	5.4365	0.043
5.2366	-2.1766	0.0004	0.0015	-0.0003	-0.0002	1.3468	0.012	-1.2023	5.3464	0.0374
5.2518	-0.91	-0.0003	-0.0019	0.0022	0.0202	1.2624	0.0117	-1.0778	3.4526	0.0312
5.3279	-2.9376	0.0019	0.0002	-0.0021	0.0101	1.0996	0.0133	-1.1438	5.2343	0.0254
5.0809	-1.2199	0.0025	-0.0027	-0.0006	0.0372	1.1247	0.0104	-1.1736	9.4415	0.0259
6.0458	-2.1606	-0.002	0.0055	-0.0019	0.0369	0.48	0.014	-1.1596	9.3088	0.0472
7.1549	-2.1505	0.0008	0.0013	0.0007	0.0408	0.8565	0.0054	-1.4182	13.74	0.0384
7.2218	-1.5064	-0.0013	0.0053	-0.0032	0.0412	0.6915	0.014	-1.1303	5.7171	0.0448
6.3757	-0.3148	-0.0035	-0.002	0.0037	0.0409	1.0275	0.0089	-1.1576	5.9548	0.0407
7.0362	-3.4372	-0.0007	0.0047	-0.0027	0.0523	0.9472	0.0014	-1.2904	10.6677	0.0358
6.7959	-2.4036	-0.0011	0.0032	-0.001	0.0409	0.7322	0.0107	-1.1234	5.6934	0.0388
6.6968	-2.579	-0.001	0.0034	-0.0012	0.041	0.7264	0.0112	-1.119	5.6948	0.0388
6.9788	-2.0643	-0.0012	0.0027	-0.0003	0.0408	0.4147	0.0056	-1.2273	12.9059	0.0388
6.9208	-2.0877	-0.0008	0.0047	-0.0028	0.0408	0.7013	0.01	-1.1193	5.6779	0.0312
6.6498	-3.3925	-0.0016	0.0009	0.0002	0.0298	0.5355	0.0023	-1.3195	13.0157	0.0479
6.983	-1.8954	-0.0015	0.0016	0.0003	0.0319	0.5805	0.0004	-1.3011	12.9779	0.0491
7.0315	-2.3097	0.0003	0.0028	-0.0015	0.0369	0.3962	0	-1.2685	12.979	0.0399
7.2007	-3.4323	-0.0007	0.0018	-0.0007	0.0339	0.3882	0.0003	-1.2819	12.9752	0.0399
6.8962	-4.9881	-0.0014	0.0015	-0.0007	0.0315	0.6196	0.005	-1.1913	9.8531	0.0399
6.8356	-2.7265	0.0004	0.0034	-0.0023	0.0371	0.5196	0.0043	-1.1646	9.9348	0.0399
7.0706	-1.5438	0.0006	0.0018	0.0011	0.0373	0.814	0.0067	-1.4495	13.7261	0.0384
7.0362	-0.8202	-0.0016	0.006	-0.003	0.0369	0.4791	0.0153	-1.1615	9.3182	0.0448
6.8539	-0.6709	-0.0041	-0.0024	0.0041	0.0369	0.9889	0.0101	-1.2131	9.778	0.0407
7.1871	-1.4525	-0.0018	0.0077	-0.0043	0.0372	0.6704	0.0084	-1.2913	13.2048	0.0407
7.0915	-1.5627	0.0002	0.004	-0.0016	0.0372	1.0619	0.0097	-1.3414	13.7053	0.0335
7.6021	-2.8497	-0.001	0.0053	-0.0027	0.0495	0.8676	0.0026	-1.3217	10.6588	0.0358
7.1675	-1.733	-0.0014	0.0038	-0.0008	0.0369	0.5053	0.0121	-1.1546	9.279	0.0388
7.0044	-1.4543	-0.0015	0.0032	-0.0001	0.0372	0.3918	0.007	-1.2586	12.8957	0.0388
6.9586	-1.4735	-0.0011	0.0054	-0.0028	0.0372	0.4917	0.0115	-1.1505	9.2533	0.0312
5.3872	2.9084	-0.0004	0.0013	-0.0004	0.0177	0.9492	0.0151	-1.1696	5.2581	0.0324
6.5376	0.3418	-0.0004	0.0014	-0.0012	0.0322	0.6276	0.007	-1.2286	10.019	0.0347
6	1.3402	-0.0003	0.0023	-0.0015	0.0349	0.5054	0.0123	-1.192	9.4449	0.0332
5.7011	1.6606	0.001	-0.0009	-0.0002	0.0349	0.8956	0.0113	-1.2151	9.6718	0.0299
5.7212	-0.7388	-0.0014	0.0013	0	0.0137	0.4583	0.0058	-1.2234	9.3612	0.0255
6.1739	-1.452	-0.0015	0.0018	-0.0005	0.0153	0.4368	0.0033	-1.2227	9.3927	0.0352
6.4559	-0.6426	-0.0001	0.0028	0	0.0369	0.521	0.0146	-1.1606	9.3141	0.0384
6.3872	-2.0903	0.0002	0.0023	-0.0006	0.0338	0.5681	0.0157	-1.2275	9.3582	0.0445
5.5528	-3.2418	-0.0018	0.0064	-0.0027	0.0415	0.3926	0.0006	-0.987	9.1856	0.0335
6.0555	-0.9042	-0.0026	0.007	-0.0024	0.0496	0.5316	0.0096	-2.0525	9.2325	0.0359
6.4559	-1.2497	-0.0004	0.0025	-0.0014	0.0306	0.5479	0.0054	-1.1186	9.3189	0.036
4.6478	-2.3804	-0.0029	-0.003	0.0035	0.0097	1.2112	0.0119	-1.1713	5.374	0.0374
6.8239	-1.8779	-0.0014	0.0041	-0.0011	0.0368	0.5032	0.0125	-1.1502	9.2804	0.0388
5.2518	-1.4472	0.0022	-0.0043	0.0014	0.0369	1.1485	0.011	-1.1777	9.4713	0.0334
5.5086	-0.4013	-0.0015	-0.0021	0.0042	0.0369	0.586	0.006	-1.1546	9.3177	0.0388
6.2757	-0.5513	-0.0009	0.0029	-0.0008	0.0343	0.408	0.0088	-1.3254	12.9962	0.0425
6.0458	-0.6709	-0.0041	-0.0024	0.0041	0.0369	0.9889	0.0101	-1.2131	9.778	0.0407
5.0752	-2.496	-0.0008	-0.0046	0.0053	0.0339	1.0059	0.0109	-1.1504	9.0539	0.0371
5.4949	-2.3877	0.0006	0.0011	-0.001	0.0045	0.9641	0.0109	-1.1778	5.4114	0.032
6.0458	-4.1902	-0.001	0.0037	-0.0013	0.0343	0.4633	0.0116	-1.1648	9.2756	0.0388
6.1192	-1.5829	-0.0015	0.0022	0.0005	0.0211	0.4876	0.0018	-1.2123	9.3036	0.0342

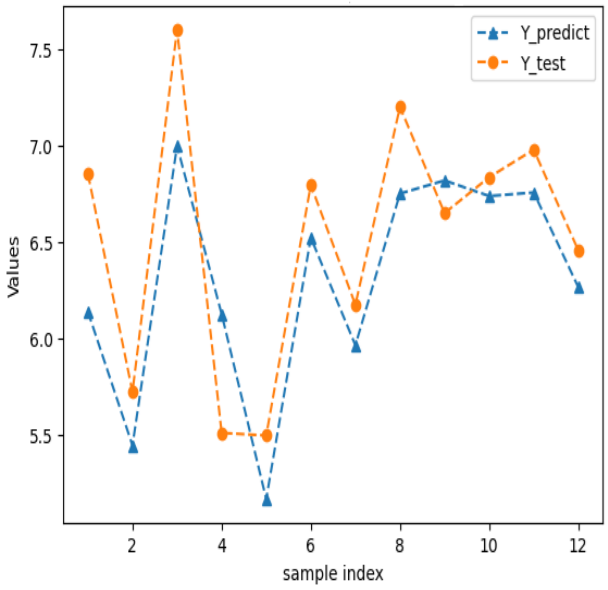


Fig. 1. Comparison of Y_predicted and Y_test

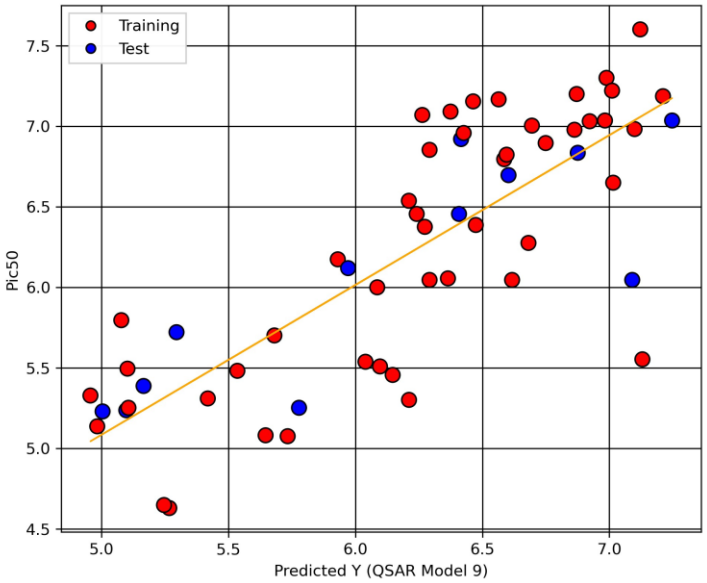


Fig. 2. Plot of predicted pIC50 against experimental pIc50 for training and test sets

Table 4: Values of observed and predicted 5 -LOX inhibitory activity.

C.no.	Set	PIc50 (obs)	PIc50 (pre.)	Residual	C.no.	Set	PIc50 (obs)	PIc50 (pre.)	Residual
1	Training	6.11919	5.9715	0.14769	38	Training	6.97881	6.8626	0.11621
2		6.04576	6.617	-0.57124	39		6.69680	6.6029	0.0939
3		5.49485	5.1029	0.39195	40		6.79588	6.5859	0.20998
4		5.07520	5.7335	-0.6583	41		7.03621	7.2470	-0.21079
5		6.04576	6.2921	-0.24634	45		6.04576	7.090	-1.04424
6		6.27572	6.6814	-0.40568	46		5.08092	5.6466	-0.56568
7		5.50864	6.0969	-0.58826	47		5.32790	4.9570	0.3709
8		5.25181	5.7782	-0.52639	48		5.25181	5.1067	0.14511
9		6.82391	6.5954	0.22851	49		5.23657	5.0982	0.13837
10		4.64782	5.2467	-0.59888	50		5.13668	4.9838	0.15288
11		6.45593	6.408	0.04793	51		4.62893	5.2674	-0.63847
12		6.05552	6.3645	-0.30898	53		5.53760	6.0398	-0.5022
13		5.55284	7.13	-1.57716	54		5.48149	5.5355	-0.05401
14		6.38722	6.4742	-0.08698	55		5.30103	6.2112	-0.91017
15		6.45593	6.2415	0.21443	57		5.45593	6.1469	-0.69097
16		6.17393	5.9308	0.24313	58		5.30980	5.4191	-0.1093
17		5.72125	5.2959	0.42535	59		7.30103	6.9887	0.31233
18		5.70115	5.6812	0.01995	19	Test	6.00000	6.0861	-0.0861
20		6.53760	6.2106	0.327	21		5.38722	6.695	-1.30778
22		6.95861	6.4264	0.53221	23		7.00436	6.6950	0.30936
24		7.16749	6.564	0.60349	26		7.09151	6.3747	0.71681
25		7.60206	7.1216	0.48046	28		6.85387	6.2921	0.56177
27		7.18709	7.2111	-0.02401	31		6.83565	6.8755	-0.03985
29		7.03621	6.9826	0.05361	34		7.03152	6.9227	0.10882
30		7.07058	6.2635	0.80708	42		6.37572	6.2734	0.10232
32		6.89620	6.7492	0.147	43		7.22185	7.0106	0.21125
33		7.20066	6.8712	0.32946	44		7.15490	6.4637	0.6912
35		6.98297	7.0992	-0.11623	52		5.22915	5.0049	0.22425
36		6.64975	7.0151	-0.36535	56		5.79588	5.0786	0.71728
37		6.92082	6.4162	0.50462					

V. Conclusions

A QSAR model was generated to evaluate the differences in residual values between the observed and predicted inhibitory activity against 5-lipoxygenase for 59 selected molecules of thiazole derivatives. Ten molecular descriptors were used to develop and evaluate the performance of the QSAR models. QSAR studies on these compounds in the training set revealed an acceptable correlation between the predicted pIC50 values and the experimental values using the MLR statistical method. The test set values suggest a high level of accuracy of the model predictions as they

closely match the observed values. The QSAR model was generated using an open-source Python script and could be a valuable tool for designing compounds with promising biological activity. The results obtained in this study could serve as input for the design of new anti-inflammatory drugs.

ACKNOWLEDGMENT

The authors wish to thank the faculty of basic sciences of the Universidad Tecnológica de Bolívar-UTB

REFERENCES

[1] R. Mashima and T. Okuyama, "The role of lipoxygenases in pathophysiology; new insights and future perspectives," (in eng), *Redox Biol,* vol. 6, pp. 297-310, Dec 2015.

[2] S. C. Dyall *et al.*, "Polyunsaturated fatty acids and fatty acid-derived lipid mediators: Recent advances in the understanding of their biosynthesis, structures, and functions," vol. 86, p. 101165, 2022.

[3] C. J. N. E. J. o. M. Godson, "Balancing the effect of leukotrienes in asthma," vol. 382, no. 15, pp. 1472-1475, 2020.

[4] T. Honda and K. J. I. I. Kabashima, "Prostanoids and leukotrienes in the pathophysiology of atopic dermatitis and psoriasis," vol. 31, no. 9, pp. 589-595, 2019.

[5] H. Asakura and T. Kitahora, "Antioxidants and polyphenols in inflammatory bowel disease: ulcerative colitis and Crohn disease," in *Polyphenols: prevention and treatment of human disease*: Elsevier, 2018, pp. 279-292.

[6] M. Aparicio-Soto, M. Sánchez-Hidalgo, and C. J. N. R. R. Alarcón-de-la-Lastra, "An update on diet and nutritional factors in systemic lupus erythematosus management," vol. 30, no. 1, pp. 118-137, 2017.

[7] C. Bortuzzo *et al.*, "The effect of leukotrienes B and selected HETEs on the proliferation of colon cancer cells," vol. 1300, no. 3, pp. 240-246, 1996.

[8] D. Wang and R. N. J. N. R. C. DuBois, "Eicosanoids and cancer," vol. 10, no. 3, pp. 181-193, 2010.

[9] R. Wisastra and F. J. Dekker, "Inflammation, Cancer and Oxidative Lipoxygenase Activity are Intimately Linked," (in eng), *Cancers (Basel),* vol. 6, no. 3, pp. 1500-21, Jul 17 2014.

[10] S. Sinha, M. Doble, S. J. B. Manju, and M. Chemistry, "5-Lipoxygenase as a drug target: A review on trends in inhibitors structural design, SAR and mechanism based approach," vol. 27, no. 17, pp. 3745-3759, 2019.

[11] N. R. Pinelli, R. Cha, M. B. Brown, and L. A. J. A. o. P. Jaber, "Addition of thiazolidinedione or exenatide to oral agents in type 2 diabetes: a meta-analysis," vol. 42, no. 11, pp. 1541-1551, 2008.

[12] M. E. Ernst and M. A. J. A. J. o. H. Fravel, "Thiazide and the thiazide-like diuretics: review of hydrochlorothiazide, chlorthalidone, and indapamide," vol. 35, no. 7, pp. 573-586, 2022.

[13] P. C. Sharma, K. K. Bansal, A. Sharma, D. Sharma, and A. Deep, "Thiazole-containing compounds as therapeutic targets for cancer therapy," (in eng), *Eur J Med Chem,* vol. 188, p. 112016, Feb 15 2020.

[14] K. B. Gangurde, R. A. More, V. A. Adole, and D. S. J. J. o. M. S. Ghotekar, "Design, synthesis and biological evaluation of new series of benzotriazole-pyrazole clubbed thiazole hybrids as bioactive heterocycles: Antibacterial, antifungal, antioxidant, cytotoxicity study," vol. 1299, p. 136760, 2024.

[15] E. d. F. Santiago *et al.*, "Evaluation of the anti-Schistosoma mansoni activity of thiosemicarbazones and thiazoles," vol. 58, no. 1, pp. 352-363, 2014.

[16] R. Kalkhambkar, G. Kulkarni, H. Shivkumar, and R. N. J. E. J. o. M. C. Rao, "Synthesis of novel triheterocyclic thiazoles as anti-inflammatory and analgesic agents," vol. 42, no. 10, pp. 1272-1276, 2007.

[17] K. J. A. i. P. Roy, Chemical, Food, Agricultural, E. Sciences, and S. Springer: Cham, "Advances in QSAR modeling," vol. 555, p. 39, 2017.

[18] A. Marković *et al.*, "Thiazole derivatives as dual inhibitors of deoxyribonuclease I and 5-lipoxygenase: A promising scaffold for the development of neuroprotective drugs," vol. 381, p. 110542, 2023.

[19] J. Suh, E. K. Yum, H. G. Cheon, Y. S. J. C. b. Cho, and d. design, "Synthesis and Biological Evaluation of N-aryl-4-aryl-1, 3-Thiazole-2-Amine Derivatives as Direct 5-Lipoxygenase Inhibitors," vol. 80, no. 1, pp. 89-98, 2012.

[20] C. B. Rödl *et al.*, "Multi-dimensional target profiling of N, 4-diaryl-1, 3-thiazole-2-amines as potent inhibitors of eicosanoid metabolism," vol. 84, pp. 302-311, 2014.

[21] S. B. Kretschmer *et al.*, "Characterization of the molecular mechanism of 5-lipoxygenase inhibition by 2-aminothiazoles," vol. 123, pp. 52-62, 2017.

[22] J. Dong *et al.*, "ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation," vol. 7, no. 1, pp. 1-10, 2015.

[23] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C. P. Varghese, and R. K. J. I. J. D. D. D. Agrawal, "Validation of QSAR models-strategies and importance," vol. 3, pp. 511-519, 2011.