

Quantitative Structure–activity Relationship Modeling of Some Naphthoquinone Derivatives as Inhibitors of Pathogenic Agent IDO1

Abstract

Quantitative structure–activity relationship (QSAR) was performed to analyze naphthoquinone derivatives as an inhibitor of indoleamine 2,3-dioxygenase pathogen via multivariate regression (MLR) and artificial neural network. The best descriptors were picked to construct the QSAR. Two sets of exercises and experiments were also performed using Principal Component Analysis for multiple linear regression (MLR). A quantitative model was then proposed based on these analyses and the activity of the compounds based on multivariate statistical analysis was interpreted. The study finally revealed that although the MLR model can predict the activity of the compounds to some extent, the artificial neural network (ANN) model results indicate that the predictions obtained by the neural network are much better and more efficient than other models. The neural network was also used where three coefficients of correlation were used. The results uncovered that the ANN model is statistically significant and has good stability for data validation for the validation method. Share Descriptive relationships of structure and activity were also examined.

Keywords: Artificial neural network (ANN), multiple linear regression (MLR), naphthoquinone derivatives, pathogenic agent, quantitative structure–activity relationship (QSAR)

Introduction

The interaction between the immune system and the growth of tumors is complex and active. Although the host immune system has the capacity to detect and identify tumor cells, many malignancies can actively hinder the immune response.^[1,2] This can be inhibited by several mechanisms, such as the recruitment of immunosuppressive cells, the removal of T cells, and the activation of inspection pathways.^[3] There is increasing evidence that indoleamine 2,3-dioxygenase 1 (IDO1) plays a seminal role in suppressing immunity in the microscopic environment of the tumor. These studies indicate that IDO1 might be a valid therapeutic target in immunotherapy.^[4-6]

Many studies have highlighted the point that IDO1 is suppressed in normal human tissues including the spleen, intestine, and lung.^[7,8] Also, using many types of human tumors, they are constitutively expressed by inflammatory stimuli such as interferon- γ (IFN- γ) and transforming growth factor- β .^[9,10] Overexpression of IDO1 is abolished by tumors that destroy tryptophan and accumulate large amounts of kynurenine

and its downstream metabolites in the tumor microenvironment. Tryptophan deficiency can also be mediated by mTORC1 inhibition^[11] and GCN2 activation,^[12] affecting T-cell sensitivity and production of kynurenine active pathway metabolites, leading to increased trade differentiation as a result of aryl hydrocarbon receptor (AHR) activation.^[13] Overexpression of IDO1 also affects dendritic cells and macrophages, conversion to normal killer cells, and production of reactive oxygen species. It, thus, enables tumor cells to suppress host immune responses.^[13-16] New research on COVID-19 has also shown that after entry into cells, coronaviruses activate AHRs by an IDO1-independent mechanism, bypassing the IDO1-kynurenine-AhR pathway.^[17]

Attempts to quantitative analysis of the relationship between the structure and activity of compounds provide an understanding of the effect of structure on their activity, which may not be easy when having large amounts of data. In addition, this method makes it possible to make predictions that lead to the synthesis of new compounds similar to the expected amount of activity. The quantitative structure–activity relationship (QSAR)

**Sajjad Jazayeri Farsani,
Saeid Asadpour,
Abolfazi Semnani,
Shima Ghanavati
Nasab**

*Department of Chemistry,
Faculty of Sciences, University of
Shahrekord, Shahrekord, Iran*

Received: 13 Oct 2020
Accepted: 03 Oct 2021
Published: 17 Dec 2021

Address for correspondence:
Dr. Saeid Asadpour,
Department of Chemistry,
Faculty of Science, Shahrekord
University, Shahrekord, Iran.
E-mail: s.asadpour@sku.ac.ir

Access this article online

Website:

www.jrpsjournal.com

DOI:10.4103/jrpts.JRPTPS_124_20

Quick Response Code:



How to cite this article: Jazayeri Farsani S, Asadpour S, Semnani A, Ghanavati Nasab S. Quantitative structure–activity relationship modeling of some naphthoquinone derivatives as inhibitors of pathogenic agent IDO1. *J Rep Pharma Sci* 2021;10:317-34.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

method covers a wide range of chemical measurements and biological tests, statistical methods, and interpretation of results.^[18–20] The QSAR method can be used for any molecular design purpose, including predicting the biological activity and physicochemical properties, better understanding the mechanism of action in a number of chemicals, saving and reducing product costs (drugs, pesticides, and new chemicals), and replacing the use of laboratory animals.^[21] QSAR modeling studies have also been performed on different types of IDO1 inhibitors.^[22]

Given the above-mentioned points, this study aimed to develop QSAR models for the use of naphthoquinone derivatives as inhibitors of the pathogenic factor IDO1 using several statistical approaches, including Principal Component Analysis (PCA), multivariate regression (MLR), and artificial neural network (ANN). The validation method was chosen to evaluate the performance and stability of this model.

Materials and Methods

A QSAR study consists of three parts as follows: (1) data related to the activity or feature under study (in this study, IDO1 inhibitors) that should be modeled and predicted, (2) the descriptors on which the model is based, and (3) a mathematical or statistical method used to formulate the model such as MLR and ANN.^[23–25]

Data sources

In the current study, the compounds of naphthoquinone derivatives as inhibitors of the pathogenic factor IDO1 were investigated and the data were retrieved from the publication of Xiangbao Meng *et al.*^[26] Figure 1 shows the basic structure of the naphthoquinone. Also, Table 1 presents the values of the percent inhibition of substituted compounds studied from naphthoquinone derivatives.

Molecular descriptors

All molecules were extracted with ChemDraw software, optimized with MM2 molecular force field, and calculated with the DRAGON bundle.

Dragon software calculates a large number of descriptors from which the most effective descriptors need to be selected. In the first step, descriptors with constant and zero values were omitted because they could not show the relationship between structure and activity well. In the second step, the correlation between the descriptors and the dependent variable was established and the descriptors that had low correlation with the dependent variable were eliminated.

In the third stage, since there is a correlation between the descriptors, the descriptors whose correlation coefficient was greater than 0.95 indicate a linear correlation between them because both descriptors have identical information. Therefore, some of these descriptors as well as those that were less correlated with the dependent variable were removed.

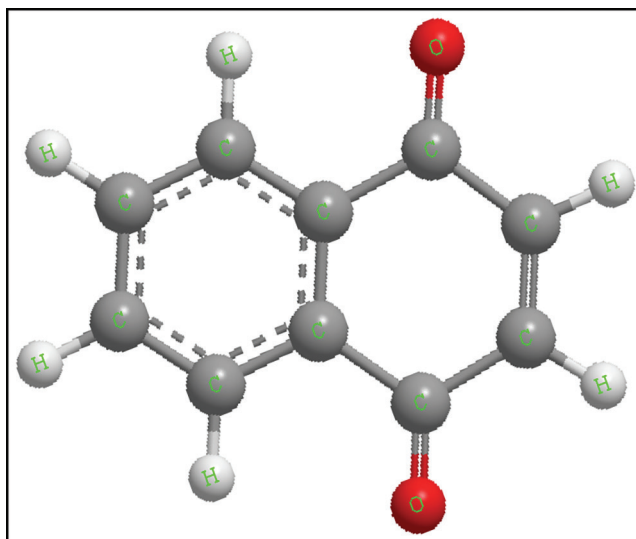


Figure 1: Chemical structure of naphthoquinone

Of all the descriptors, six were selected to predict IDO1 pathway inhibitor activity as follows:

Geary autocorrelation of lag 2 weighted by van der Waals volume (GATS2v), R maximal autocorrelation of lag 3/weighted by Sanderson electronegativity (R3e⁺), signal 16/weighted by mass (Mor16m), first component symmetry directional WHIM index/weighted by mass (G1m), second component symmetry directional WHIM index/unweighted (G2u), and signal 31/weighted by mass (Mor31m).

Statistical analysis

The purpose of quantitative structure–activity relationship (QSAR) analysis is to predict the biological activities of the compounds by chemical structures using models. In the QSAR analysis, quantitative descriptors and analysis results are used in a mathematical model to depict the chemical structure that describes the relationship between chemical structure and biological activity.

The target molecules were divided into two groups of training and testing using Minitab software. PCA is a high-performance statistical method for summarizing all encrypted information in the structure of compounds

Among the methods of regression analysis, MLR is typically taken as a regression-based method for QSAR or QSPR analysis. Each variable is added to the equation before another and regression is performed. The new expression remains if an experiment confirms the significance of the equation. This regression method is useful when a large number of variables and key descriptors are unknown.^[27] In addition, it selects descriptors that are used as input parameters to the ANN.

MLR is generated using Excel software. Various parameters are used to evaluate the model such as correlation coefficient (R), mean square error (RMSE), and cross-correlation correlation coefficient. ANN analysis is performed via the Matlab Toolkit in the Components Database. A number of unique ANN

Table 1: Observed inhibition% of studied naphthoquinone derivatives

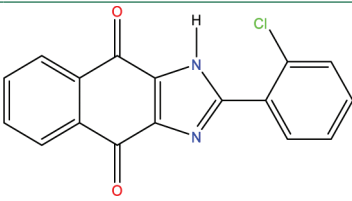
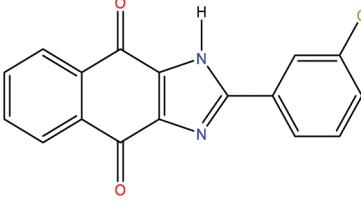
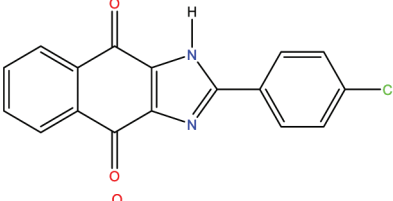
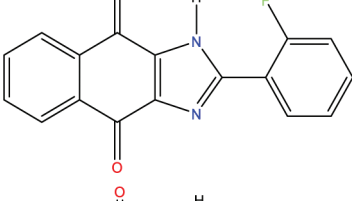
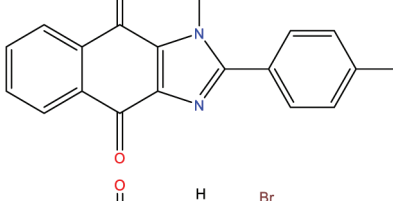
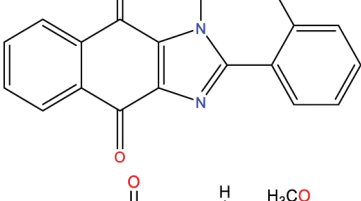
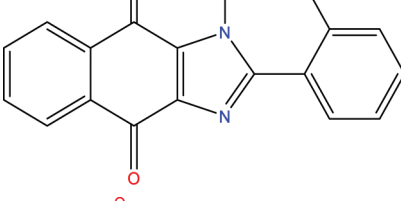
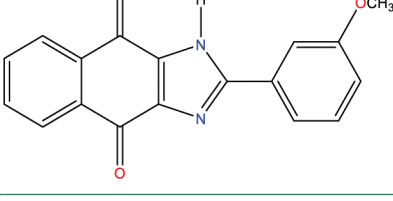
Number	Chemical name	Chemical structure	Inhibition%
1	2-(2-Chlorophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		34
2	2-(3-Chlorophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		59
3	2-(4-Chlorophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		20.5
4	2-(2-Fluorophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		38
5	2-(4-Fluorophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		16.5
6	2-(2-Bromophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		33.5
7	2-(2-Methoxyphenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		38
8	2-(3-Methoxyphenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		30.5

Table 1: Continued

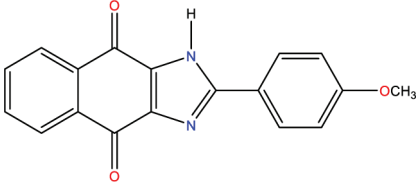
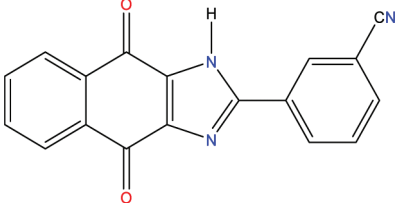
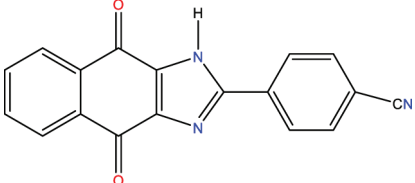
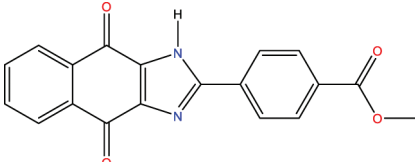
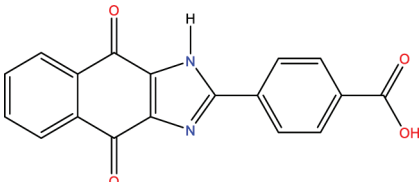
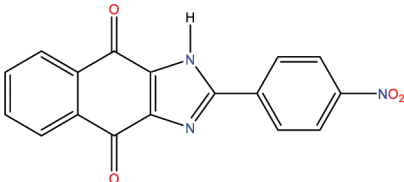
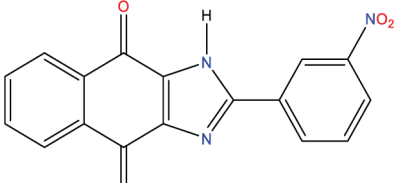
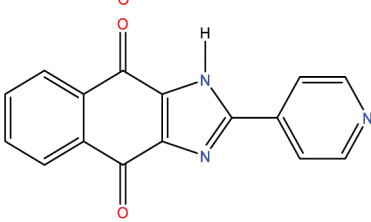
Number	Chemical name	Chemical structure	Inhibition%
9	2-(4-Methoxyphenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		27.5
10	3-(4,9-Dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-2-yl) benzonitrile		34.5
11	4-(4,9-Dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-2-yl) benzonitrile		66.5
12	Methyl 4-(4,9-dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-2-yl) benzoate		38
13	4-(4,9-Dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-2-yl) benzoic acid		28
14	2-(4-Nitrophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		62
15	2-(3-Nitrophenyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		46
16	2-(Pyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		79

Table 1: Continued

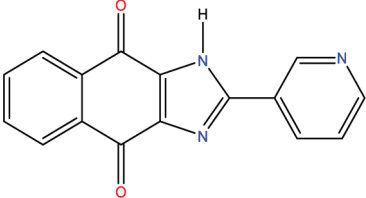
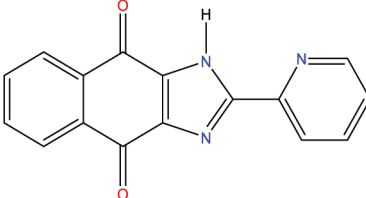
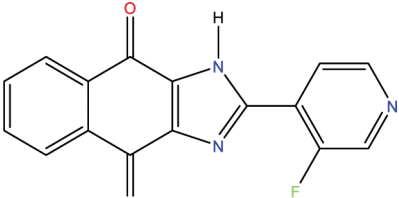
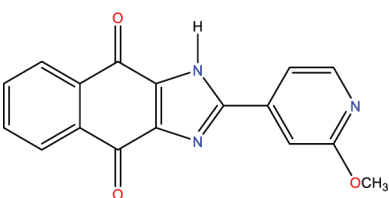
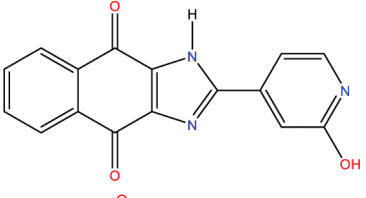
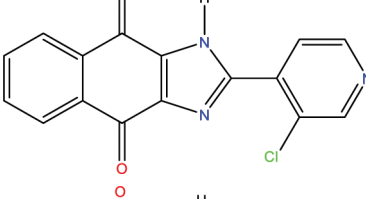
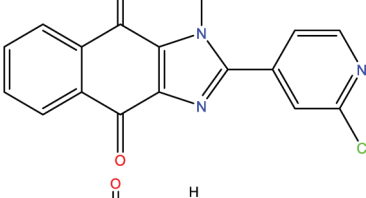
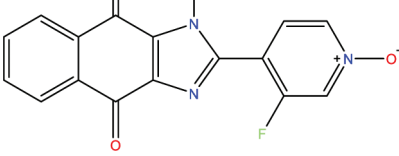
Number	Chemical name	Chemical structure	Inhibition%
17	2-(Pyridin-3-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		58
18	2-(Pyridin-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		56
19	2-(3-Fluoropyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		42
20	2-(2-Methoxypyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		51.5
21	Procedure for the synthesis of 2-(2-hydroxypyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		37
22	2-(3-Chloropyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		45
23	2-(2-Chloropyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		45
24	4-(4,9-Dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-2-yl)pyridine 1-oxide		30.5

Table 1: Continued

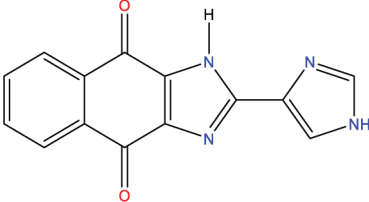
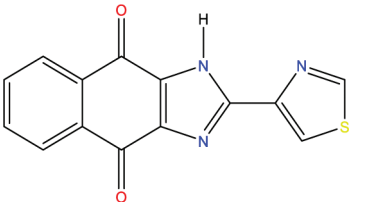
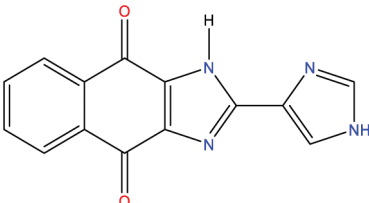
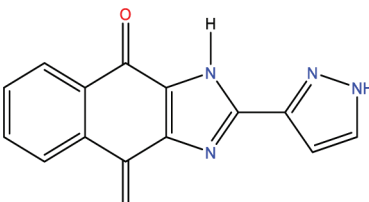
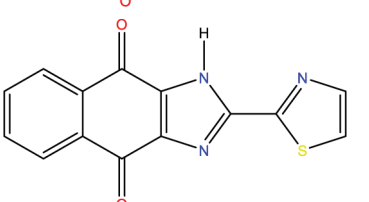
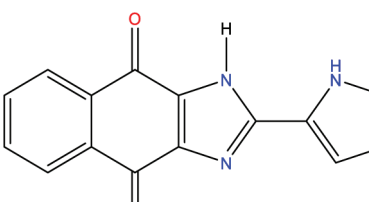
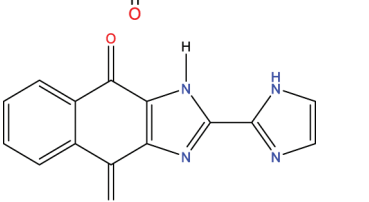
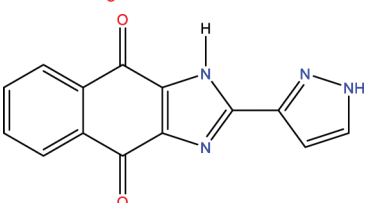
Number	Chemical name	Chemical structure	Inhibition%
25	<i>2-(Thiazol-5-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		38
26	<i>2-(Thiazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		74.5
27	<i>2-(1H-imidazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		34
28	<i>2-(1H-pyrazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		53.5
29	<i>2-(Thiazol-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		37.5
30	<i>2-(1H-pyrrol-3-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		67.5
31	<i>2-(1H-imidazol-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		74.5
32	<i>2-(1H-pyrazol-3-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		46

Table 1: Continued

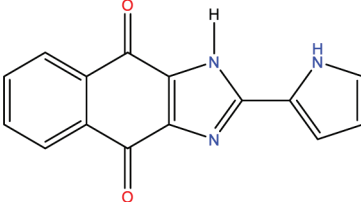
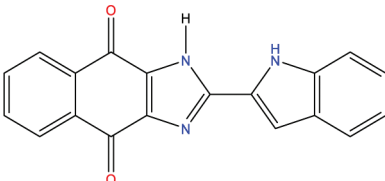
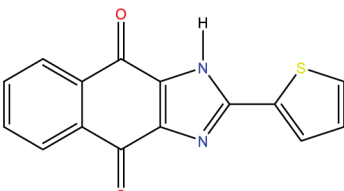
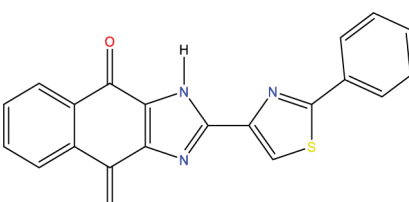
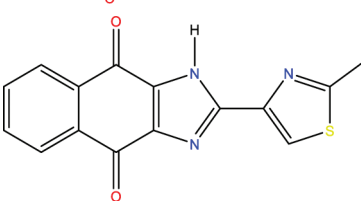
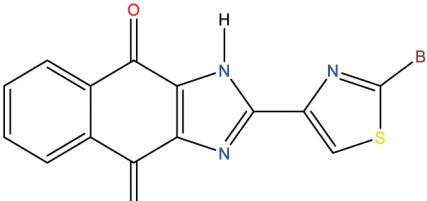
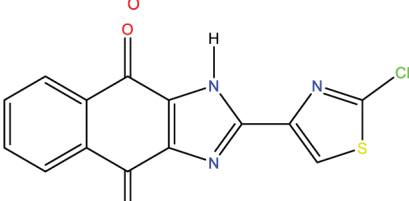
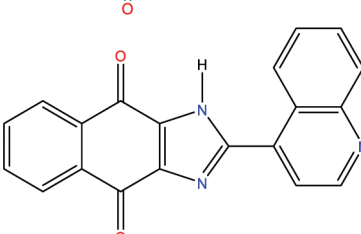
Number	Chemical name	Chemical structure	Inhibition%
33	<i>2-(1H-pyrrol-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		44.5
34	<i>2-(1H-indol-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		17.5
35	<i>2-(Thiophen-2-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		57
36	<i>2-(2-Phenylthiazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		12
37	<i>2-(2-Methylthiazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		50
38	<i>2-(2-Bromothiazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		37
39	<i>2-(2-Chlorothiazol-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		15
40	<i>2-(Quinolin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		28

Table 1: Continued

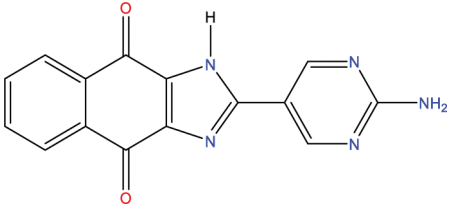
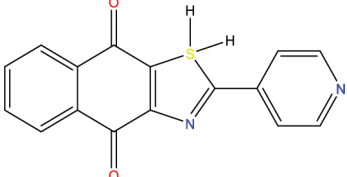
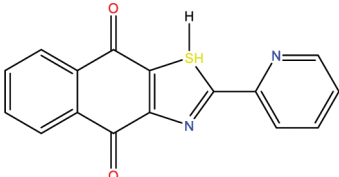
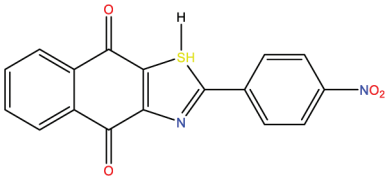
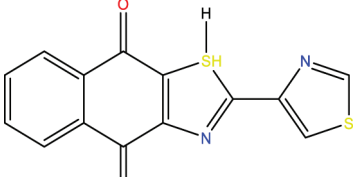
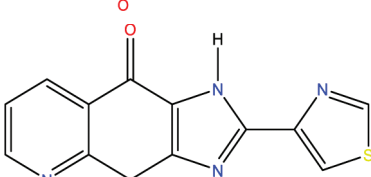
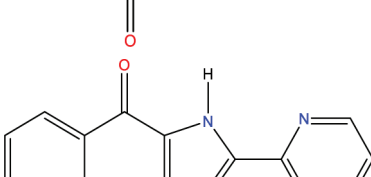
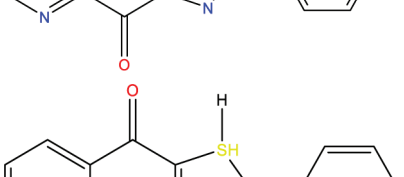
Number	Chemical name	Chemical structure	Inhibition%
41	<i>(2-Aminopyrimidin-5-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		53.5
42	<i>2-(2-Aminopyrimidin-5-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione</i>		88
43	<i>2-(Pyridin-2-yl)naphtho[2,3-d]thiazole-4,9-dione</i>		74
44	<i>2-(Thiazol-4-yl)naphtho[2,3-d]thiazole-4,9-dione</i>		76
45	<i>2-(Pyridin-4-yl)-3H-imidazo[4,5-g]quinoline-4,9-dione</i>		65.5
46	<i>2-(Thiazol-4-yl)-3H-imidazo[4,5-g]quinoline-4,9-dione</i>		76
47	<i>2-(Pyridin-2-yl)-3H-imidazo[4,5-g]quinoline-4,9-dione</i>		71.5
48	<i>2-(Pyridin-4-yl)thiazolo[5,4-g]quinoline-4,9-dione</i>		73

Table 1: Continued

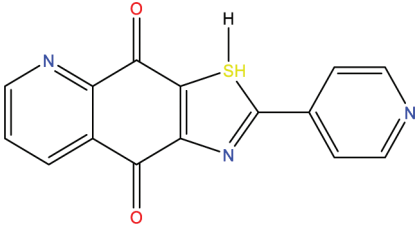
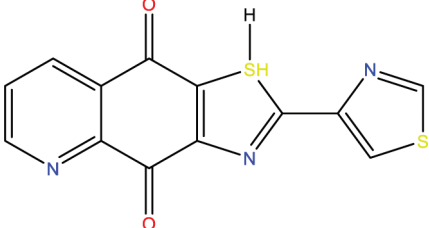
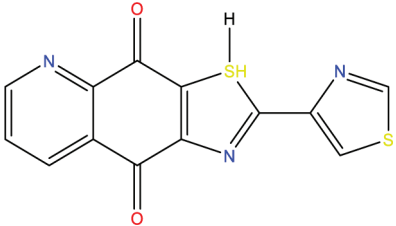
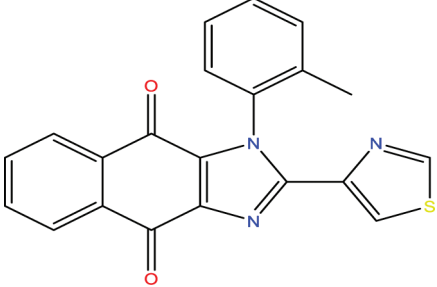
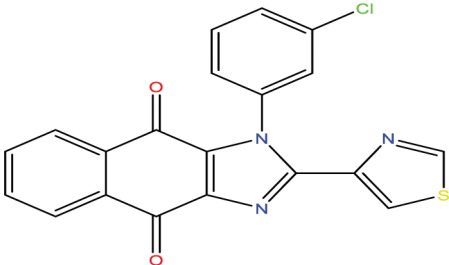
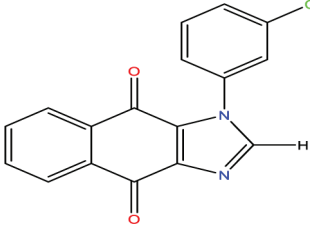
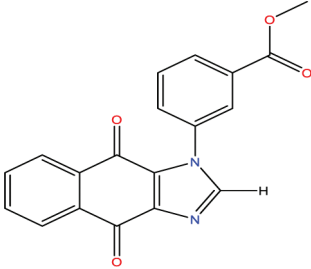
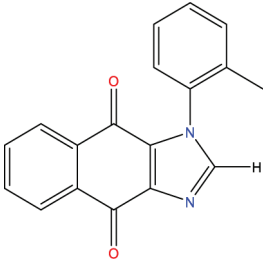
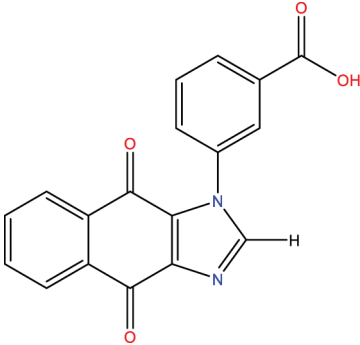
Number	Chemical name	Chemical structure	Inhibition%
49	2-(Pyridin-4-yl)thiazolo[5,4-g]quinoline-4,9-dione		70.5
50	2-(Thiazol-4-yl)thiazolo[4,5-g]quinoline-4,9-dione		73.5
51	2-(Thiazol-4-yl)thiazolo[5,4-g]quinoline-4,9-dione		69.5
52	1-(2-Methylbenzyl)-2-(pyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		9
53	1-(3-Chlorobenzyl)-2-(pyridin-4-yl)-1H-naphtho[2,3-d]imidazole-4,9-dione		26
54	1-(3-Chlorobenzyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		16

Table 1: Continued

Number	Chemical name	Chemical structure	Inhibition%
55	Methyl 3-((4,9-dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-1-yl)methyl)benzoate		9
56	1-(2-Methylbenzyl)-1H-naphtho[2,3-d]imidazole-4,9-dione		10
57	3-((4,9-Dioxo-4,9-dihydro-1H-naphtho[2,3-d]imidazol-1-yl)methyl)benzoic acid		16

models were designed, manufactured, and trained. Regarding the structure of a neural network, it entails three underlying elements of the processing elements or nodes, the topology of the connections between the nodes, and the learning rule by which new information is encoded in the network. Among the different models of ANN, the forward feed distribution network was decided to be used in this study. In this type of network, neurons are set as the input layer, a hidden layer, and an output layer. Each neuron in each layer is entirely related to the neurons of a single layer, and there is no correlation among the neurons belonging to one layer.

Results and Discussion

Dataset for analysis

A QSAR study was performed for 57 derivatives of naphthoquinone as an inhibitor IDO1 to determine a quantitative relationship between the structure and inhibitory activity.^[26] Table 2 represents the values of the six descriptors.

Correlation analysis was performed to identify the relationship between different variables. Table 3 presents the correlation

coefficients matrix for the relevance between the six descriptors selected.

The obtained matrix provides information on the degree of correlation between the variables. In general, the results indicate a low correlation ($r < 0.5$) between most variables. Although a high interrelationship was observed between GATS2v and Mor31m ($r = -0.5032$), a low interrelationship was observed between R3e+ and G2u ($r = 0.03975$).

MLR model results

Initially, six final descriptors coding 57 molecules (as mentioned above) were sent to the PCA for classification of compounds into train and test sets to validate the MLR model. In total, 45 molecules were included in the train set to construct QSAR models, whereas the remaining 12 molecules constituted the test set. Sections were randomly selected using PCA. Figure 2 shows the results of data classification by the PCA analysis.

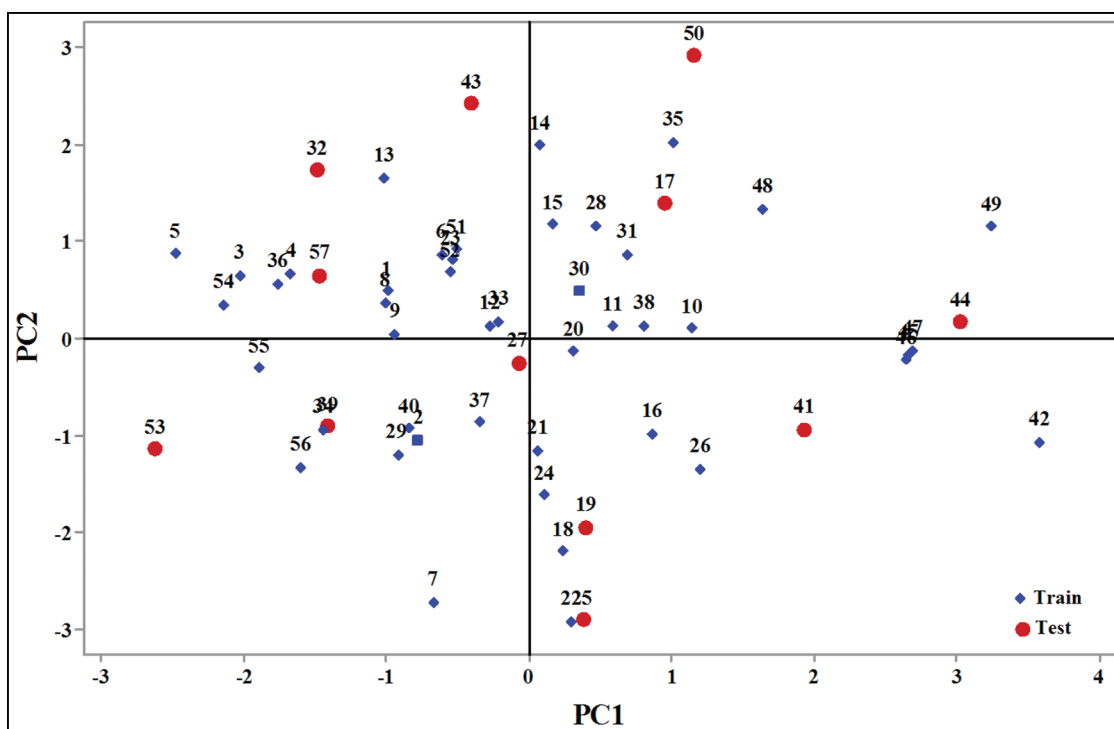
Many attempts have been made to establish an acceptable relationship between the molecular descriptors and the values of Inhibition%. However, the best relationship ultimately

Table 2: Values of the parameters of the derivatives of naphthoquinone investigated

Number	%Inhibition	Descriptors entered in the model					
		GATS2v	R3e+	Mor16m	G1m	G2u	Mor31m
1	34	-0.46218	-0.42708	-0.45359	-0.60586	-0.65671	0.563068
2	59	-0.56904	-0.90728	-0.69039	1.456645	-0.11536	0.780917
3	20.5	-0.56904	-0.86362	-1.12892	-0.73477	-1.39141	1.024908
4	38	-2.03839	-0.47074	-0.65531	-0.47695	-1.08206	0.284221
5	16.5	-2.09183	2.541392	-0.90527	0.038672	0.619334	1.251471
6	33.5	-1.51744	0.14042	0.43663	-0.99258	-0.23136	-0.20376
7	38	-1.51744	-0.90728	0.43663	3.13243	1.083352	-0.37804
8	30.5	0.165636	-0.77631	-0.85703	-0.47695	-1.31407	0.50207
9	27.5	0.165636	-0.73266	-0.38781	-0.47695	-0.23136	0.824487
10	34.5	0.766735	0.795231	0.752371	-0.47695	0.464662	-0.83117
11	66.5	0.098847	0.271382	-0.05891	-0.47695	0.812675	0.397502
12	38	-0.46218	-0.42708	0.068265	-0.60586	0.000644	-0.00334
13	28	-0.72933	0.664269	-0.28694	-1.12149	-1.50741	0.284221
14	62	-0.76941	3.06524	0.015641	-0.60586	0.000644	0.615351
15	46	-0.06145	-0.0342	0.739216	-0.47695	-1.31407	0.153511
16	79	0.539653	-0.25247	0.037567	1.972271	-0.23136	0.240651
17	58	0.366003	1.799273	0.4454	-0.60586	-0.34737	-0.04691
18	56	-0.88963	0.489653	-1.12454	1.585551	1.586037	-1.23201
19	42	0.192352	-0.0342	-1.29995	0.812112	0.967347	-1.38015
20	51.5	0.112205	-0.6017	-0.38781	0.167579	-1.12073	-0.65689
21	37	0.646516	-0.64535	-0.77371	0.167579	0.464662	-0.32576
22	45	0.098847	-0.6017	-1.28241	1.327738	2.011387	-1.06644
23	45	-0.46218	-0.42708	-0.38342	-0.60586	-1.35274	0.162225
24	30.5	1.234257	-0.6017	-0.93597	1.327738	-0.38604	-0.63075
25	38	0.659873	-0.6017	-1.08507	1.327738	2.011387	-0.88345
26	74.5	1.234257	-0.6017	-0.93597	1.327738	-0.38604	-0.63075
27	34	0.673231	-0.68901	-0.38342	-0.21914	-0.50204	-0.05562
28	53.5	0.446149	-0.95093	1.054957	-1.25039	-0.65671	0.754775
29	37.5	-0.64919	-0.64535	-0.25625	1.327738	0.348658	0.475928
30	67.5	0.753378	-0.0342	0.1472	0.296485	-0.54071	1.120762
31	74.5	1.047248	0.751577	-0.78248	-0.47695	-0.57938	0.641493
32	46	-0.38203	0.576961	-0.629	-0.60586	-1.35274	1.913733
33	44.5	0.125563	-0.64535	-0.08522	0.167579	-1.12073	0.205795
34	17.5	1.047248	-0.64535	-1.37888	0.167579	0.464662	1.399609
35	57	-0.00801	1.319079	1.993412	-1.3793	0.387326	0.789631
36	12	-1.11671	-0.77631	0.208594	-0.60586	-0.65671	0.807059
37	50	-0.10152	-0.73266	0.581345	1.714458	0.000644	0.955196
38	37	0.539653	-0.68901	1.449634	-0.60586	0.309989	-0.04691
39	15	-1.98496	-0.81997	1.550496	0.941018	1.276692	0.693777
40	28	0.432791	-0.55804	-1.40958	-0.09023	0.348658	0.258079
41	53.5	1.27433	-0.1215	0.313841	-0.21914	0.967347	-1.31915
42	88	1.675063	0.184074	0.449786	0.296485	0.619334	-2.37354
43	74	-0.38203	2.366775	-0.83072	-0.60586	-1.35274	1.347325
44	76	0.846882	-0.29612	1.664514	-0.47695	-0.23136	-1.9117
45	65.5	0.846882	-0.38343	2.024109	-0.47695	0.812675	-1.25815
46	76	1.27433	0.009458	0.818151	0.167579	-0.3087	-1.54571
47	71.5	1.27433	-0.07785	1.633817	0.167579	0.077981	-1.15358
48	73	1.100679	-0.20881	0.169126	-1.89493	-0.77272	-0.34319
49	70.5	1.100679	1.799273	1.989026	-1.50821	1.160688	-1.33658
50	73.5	-0.46218	2.410429	1.782917	-0.60586	-1.35274	0.380074
51	69.5	-2.03839	-0.07785	0.414703	0.296485	-1.08206	0.336504
52	9	-0.46218	0.096766	0.568189	-1.12149	-0.3087	-0.36933
53	26	-1.70445	-0.55804	-1.68147	1.714458	-0.65671	0.746061
54	16	-2.09183	2.585046	0.432245	0.941018	1.663374	1.704598
55	9	-0.14159	-0.95093	-1.107	-1.25039	0.812675	1.216615
56	10	-0.14159	-0.20881	-0.94913	0.812112	0.812675	0.658921
57	16	-0.14159	-0.55804	-0.96667	-1.25039	-0.92739	0.563068

Table 3: Correlation matrix between different obtained descriptors

	%Inhibition	GATS2v	R3e+	Mor16m	G1m	G2u	Mor31m
%Inhibition	1						
GATS2v	0.455393	1					
R3e+	0.249624	-0.1764	1				
Mor16m	0.381916	0.12386	0.230514	1			
G1m	0.010826	-0.1490	-0.2445	-0.2771	1		
G2u	-0.11083	0.06314	0.03975	0.046	0.3990	1	
Mor31m	-0.42625	-0.5032	0.13451	-0.2718	-0.1075	-0.319	1

**Figure 2: PCA analysis and score plots of the analyzed naphthoquinone derivatives**

obtained with this method was the one related to the linear combination of the selected several descriptors, GATS2v, R3e+, Mor16m, G1m, G2u and Mor31m.

The resulting equation is as follows:

$$\text{INHIBITION} = 45.291 + 8.754\text{GATS2v} + 10.148\text{R3e} + 6.740\text{Mor16m} + 8.322\text{G1m} - 9.424\text{G2u} - 6.474\text{Mor31m} \quad (1)$$

$$N_{\text{train}} = 45; N_{\text{test}} = 12; R^2 = 0.569; \text{RMSE} = 13.827; Q^2 = 0.467.$$

In this equation, N is the number of compounds, R^2 is the squared correlation coefficient, RMSE is the root RMSE, Q^2 is the cross-correlation coefficient evaluation. A higher correlation coefficient and lower root RMSE indicate that the model is more reliable. The QSAR model expressed by Eq. (1) is cross-validated by its appreciable R^2 values ($R^2 = 0.569$), Q^2 values ($Q^2 = 0.467$).

The developed QSAR model reveals that inhibitors of Pathogenic agent IDO1 might be explained by a number of

3D-MorSE and 2D autocorrelations and GETAWAY and WHIM factors. Whereas the negative correlation between the WHIM descriptor (G2u) and the 3D-MorSE descriptor (Mor31m) with inhibition activity indicates that the increase in these values represents a devaluation of inhibition, a positive correlation between the 2D autocorrelations (GATS2v) and GETAWAY descriptor (R3e+) and 3D-MorSE descriptor (Mor16m) and WHIM descriptor (G1m) with inhibition activity indicates an increase in the inhibition value. Based on Equation (1), the mechanism of Pathogenic agent IDO1 activity for the derivatives of naphthoquinone is as follows:

- (1) The inhibitors of pathogenic agent IDO1 activity of the naphthoquinone derivatives decrease with the increase of G2u, Mor31m. Thus, these descriptors are against the Inhibitory activity of the naphthoquinone derivatives.
- (2) The inhibitors of pathogenic agent IDO1 activity of the naphthoquinone derivatives increase with the increase of GATS2v, R3e+, Mor16m, and G1m. Thus, the descriptor is directly related to the Inhibitor activity of the derivatives of the naphthoquinone.

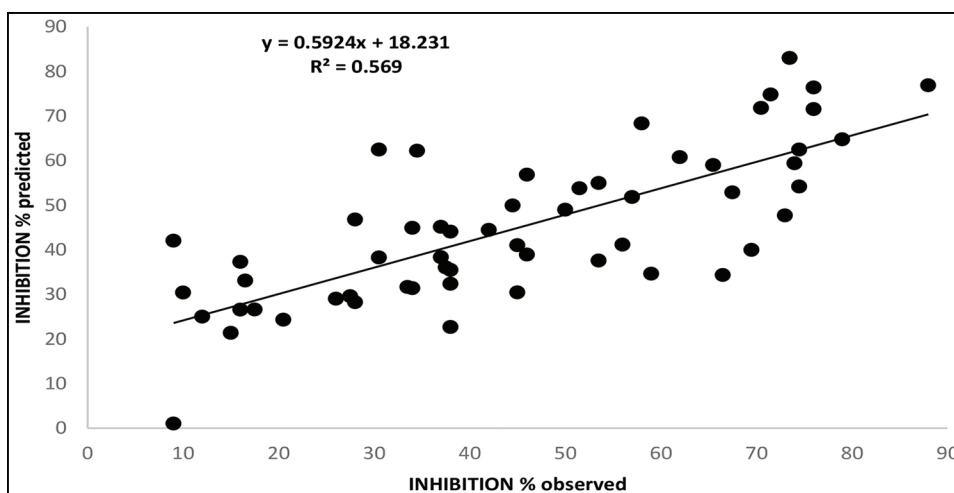


Figure 3: Graphical representation of calculated and observed activity by MLR

Figure 3 shows the correlations of the predicted and observed activities. The descriptors proposed in Eq. (1) by MLR are, therefore, used as the input parameters in ANN. Table 4 presents the predicted values of the inhibition percent of the train set and the test set using the MLR equation.

ANN model results

Among the types of models available, ANNs can produce predictive models of QSAR between the molecular descriptors derived from the MLR model and the activity observed from the compounds. The predicted activity of the compounds was prepared via the ANN model using the properties of several compounds studied. Figure 4 shows the degree of compliance with the anticipated and observed activities.

The squared correlation coefficient (R^2) obtained from the neural network model for this set of naphthoquinone derivatives was calculated as 0.983. Given the acceptable value of this coefficient for the model, it is confirmed that ANN is a superior method for constructing quantitative structure–activity relevance model to predict the desired activity in the compounds mentioned.

In addition, the high value of this coefficient ($R^2 = 0.983$) confirms that the obtained QSAR model can well predict inhibitory activity against pathogenic agent IDO1 for other similar compounds. Table 5 represents the predicted values of inhibition percent for the training set, validation set, and test set using the ANN model.

To assess the predictive power of MLR and ANN models for case activity, we need to use a set of compounds that are different from the training set to create the QSAR model and are not used in model construction. The model MLR established in the computation procedure using the 45 thiazolidinedione derivatives are used to predict the activity of the remaining 12 compounds and model ANN using the 39 thiazolidinedione derivatives are used to predict the activity of the remaining 18 compounds (test = 9 and validation = 9).

The results of regression residual investigation displayed that the error dispensation of these models are unsystematic with normal distribution and have homogenous variance. The agreement observed between the predicted experimental values in Figures 3 and 4 and the random distribution of residuals about zero mean in Figure 5 confirms the high predictive capability of MLR and ANN modeling. Additionally, as can be seen from the figure, the distribution of percent residuals shows that the ANN approach results in fewer prediction errors across the entire data.

Table 6 shows the main performance parameters of the two models. As expected, all statistical parameters for the neural network model are better than those for the MLR model.

In sum, we evaluated the best linear QSAR regression equations specified in this study. As expected, based on the results obtained, the quality evaluation of the MLR model shows that the ANN model is significantly more predictive than the MLR model because the results of ANN method are better than those of MLR model. Therefore, ANN establishes an acceptable relationship between the types of molecular descriptors and the activity of the compounds studied.

Conclusion

In this study, two methods of MLR and ANN were used to predict the inhibitor of pathogenicity of IDO1 with naphthoquinone derivatives. Six types of descriptors were picked to construct the MLR model and the neural network for naphthoquinone derivatives, which include GATS2v, R3e +, Mor16m, G1m, G2u, and Mor31m. The results of the models showed that the model of the neural network is better than the MLR model. To compare the accuracy and prediction of proposed models, key statistical indicators, such as R^2 and RMSE were presented in different models using different statistical tools and descriptors. The results of comparing the two models showed that the ANN model was significantly superior to the MLR model. The slope of the predicted line equations was close to that of

Table 4: Observed and calculated values of inhibition percent according to MLR method for the train and test sets

Number	Inhibition % observed	Inhibition % predicted	Residual	% Relative error
Train set				
1	34	31.36	-2.64	-7.76
2	59	34.6	-24.4	-41.36
3	20.5	24.3	3.8	18.54
4	38	22.64	-15.36	-40.42
5	16.5	33.05	16.55	100.3
6	33.5	31.61	-1.89	-5.64
7	38	44.05	6.05	15.92
8	30.5	38.25	7.75	25.41
9	27.5	29.57	2.07	7.53
10	34.5	62.18	27.68	80.23
11	66.5	34.31	-32.19	-48.41
12	38	32.34	-5.66	-14.89
13	28	46.75	18.75	66.96
14	62	60.74	-1.26	-2.03
15	46	56.81	10.81	23.5
16	79	64.74	-14.26	-18.05
18	56	41.12	-14.88	-26.57
20	51.5	53.76	2.26	4.39
21	37	38.31	1.31	3.54
22	45	30.4	-14.6	-32.44
23	45	40.98	-4.02	-8.93
24	30.5	62.45	31.95	104.75
26	74.5	62.45	-12.05	-16.17
28	53.5	37.55	-15.95	-29.81
29	37.5	36.01	-1.49	-3.97
30	67.5	52.84	-14.66	-21.72
31	74.5	54.15	-20.35	-27.32
33	44.5	49.89	5.39	12.11
34	17.5	26.57	9.07	51.83
35	57	51.8	-5.2	-9.12
36	12	24.96	12.96	108
37	50	48.96	-1.04	-2.08
38	37	45.13	8.13	21.97
40	28	28.21	0.21	0.75
42	88	76.85	-11.15	-12.67
45	65.5	58.97	-6.53	-9.97
46	76	76.37	0.37	0.49
47	71.5	74.8	3.3	4.62
48	73	47.68	-25.32	-34.68
49	70.5	71.76	1.26	1.79
51	69.5	39.94	-29.56	-42.53
52	9	42.02	33.02	366.89
54	16	37.24	21.24	132.75
55	9	1	-8	-88.89
56	10	30.37	20.37	203.7
Test set				
17	58	68.29	10.29	17.74
19	42	44.44	2.44	5.81
25	38	35.46	-2.54	-6.68
27	34	44.88	10.88	32
32	46	38.88	-7.12	-15.48
39	15	21.35	6.35	42.33
41	53.5	54.93	1.43	2.67
43	74	59.35	-14.65	-19.8
44	76	71.51	-4.49	-5.91

Table 4: Continued

Number	Inhibition % observed	Inhibition % predicted	Residual	% Relative error
50	73.5	82.97	9.47	12.88
53	26	29	3	11.54
57	16	26.56	10.56	66

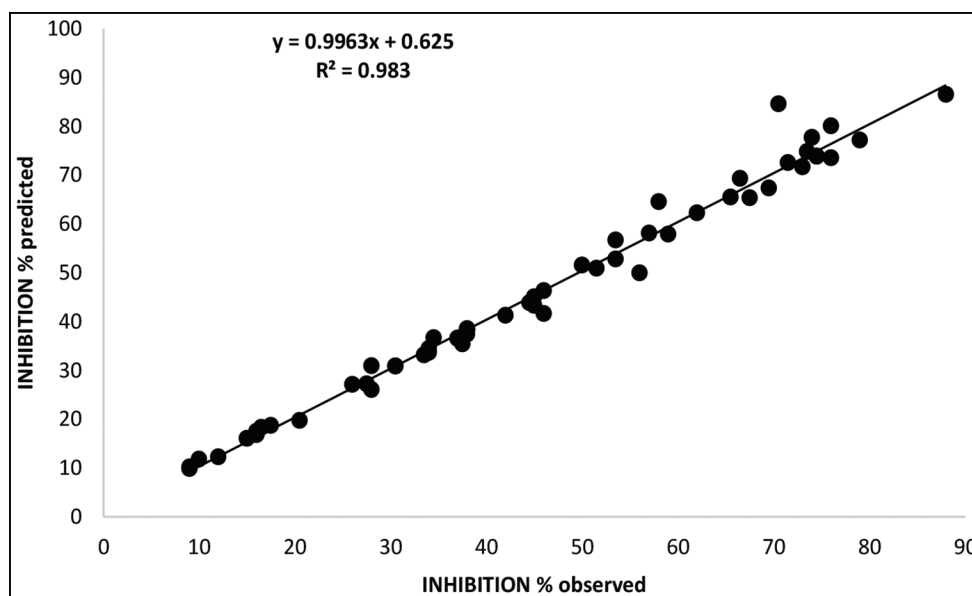


Figure 4: Graphical representation of calculated and observed activity by ANN

$N_{\text{train}} = 39$; $N_{\text{test}} = 9$; $N_{\text{validation}} = 9$; $R^2_{\text{all}} = 0.983$; RMSE = 6.534

Table 5: Observed values and calculated values of Inhibition percent according to ANN method

Number	Inhibition % observed	Inhibition % predicted	Residual	%Relative error
Training set				
3	20.5	19.71	-0.79	-3.85
4	38	37.39	-0.61	-1.61
5	16.5	18.3	1.8	10.91
8	30.5	30.87	0.37	1.21
10	34.5	36.68	2.18	6.32
11	66.5	69.29	2.79	4.2
12	38	38.36	0.36	0.95
15	46	41.6	-4.4	-9.57
17	58	64.54	6.54	11.28
18	56	49.95	-6.05	-10.8
19	42	41.25	-0.75	-1.79
20	51.5	50.89	-0.61	-1.18
21	37	36.47	-0.53	-1.43
22	45	43.28	-1.72	-3.82
24	30.5	30.85	0.35	1.15
25	38	38.52	0.52	1.37
29	37.5	35.4	-2.1	-5.6
30	67.5	65.36	-2.14	-3.17
31	74.5	73.86	-0.64	-0.86
33	44.5	43.85	-0.65	-1.46
34	17.5	18.7	1.2	6.86
35	57	58.09	1.09	1.91
36	12	12.25	0.25	2.08
37	50	51.53	1.53	3.06
39	15	16.04	1.04	6.93

Table 5: Continued

Number	Inhibition % observed	Inhibition % predicted	Residual	%Relative error
40	28	30.93	2.93	10.46
41	53.5	56.68	3.18	5.94
42	88	86.53	-1.47	-1.67
44	76	73.53	-2.47	-3.25
45	65.5	65.48	-0.02	-0.03
46	76	80.07	4.07	5.36
48	73	71.66	-1.34	-1.84
49	70.5	84.58	14.08	19.97
50	73.5	74.81	1.31	1.78
51	69.5	67.34	-2.16	-3.11
52	9	9.84	0.84	9.33
55	9	10.17	1.17	13
56	10	11.77	1.77	17.7
57	16	17.5	1.5	9.38
Validation set				
1	34	33.64	-0.36	-1.06
2	59	57.88	-1.12	-1.9
6	33.5	33.16	-0.34	-1.01
7	38	38.28	0.28	0.74
9	27.5	27.2	-0.3	-1.09
13	28	26.05	-1.95	-6.96
14	62	62.23	0.23	0.37
43	74	77.72	3.72	5.03
47	71.5	72.54	1.04	1.45
Test set				
16	79	77.18	-1.82	-2.3
23	45	45.03	0.03	0.07
26	74.5	73.85	-0.65	-0.87
27	34	34.43	0.43	1.26
28	53.5	52.8	-0.7	-1.31
32	46	46.3	0.3	0.65
38	37	36.6	-0.4	-1.08
53	26	27.11	1.11	4.27
54	16	16.79	0.79	4.94

the experimental value, which was closer to one and expresses the correct prediction for the naphthoquinone derivatives, with a high R^2 value and a low RMSE. Finally, we conclude that descriptor studies using the neural network have a much better ability to predict the inhibitor of the pathogenic IDO1 agent.

Financial support and sponsorship

Nil.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that might influence the work reported in this study. Also, there are no conflicts of interest.

References

- Li H, Chiappinelli KB, Guzzetta AA, Easwaran H, Yen RW, Vatapalli R, *et al.* Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. *Oncotarget* 2014;5:587-98.
- Vinay DS, Ryan EP, Pawelec G, Talib WH, Stagg J, Elkord E, *et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Semin Cancer Biol* 2015;35 Suppl:185-98.
- Munn DH, Mellor AL. IDO in the tumor microenvironment: Inflammation, counter-regulation, and tolerance. *Trends Immunol* 2016;37:193-207.
- Prendergast GC. Immune escape as a fundamental trait of cancer: Focus on IDO. *Oncogene* 2008;27:3889-900.
- Katz JB, Muller AJ, Prendergast GC. Indoleamine 2,3-dioxygenase in T-cell tolerance and tumoral immune escape. *Immunol Rev* 2008;222:206-21.
- Zamanakou M, Germenis AE, Karanikas V. Tumor immune escape mediated by indoleamine 2,3-dioxygenase. *Immunol Lett* 2007;111:69-75.
- Platten M, Wick W, Van den Eynde BJ. Tryptophan catabolism in cancer: Beyond IDO and tryptophan depletion. *Cancer Res* 2012;72:5435-40.
- Théate I, van Baren N, Pilotte L, Moulin P, Larriue P, Renaud JC, *et al.* Extensive profiling of the expression of the indoleamine 2,3-dioxygenase 1 protein in normal and tumoral human tissues. *Cancer Immunol Res* 2015;3:161-72.

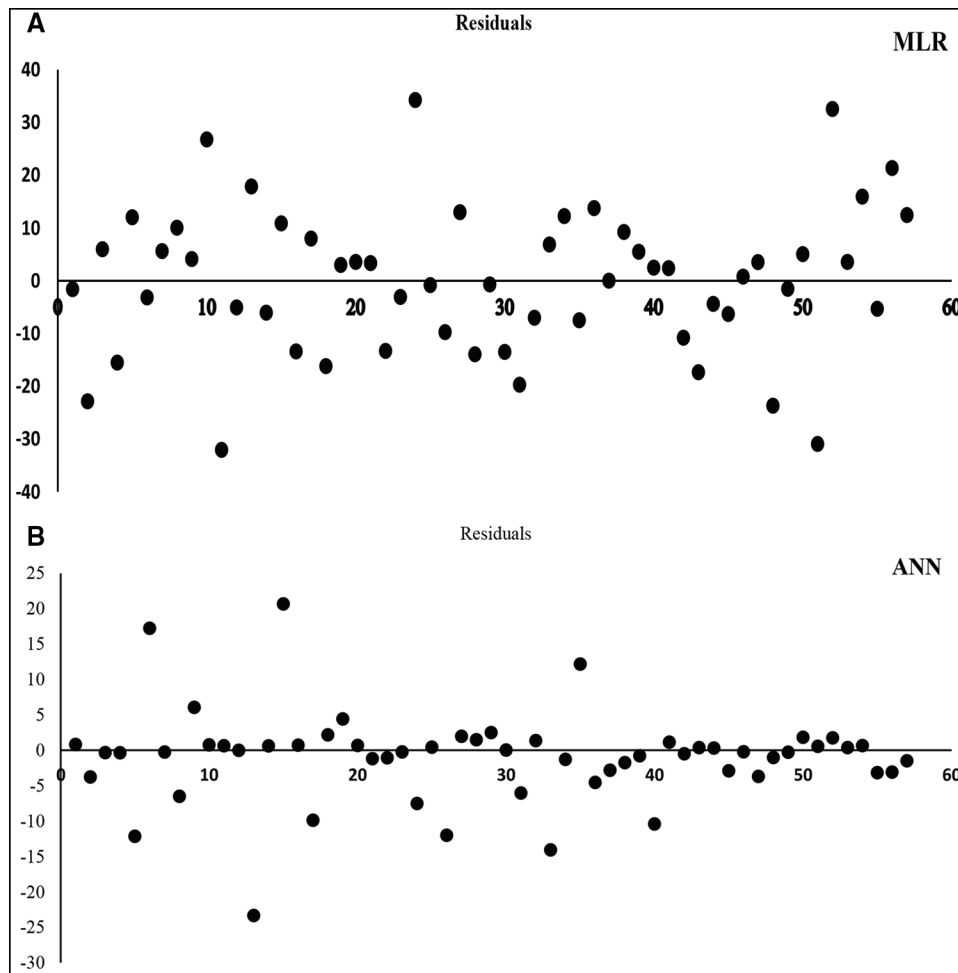


Figure 5: Residual versus experimental values in MLR (a) and ANN (b) models

Table 6: Performance comparison between models obtained by MLR and ANN

Model	R^2	RMSE
MLR	0.569	13.827
ANN	0.983	6.534

- Pallotta MT, Orabona C, Volpi C, Vacca C, Belladonna ML, Bianchi R, *et al.* Indoleamine 2,3-dioxygenase is a signaling protein in long-term tolerance by dendritic cells. *Nat Immunol* 2011;12:870-8.
- Taylor MW, Feng GS. Relationship between interferon-gamma, indoleamine 2,3-dioxygenase, and tryptophan catabolism. *Faseb J* 1991;5:2516-22.
- Metz R, Rust S, Duhadaway JB, Mautino MR, Munn DH, Vahanian NN, *et al.* IDO inhibits a tryptophan sufficiency signal that stimulates mTOR: A novel IDO effector pathway targeted by D-1-methyl-tryptophan. *Oncoimmunology* 2012;1:1460-8.
- Munn DH, Sharma MD, Baban B, Harding HP, Zhang Y, Ron D, *et al.* GCN2 kinase in T cells mediates proliferative arrest and energy induction in response to indoleamine 2,3-dioxygenase. *Immunity* 2005;22:633-42.
- Mezrich JD, Fechner JH, Zhang X, Johnson BP, Burlingham WJ, Bradfield CA. An interaction between kynurenine and the aryl hydrocarbon receptor can generate regulatory T cells. *J Immunol* 2010;185:3190-8.
- Jinushi T, Shibayama Y, Kinoshita I, Oizumi S, Jinushi M, Aota T, *et al.* Low expression levels of microRNA-124-5p correlated with poor prognosis in colorectal cancer via targeting of SMC4. *Cancer Med* 2014;3:1544-52.
- Shirey KA, Jung JY, Maeder GS, Carlin JM. Upregulation of IFN-gamma receptor expression by proinflammatory cytokines influences Ido activation in epithelial cells. *J Interferon Cytokine Res* 2006;26:53-62.
- Song H, Park H, Kim YS, Kim KD, Lee HK, Cho DH, *et al.* L-kynurenine-induced apoptosis in human NK cells is mediated by reactive oxygen species. *Int Immunopharmacol* 2011;11:932-8.
- Turski WA, Wnorowski A, Turski GN, Turski CA, Turski L. Ahr and IDO1 in pathogenesis of COVID-19 and the “systemic ahr activation syndrome:” a translational review and therapeutic perspectives. *Restor Neurol Neurosci* 2020;38:343-54.
- Masoomi Sefiddashti F, Haddadi H, Asadpour S, Ghanavati Nasab S. Prediction of IC50 Values of 2-benzyloxy benzamide Derivatives using Multiple Linear Regression and Artificial Neural Network Methods. *Iran J Math Chem* 2020;11:179-99.
- Norouzian MA, Asadpour S. Prediction of feed abrasive value by artificial neural networks and multiple linear regression. *Neural Comput Appl* 2012;21:905-9.
- Asadpour S, Chamsaz M, Haron MJ. Application of MLR, PLS and artificial neural networks for prediction of GC/ECD retention times

- of chlorinated pesticides, herbicides, and organohalides. *Res J Pharm Biol Chem Sci* 2012;3;21:850-60.
21. Gramatica P, Consonni V, Todeschini R. QSAR study on the tropospheric degradation of organic compounds. *Chemosphere* 1999;38:1371-8. Available from: <https://www.sciencedirect.com/science/article/pii/S0045653598005396>
 22. Zhang L, Lai F, Chen X, Xiao Z. Identification of potential indoleamine 2, 3-dioxygenase 1 (IDO1) inhibitors by an FBG-based 3D QSAR pharmacophore model. *J Mol Graph Model* 2020;99:107628. Available from: <https://www.sciencedirect.com/science/article/pii/S1093326320302497>
 23. Javidfar M, Ahmadi S. QSAR modelling of larvicidal phytochemicals against *Aedes aegypti* using index of ideality of correlation. *SAR QSAR Environ Res* 2020;31:717-39.
 24. Ahmadi S, Ghanbari H, Lotfi S, Azimi N. Predictive QSAR Modeling for the antioxidant activity of natural compounds derivatives based on Monte Carlo method. *Mol Divers* 2021;25:87-97.
 25. Ghiasi T, Ahmadi S, Ahmadi E, Talei Babil Olyai MR, Khodadadi Z. The index of ideality of correlation: QSAR studies of hepatitis C virus NS3/4A protease inhibitors using SMILES descriptors. *SAR QSAR Environ Res* 2021;32:495-520.
 26. Pan L, Zheng Q, Chen Y, Yang R, Yang Y, Li Z, *et al.* Design, synthesis and biological evaluation of novel naphthoquinone derivatives as IDO1 inhibitors. *Eur J Med Chem* 2018;157:423-36.
 27. Habibi-Yangjeh A, Danandeh-Jenagharad M, Nooshyar M. Application of artificial neural networks for predicting the aqueous acidity of various phenols using QSAR. *J Mol Model* 2006;12:338-47.