Int. j. adv. multidisc. res. stud. 2022; 2(1):4-11

Received: 20-11-2021
Accepted: 28-12-2021

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

2D QSAR Studies on an assay of Pyrazole derivatives to inhibit IL4-JAK1-STAT6 pathway to find out herbal remedy for Asthma

[1] SV Manoj, [2] DB Ambili Raj, [3] Sarithamol S, [4] Arya Rajan P, [5] Divya V, [6] Lini K Mathew

[1, 3, 4] Department of Chemistry, Sree Narayana College, Kollam, Kerala, India
[2] Department of Chemistry, Sree Narayana College, Chempazhanthi, Thiruvananthapuram, Kerala, India
[5] Department of Chemistry, MSM College, Kayamkulam, Kerala, India
[6] Department of Chemistry, St.Cyril's College, Adoor, Pathanamthitta, Kerala, India

Corresponding Author: **Sarithamol S**

Abstract

Asthma is characterized as airway inflammation resulting in breathing and chronic coughing problems. Treatments of asthma include inhaled corticosteroids, broncho dilators, DNA vaccines and leukotriene modifiers. But their drawbacks include vomiting, weakness, mental and growth retardation. Asthmatic inflammation can be reduced by the inhibition of JAK (Janus Kinase) family of tyrosine kinases. Our present study focuses on 2D Quantitative Structural Activity Relationship studies on pyrazole derivatives having minimum inhibitory concentration against JAK 1 protein which was identified by Brubaker *et al*. The 2D QSAR model furnished in the study proved its prediction ability (R^2 = 0.6522, p > 0.000, F = 9.548, cross correlation coefficient of the test set molecules Q^2_{test} =0.602542) and showed that positive ionization (POS), hydrogen bond donor sites (HBD) and topological descriptors influenced the prediction of the dataset. The developed 2 D QSAR model was used to check anti-asthmatic activity of 13 phytochemicals taken from Duke's phytochemicals database. Out of these cyanidin exhibited highest activity. Our findings indicate the potency of cyanidin against JAK 1 protein for the inhibition of asthma.

Keywords: Asthma, 2-D QSAR, JAK-STAT Pathway, Regression, Cross-Validation, Correlation Coefficient, Phytochemicals

1. Introduction

Asthma is a chronic lung disease characterized by variable obstruction, airway hyper responsiveness (AHR), airway inflammation and remodeling [1]. Histological studies shows that Cytokines like IL-4 and IL-12 binds to its receptor sites to activate JAK protein which activates STAT (Signal Transducer and Activator of Transcription) through tyrosine phosphorylation [2] .This allows stat protein to dimerise and move to nucleus to mediate change in gene expression by binding specific DNA elements [3]. IL-4-JAK1-STAT6 signaling is responsible for asthma over-expression [4]. Hence the investigation focus on finding JAK1 inhibitors against asthma. There are many treatments for asthma such as corticosteroid inhalers, leukotriene and bronchodilators. But they do have side-effects such as vomiting, restlessness, nausea, growth retardation etc [5]. So, an alternative molecule is required as JAK1 inhibitor against asthma.

QSAR (Quantitative Structure-Activity Relationship) approach helps to correlate biological activity of molecules with physico chemical properties in terms of descriptors. This model can be used to determine the activities of untested compounds quantitatively [6, 7]. A set of 95 molecules from a bioassay having nano molar inhibition towards JAK 1 are taken for the present study. QSAR models were created and statistical studies were done using the software Build QSAR. Superior set (training set) molecules were then used to generate candidate set (test set). About 13 phytochemicals were taken from Dr.Dukes phytochemical database to check their anti-asthmatic property using the model generated.

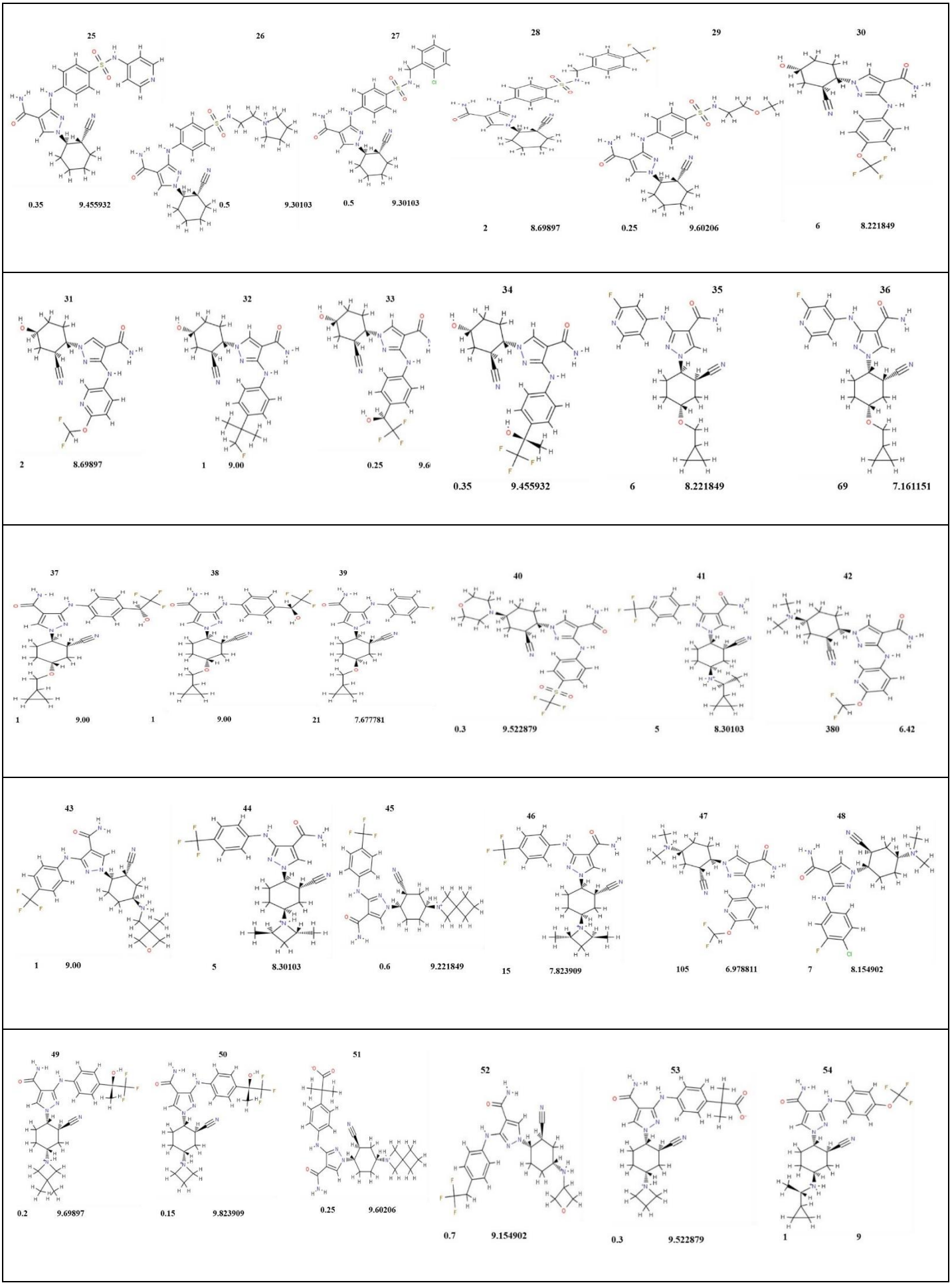
2. Materials and methods

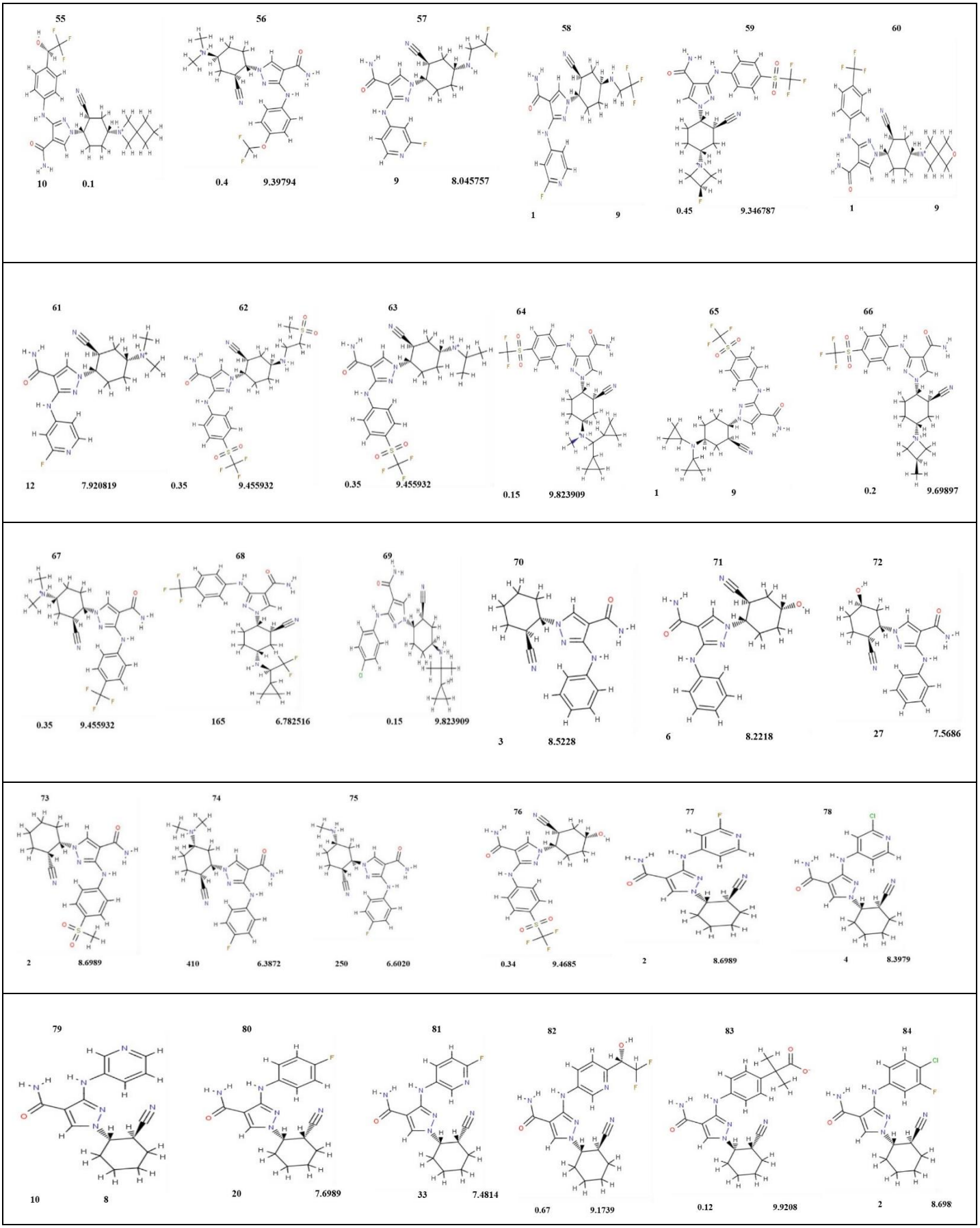
Fig 1: General structure of data set molecules

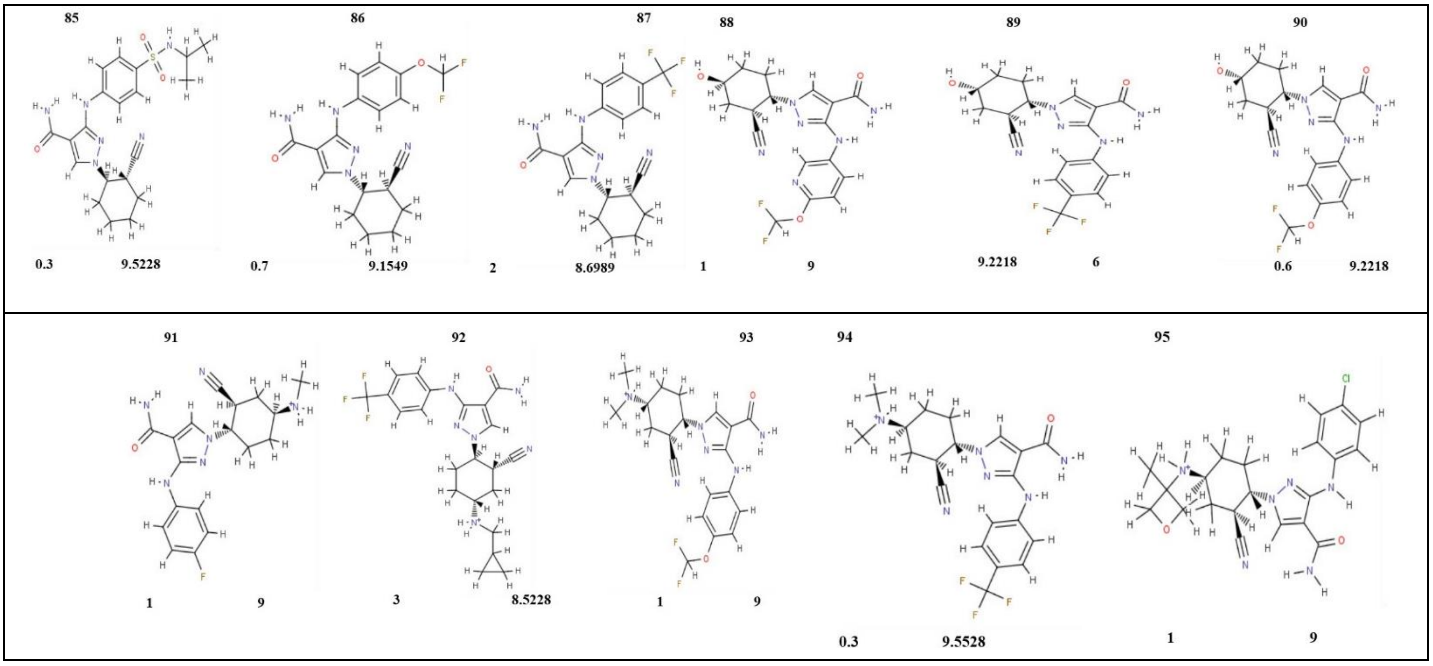
The studies started with choosing 95 pyrazole carboxamide based compounds from Brubaker *et al* [8]. These compounds were randomly divided into 69 training set and 26 test set compounds. Structures of these compounds were used for 2D QSAR studies and using their JAK1 inhibitory activity (IC50-molar concentration of drug leading to 50% inhibition of enzyme integrate) and their negative logarithm (pIC50) were calculated [9, 10] (Table 1). Structures of these compounds were sketched using Marvin Sketch. These compounds were made into single file using Maestro 10.5 interface.

Table 1: Structure and biological activity of training set (1-69) and test set (70-95) compounds and Corresponding IC50 (nM) and pIC50

1	2	3	4	5	6
0.04 10.39794	700 6.154902	305 6.5157	14 7.853872	12 7.920819	5 8.30103
7	8	9	10	11	12
158 6.801343	2 8.69897	2 8.69897	0.35 9.455932	0.17 9.769551	0.36 9.443697
13	14	15	16	17	18
0.6 9.221849	2 8.69897	0.19 9.721246	2 8.69897	18 7.744727	3 8.522879
19	20	21	22	23	24
17 7.769551	41 7.387216	2 8.69897	60 7.221849	6 8.221849	6 8.221849


25
0.35 9.455932
26
0.5 9.30103
27
0.5 9.30103
28
2 8.69897
29
0.25 9.60206
30
6 8.221849
31
2 8.69897
32
1 9.00
33
0.25 9.6
34
0.35 9.455932
35
6 8.221849
36
69 7.161151
37
1 9.00
38
1 9.00
39
21 7.677781
40
0.3 9.522879
41
5 8.30103
42
380 6.42
43
1 9.00
44
5 8.30103
45
0.6 9.221849
46
15 7.823909
47
105 6.978811
48
7 8.154902
49
0.2 9.69897
50
0.15 9.823909
51
0.25 9.60206
52
0.7 9.154902
53
0.3 9.522879
54
1 9


55
10 0.1
56
0.4 9.39794
57
9 8.045757
58
1 9
59
0.45 9.346787
60
1 9
61
12 7.920819
62
0.35 9.455932
63
0.35 9.455932
64
0.15 9.823909
65
1 9
66
0.2 9.69897
67
0.35 9.455932
68
165 6.782516
69
0.15 9.823909
70
3 8.5228
71
6 8.2218
72
27 7.5686
73
2 8.6989
74
410 6.3872
75
250 6.6020
76
0.34 9.4685
77
2 8.6989
78
4 8.3979
79
10 8
80
20 7.6989
81
33 7.4814
82
0.67 9.1739
83
0.12 9.9208
84
2 8.698



2.1 Descriptor calculation

Single file is then fed into PowerMV to generate descriptors like pharmacophore fingerprints, weighted burden number, topological indexes. 2D auto-correlation descriptors are obtained from Marvin Sketch and E-dragon softwares. Descriptor set with higher correlation coefficient (r^2=0.6522) were chosen and others were discarded [11].

2.2 Regression analysis

Selection of training set and test set molecules was done on the basis of having sufficient diversity in structure and covering wide range of biological activity [12]. Data set was divided into 69 training set and 26 test set compounds through activity sampling [13]. Training set and test molecules were then subjected to MLR (Multiple Linear Regression) analysis to predict anti-asthmatic activity of test set [14].

2.3 Validation using test set

QSAR models were used to calculate squared correlation coefficient (r^2), adjusted squared correlation coefficient (r_{adj}), Kubinsyi function (F), standard deviation of sum of square of difference between predicted and observed value (S_{PRESS}), standard deviation of error prediction (S_{DEP}) and cross validation coefficient (Q^2) of test set [15]. These data were finally used to evaluate r^2_{pred} of both set of compounds. A QSAR model is predictive if following conditions fulfill:

$$r^2>0.6,\ Q2>0.6,\ r_{pred}2>0.5,\ p<0.0001$$

Cross validation coefficient was calculated using the equation;

$$Q^2=1-\{\sum (y_{obs}-y_{pred})^2/\sum (y_{obs}-y_m)^2\}$$

Where y_{obs} and y_{pred} are the observed and predicted activities of the test molecules respectively and y_m is the mean activity of training set molecules. External validation was also done to check the predictive power of the model created and pred_r^2 values were calculated [16].

$$pred\_r^2=1-\{\sum (y^t_{obs}-y^t_{pred})^2/\sum (y^t_{obs}-y_m)^2\}$$

2.4 Prediction of activity of phytochemicals against asthma

The generated model was used to predict the inhibitory property of phytochemicals obtained from Duke's Phytochemical database [17] against JAK 1. Their properties as well as their activity were calculated and are shown in Table 2.

Table 2: Phytochemicals and their descriptors

Compounds	POS_03_HYP	IDDE	IDMT	IVDE	IDDM	IDET	IDM	IAC	HBD_03_ARC	ATS4e
Alpha Boswellic Acid	0	4.74	23645.31	1.95	5.02	1783.82	8.82	93.49	0	4.39
Androsin	0	4.05	9328.21	1.57	4.50	829.97	7.76	64.29	0	3.75
Beta Boswellic	0	4.80	23441.55	1.95	5.02	1773.20	8.83	93.49	0	4.43
Corilagin	0	5.12	64959.83	1.49	5.46	3768.43	9.73	104.88	0	4.77
Glycyrrhetic acid	0	4.59	25579.56	1.96	5.06	1903.03	8.92	96.46	0	4.43
Glycyrrhetinic Acid	0	4.59	25579.56	1.96	5.06	1903.03	8.92	96.46	0	4.43
Glycyrrhizin	0	5.42	145050.6	1.89	5.83	6916.44	10.44	169.19	0	5.03
Lobeline	0	3.44	13039.08	1.32	4.62	1062.78	7.98	67.93	1	3.60
Oxymatrine	0	4.04	4232.45	1.58	4.23	454.49	7.25	60.69	0	3.61
Papaverine	0	4.24	12262.38	1.44	4.62	1027.20	7.99	67.41	0	3.77
Piperlongumine	0	3.82	9929.23	1.51	4.50	854.43	7.75	64.67	0	3.70
Quercitine	0	3.97	7533.79	1.54	4.43	717.25	7.64	48.53	1	3.81
Cyanidin	0	4.39	6700.15	1.55	4.37	654.48	7.49	47.83	0	3.63

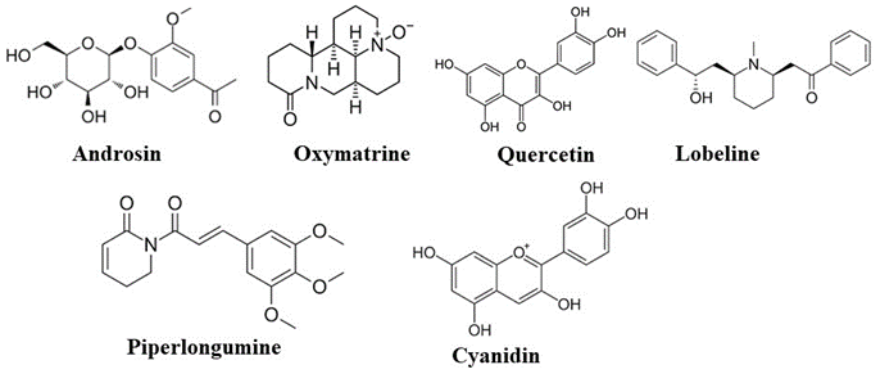


Fig 2: Structure of a few phytochemicals having high predictivity:

3. Results and discussion

In the present study more than 5000 descriptors were generated using the tools power MV and E-dragon. They included topological descriptors, pharmacophore fingerprints, weighted burden, 2D correlation, electronic, steric, lipophilic descriptors etc. MLR regression analysis was done on the dataset using Build QSAR software.10 statistically significant descriptors were identified from 5000 descriptors by the continuous evaluation (table 3).2D QSAR equation was developed using the significant descriptors. (Table 4)

Table 3: List of descriptors used in the study

Sl.No	Descriptor	Definition
1	POS_03_HYP	Positive ionization due to hyperconjugation
2	IDDE	Mean information content on the distance degree equality
3	IDMT	Total information content on the distance magnitude
4	IVDE	Mean information content on the vertex degree equality
5	IDDM	Mean information content on the distance degree magnitude
6	IDET	Total information content on the distance equality
7	IDM	Mean information content on the distance magnitude
8	IAC	Mean information content on the distance magnitudetotal information index on atomic composition information indexes
9	HBD_03_ARC	Hydrogen bond donor that is 3 bond away from aromatic ring
10	ATS4e	Broto-moreau auto-correlatio of lag4(log function) weighted by Sanderson electronegativity

Table 4: Selected 2D QSAR equation

$$Y_{cal} = +0.6679X_1 + 0.7581X_2 - 0.0013X_3 + 1,9969X_4 + 106.0597X_5 + 0.0350X_6 - 68.7779X_7 + 0.0604X_8 + 0.2945X_9 - 3.5740X_{10} + 52.3673$$

Using the above equation, observed and calculated activities of training and test sets were calculated and a graph was plotted between them as shown in Fig 3.

Table 5: Observed and calculated activities of training set molecules (as per the equation)

Compound	Observed activity (y_{obs})	Calculated activity (y_{cal})	Compound	Observed activity (y_{obs})	Calculated activity (y_{cal})
1	10.39	9.67	36	7.16	7.71
2	6.15	6.72	37	9.00	9.05
3	6.51	6.82	38	9.00	9.05
4	7.85	7.73	39	7.67	7.82
5	7.92	7.97	40	9.52	9.36
6	8.30	7.77	41	8.30	8.59
7	6.80	7.57	42	6.4	8.39
8	8.69	7.63	43	9.00	9.10
9	8.69	8.46	44	8.30	8.32
10	9.45	9.35	45	9.22	9.77
11	9.77	8.62	46	7.82	8.32
12	9.44	9.30	47	6.97	8.39
13	9.22	9.30	48	8.15	8.00
14	8.69	9.59	49	9.69	9.90
15	9.72	9.46	50	9.82	9.67

16	8.69	8.29	51	9.60	9.85
17	7.74	8.29	52	9.15	8.73
18	8.52	8.29	53	9.52	8.89
19	7.77	7.87	54	9.00	8.61
20	7.38	8.53	55	10.00	10.12
21	8.69	8.53	56	9.39	8.41
22	7.22	7.97	57	8.04	7.80
23	8.22	8.21	58	9.00	8.46
24	8.22	8.10	59	9.34	9.69
25	9.45	8.71	60	9.00	9.17
26	9.30	9.33	61	7.92	7.37
27	9.30	9.35	62	9.45	9.34
28	8.69	9.15	63	9.45	9.53
29	9.60	9.52	64	9.82	8.88
30	8.22	9.16	65	9.00	8.38
31	8.69	8.56	66	9.69	9.77
32	9.00	9.14	67	9.45	8.51
33	9.60	9.27	68	6.78	7.90
34	9.45	9.56	69	9.82	8.60
35	8.22	7.71			

Table 6: Observed and calculated activity of test set

Compound	$Y_{observed}$	$Y_{calculated}$	Compound	$Y_{observed}$	$Y_{calculated}$
1	8.52	8.10	14	9.92	9.32
2	8.22	7.92	15	8.69	8.17
3	7.56	7.49	16	9.52	9.73
4	8.69	9.47	17	9.15	9.22
5	6.38	7.19	18	8.69	9.14
6	6.60	7.43	19	9.00	9.03
7	9.46	9.88	20	9.22	9.36
8	8.69	7.84	21	9.22	9.08
9	8.39	7.87	22	9.00	8.32
10	8.00	8.10	23	8.52	9.26
11	7.69	8.52	24	9.00	9.07
12	7.48	8.53	25	9.52	9.00
13	9.17	9.55	26	9.00	9.13

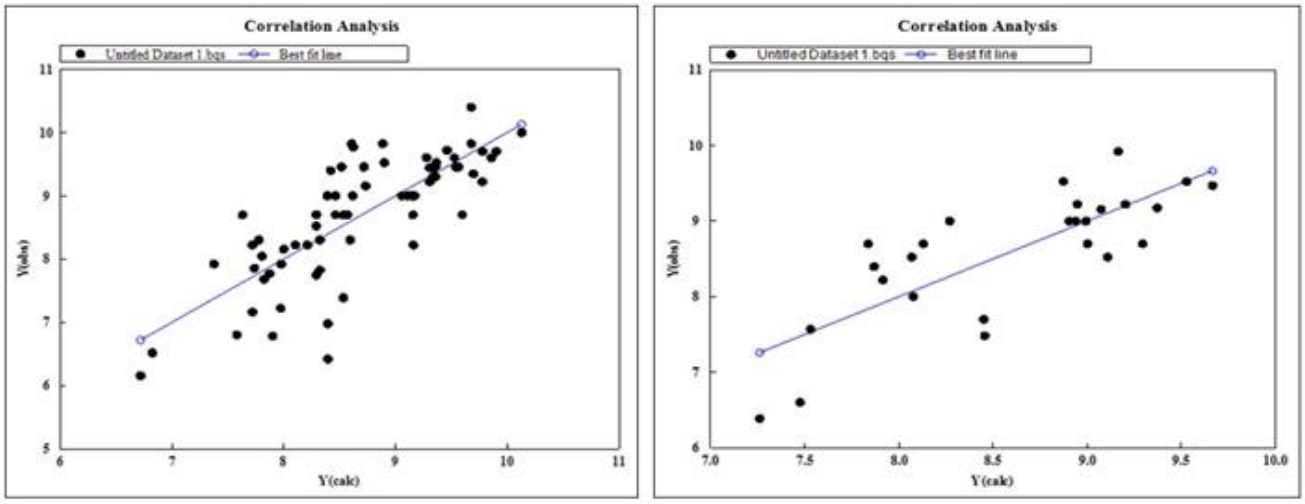


Fig 3: Y_{obs} v/s Y_{cal} (training set), Y_{obs}v/s Y_{cal} (test set)

Table 7: Regression statistics

R^2	R_{adj}^2	Q^2_{test}	F	S PRESS	S DEP	p value
0.6522	0.5923	0.602542	9.548	0.788	0.7275	0.0000

Prediction of biological activity of phytomolecules

Around 13 phytochemicals were taken from Duke's phytochemicals database which are biologically active. These phytochemicals were subjected to predict their activity against JAK 1 using the above linear equations. The phytochemical, cyanidin, shows the highest activity.

Table 8: Predicted JAK1 inhibitory activity of phytochemical compounds

Compounds	JAK1 inhibitory activity predicted (Yp)
Cyanidin	10.48
Piperlongumine	10.45

Lobeline	10.04
Oxymatrine	9.83
Androsin	9.04
Papaverine	8.73
Alpha Boswellic Acid	6.64
Quercitine	8.01
Beta Boswellic	6.36
Corilagin	6.46
Glycyrrhetic acid	6.42
Glycyrrhetinic Acid	6.42
Glycyrrhizin	6.56

4. Conclusion

The QSAR model furnished in the present study has proved its prediction ability through the results as R^2= 0.6522, p =0.000, F=9.548. and cross correlation coefficient of the test set molecules is Q^2_{test}=0.602542. Observed and calculated values have showed good agreement with each other for training and test set taken. From this we have concluded that descriptors such as hydrogen bond donor, and bond distance are highly influenced on the biological activity of molecules present in the dataset. This information can make useful for the designing of new candidates using the present QSAR model against asthma. QSAR model prediction on 13 phytochemicals taken from Duke's Phytochemicals database showed that cyanidin has highest inhibitory activity for asthma. The anti-inflammatory property of cyanidin can be used for designing a better anti-asthmatic candidate. Hence, in silico study has revealed that Cyanidin is a potent JAK1 inhibitor against asthma. The scaffold of this compound can be considered for designing better anti asthmatic candidate.

5. References

1. Heijink IH, Van Oosterhout AJM. Strategies for targeting T-cells in allergic diseases and asthma. Pharmacol. Ther. 2006; 112(2):489-500.
2. Ngoc PL, Gold DR, Tzianabos AO, Weiss ST, Celedón JC. Cytokines, allergy, and asthma. Curr. Opin. Allergy Clin. Immunol. 2005; 5(2):161-166.
3. Barnes PJ. Th2 cytokines and asthma: An introduction. Respir. Res. 2001; 3(2):64-65.
4. Williams NK, *et al*. Dissecting Specificity in the Janus Kinase: The Structures of JAK-Specific Inhibitors Complexed to the JAK1 and JAK2 Protein Tyrosine Kinase Domains. J. Mol. Biol. 2009; 387(1):219-232.
5. Luo C, Laaja P. Inhibitors of JAKs/STATs and the kinases: A possible new cluster of drugs. Drug Discov. Today. 2004; 9(6):268-275.
6. Baig MH *et al*. Computer Aided Drug Design: Success and Limitations Structure of target protein Propose new lead or optimize existing lead. Curr. Pharm. Des.2016; 22(5):572-581.
7. Shibi IG, Aswathy L, Jisha, RS, Masand VH, Divyachandran A, Gajbhiye. Molecular docking and QSAR analyses for understanding the antimalarial activity of some 7-substituted-4-aminoquinoline derivatives. Eur. J. Pharm. Sci. 2015; 77:9-23.
8. Brubaker J, Dinsmore C, Hoffman DM, Jung J, Liu D, Peterson S, *et al*. Cycloalkylnitrile pyrazole carboxamides as Janus Kinase inhibitors. 2003; 1(12):1-4. U.S. Patent 8,962,608, B2 Feb. 24, 2015.
9. Divya V, Pushpa VL, Manoj KB. Cyclin dependent kinase 4 inhibitory activity of Thieno [2, 3-d] pyrimidin-4-ylhydrazones e Multiple QSAR and docking studies. J. Mol. Struct. 2019; 1183:263-273.
10. Talkit KM, Mahajan DT, Masand VH. Physicochemical properties of soybean oil and their blends with vegetable oils for the Evaluation of Lubrication Properties. J. Chem. Biol. Phys. Sci. 2013; 3(1):490-497.
11. Masand VH, Toropov AA, Toropova AP, Mahajan DT. QSAR Models for Anti-Malarial Activity of 4-Aminoquinolines. Curr. Comput. Aided. Drug Des. 2014; 10(1):75-82.
12. Masand VH, Mahajan DT, Nazeruddin GM, Hadda TB, Rastija V, Alfeefy AM. Effect of information leakage and method of splitting (rational and random) on external predictive ability and behavior of different statistical parameters of QSAR model. Med. Chem. Res. 2015; 24(3):1241-1264.
13. Divya V, Pushpa VL, Sarithamol S, Manoj KB. Journal of Molecular Graphics and Modelling Computational approach for generating robust models for discovering novel molecules as Cyclin Dependent Kinase 4 inhibitors. J. Mol. Graph. Model. 2018; 82:48-58.
14. Rastija V, Nikolić S, Masand VH. Quantitative relationships between structure and lipophilicity of naturally occurring polyphenols. Acta Chim. Slov. 2013; 60(4):781-789.
15. Divya V, Pushpa VL, Sarithamol S, Manoj KB. Computational approach for generating robust models for discovering novel molecules as Cyclin Dependent Kinase 4 inhibitors. J. Mol. Graph. Model. 2018; 82:48-58.
16. Mahajan DT, Masand VH, Patil KN, Ben Hadda T, Rastija V. Integrating GUSAR and QSAR analyses for antimalarial activity of synthetic prodiginines against multi drug resistant strain. Med. Chem. Res. 2013; 22(5):2284-2292.
17. Duke JA, Dr. Duke's Phytochemical and Ethnobotanical Databases, 1992.