International Journal of ChemTech Research
CODEN(USA): IJCRGG ISSN : 0974-4290
Vol.4, No.4, pp 1350-1360, Oct-Dec 2012

2D QSAR Study of Substituted 2-Phenyl-Benzimidazole Derivatives as Potent Anti Allergic Agents

Navin Raj*, Sanmati K. Jain

SLT Institute of Pharmaceutical Sciences, GGU, Bilaspur-495009, Chhatisgarh, India.

***Corres.author: pharmanavin@rediffmail.com**
Ph: +919826846661

Abstract: A set of forty one substituted-2-phenyl-benzimidazole derivatives with anti-allergic activity was subjected to the two dimensional quantitative structure activity relationships studies using V life molecular design suite 3.5. Drug Designing module contain various combinations of thermodynamic, electronic, topological and spatial descriptors.Substituted-2-phenyl-benzimidazole was taken as the lead molecule and QSAR model developed using partial least square regression approach. For each set of descriptors, the best multilinear.QSAR equations were obtained by the stepwise forward backward variable selection method. Logarithmic inverse value of IC_{50} was taken as dependent variable and various physico-chemical descriptors were taken as independent variable. The best QSAR model *i.e.* model-2(r^2 = 0.7107, Fischer's test value F=17.8671, r^2 se = 0.4900) has acceptable statistical quality and predictive potential as indicated by the value of cross validated squared correlation coefficient (q^2=0.6230) and r^2 for external test set (pred_r^2 = 0.7256) .From the build model it seems to be clear that SssNHE- index, IPC Average and SaasN(Noxide)E-index contributes negatively and $SssCH_2E$-index,slogp and chi6chain contributes positively. Thus this validated model brings important structural insight to aid the design of more potent anti-allergic agents.
Keywords: Anti allergic, substituted-2-phenyl-benzimidazole derivative, PLS.

1. INTRODUCTION

An allergy is one of the immune dysfunctions, which is a serious health problem worldwide. Substances that cause allergy are called allergens, such as pollen, food, dust mites, cosmetics, mold spores and animal hairs [1].

In recent years, the rapid increase in the number of patients suffering from a variety of allergic symptoms has become of grave concern[2].Mast cells are mononuclear and granule-containing secretary cells. They are important in the development of many physiological changes during allergic and anaphylactic responses[3].Degranulation of mast cells caused by antigen–antibody reactions triggers type I allergic diseases such as bronchial asthma, allergic rhinitis, atopic dermatitis, and pollenosis [4,5].

Allergies such as asthma, allergic rhinitis, anaphylaxis and atopic dermatitis have been well known as immediate reactions following contact with certain exogenic allergens [6]. Allergies in humans are characterized by the appearance in serum and tissues of the immunoglobulin E isotype (IgE) directed toward specific environmental antigens. The studies on the cellular basis of IgE regulation have provided important insights into a disease process that affects a considerable proportion of the population worldwide. More recently, the molecular events underlying IgE synthesis has been actively investigated[7].

Allergen-induced IgE synthesis is a central feature of allergic disorders. Subsequent interactions

between IgE and allergen through a variety of mechanisms leads to the cross linking of the bound IgE, triggering the release of pharmacological active mediators such as platelet-activating factor (PAF), leukotrienes, vasoactive amines from mast cells and basophiles. These mediators cause smooth muscle contraction, increased vascular permeability and vasodilatation[8]. IgE production is regulated by CD23, a cell-surface molecule with a variety of activities also known as the low-affinity receptor for IgE (FcЄRII) [9-12].

IgE-mediated allergic diseases, called atopy, include allergic rhinitis, ophthalmia, asthma, dermatitis, drug and food allergy, anaphylaxis, and hyper-IgE syndrome. The suppression of undesirable IgE-mediated inflammation and immune reactions is believed to be useful for treatment of atopic disorders. Although there are many anti allergic remedies such as histamine antagonists and mast cell stabilizers available in the market at present, they are not very effective against allergic disorders because they impact the disease state by targeting a single mediator that modifies a response at the target organ thus unable to prevent immunization against an allergenic antigen[13, 14].

There are a number of pharmacological agents available for the treatment of allergic conditions such as asthma and allergic rhinitis. For a drug to be effective against allergic conditions, an action on a target that influences multiple mediators within the allergy cascade is required [15]. Activity of IgE interfering compounds against a diverse group of allergy mediators will provide us to establish compounds as powerful tools for the treatment of allergy based diseases. Consequently novel approaches are needed as efficient search for useful candidates to be screened as anti-allergic drugs, so that few novel therapeutic candidates can be discovered which can serve the purpose best.

Benzimidazole derivatives are of wide interest because of their diverse biological activity and clinical applications. This ring system is present in numerous antiparasitic, fungicidal, anthelmintic and anti-inflammatory drugs[16–19].

One could not, however, confirm that the compounds designed would always possess good inhibitory activity. Consequently, it is of interest to develop a prediction method for biological activities before the synthesis. Quantitative structure activity relationship (QSAR) searches information relating chemical structure to biological and other activities by developing a QSAR model. Using such an approach one could predict the activities of newly designed compounds before a decision is being made whether these compounds should really be synthesized and tested.

Accordingly, there is an urgent need to design and screen anti allergic agents with higher bioactivities. There is a need to analyze the correlation between *in-vitro* IgE inhibitory activity and physico-chemical parameters using the Quantitative Structure Activity Relationship (QSAR) methods because the quantitative analysis of such molecules can be utilized for increasing the potency and minimizing the side effects. In order to study and presume a correlation between structure and biological activity of substituted 2-phenyl-benzimidazole as anti allergic agents, we have developed QSAR models.

2. MATERIALS AND METHODS

2.1 Chemical Data

A series of 53, substituted 2-phenyl-benzimidazole derivatives having *in- vitro* IgE activity were selected from the literature[20]. The above reported substituted 2-phenyl-benzimidazole derivatives showed wide variation in their structure and potency profiles.

2.2 Biological activities

The experimental biological activities, in the form of IC_{50} (nM) were converted into pIC_{50} thus correlating the data linear to the free energy change and used as a dependant variable for the development of a valid 2D-QSAR models. Since some compounds exhibited insignificant activity, hence excluded from the present study, hence study is concerned with 41 compounds. Table 1 shows the list of 41 such compounds along with their biological activity data.

2.3 2D-QSAR Methodology

Two dimensional quantitative structure–activity relationship studies of substituted 2-phenyl-benzimidazole were carried out by using V Life Molecular Design Suite software version 3.5 running on Intel core 2 duo processor and windows XP operating system[21].The molecular structures of the compounds in selected series were drawn in 2D-APPL mode of software, exported to QSAR Plus through MDS path and converted to 3D model. Energy minimization[22] and geometry optimization was conducted using Merck Molecular Force Field (MMFF) method[23-26] with Root Mean Square (RMS) gradient set to 0.01 Kcal/mol Å and iteration limit to 10000.

2.4 Molecular Descriptors

The optimized geometries of the molecules were used to compute the necessary quantum chemical 2

D descriptors like element counts, molecular weight, molecular refractivity, Baumann alignment independent topological descriptors etc., available in the V-Life MDS 3.5. The invariable (constant) columns of independent variables (*i.e*., descriptors) were removed and later used for QSAR analysis.

2.5 Selection of Training and Test Sets

Training and test sets were selected using random selection[27] and manual selection methods.

2.5.1 Random selection: In order to build and validate the QSAR models, both internally and externally, the data sets were divided into training [90%-60% (90%, 85%, 80%, 75%, 70%,65%and 60%) of total data set] and test sets [10%-40% (10%, 15%, 20%,25%, 30%, 35% and 40%) of total data set]in a random manner. 10 trials were run in each case.

2.5.2 Manual data selection: The whole range of activities were sorted by arranging biological activity in ascending order and every 3rd - 8th compound assigned to the test set[28].

Variable-selection for the QSAR modeling was carried out by stepwise forward–backward partial least square regression method (PLS) using statistical program of V-Life MDS 3.5.with pIC50 values as dependent variable and various 2D descriptors calculated for the molecules as independent variables. The best model was selected on the basis of various statistical parameters such as squared correlation coefficient (r^2), cross validated square correlation coefficient (q^2), standard error of estimation (SE) and sequential Fischer's test (F)[29].

3. RESULTS AND DISCUSSION

Out of the various trials performed by random selection method , training sets of 85% and 90 % fashioned into significant results with training set of 34 and test set of 7 and training set of 35 and test set of 6, substituted 2-phenyl-benzimidazole derivatives respectively having different substitution, were employed. Again out of the various trial performed by manual training and test sets selection method training set of 33 and test set of 8, substituted 2-phenyl-benzimidazole derivatives fashioned into a significant model.

Following statistical measure was used to correlate biological activity and molecular descriptors; n ,number of molecules; df ,degree of freedom; r^2 ,coefficient of determination; q^2 , cross validated r^2; pred_r^2 , r^2 for external test set; pred_r^2se , coefficient of correlation of predicted data set; Z score, Z score calculated by the randomization test; best_rand_r^2; best_rand_q^2,highest q^2 value in the randomization test; α, statistical significance parameter obtained by the randomization test. Selecting training and test set by random and manual selection method, unicolumn statistics shows that the max of the test is less than max of training set and the min of the test set is greater than of training set. The data for unicolumn statistics is shown in Table 2, which is prerequisite analysis for further QSAR study. The above result shows that the test is interpolative i.e. derived within the min-max range of the training set. The mean and standard deviation of the training and test provides imminent to the relative difference of mean and point density distribution of the two sets.

3.1 Generation of QSAR models

The dataset of 41 molecules were used in the present study. The common structure of substituted 2-phenyl-benzimidazole is shown in Fig. (1).

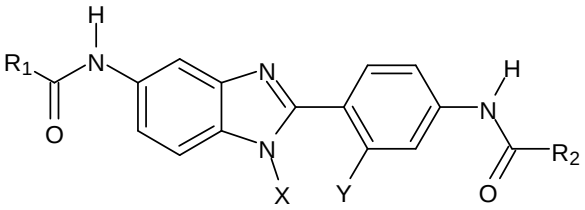


Fig.1. Parent structure of substituted-2-phenyl-benzimidazole derivatives

MODEL – 1 [Random selection method (85 %)]

After 2D QSAR study by Partial Least Squares method using stepwise forward-backward variable selection method, the final QSAR equation developed as follows:

pIC_{50} = 0.171973 $SssCH_2E$ -index + 0.44182 slogp – 1.33303 SssNHE- index – 9.57044 e^{-009} IPC Average + 2.64192 Chi6chain + 11.2293

Values of different statistical parameters generated in model during 2D QSAR studies is shown in Table – 3

The observed and predicted pIC_{50} values are shown in Table 1.

The plot of observed vs. predicted activity is shown in Fig. (2).From the plot it can be seen that the model is quite able to predict the activity of training set (maximum points are close to regression line) as well as external.

The descriptors which contribute for the pharmacological action are shown in Fig. (3).The major group of descriptors involved sub groups like $SssCH_2E$ -index, slogp, SssNHE- index, IPC Average and Chi6chain, help in understanding the effect of substituent at different position of substituted 2-phenyl-benzimidazole.

Table 1. Substituents, Experimental and Predicted activity of substituted-2-phenyl-benzimidazole derivatives used in Training and Test Set

Comp. Name	R_1	R_2	X	Y	IgE *in-vitro* IC_{50} (nM)	IgE *in- vitro* pIC_{50} (M)				
						Experimental activity	Predicted activity			
							Model-1	Model-2	Model-3	Model-4
1a	Phenyl	Phenyl	H	H	20	7.698970	6.95114	7.05533	7.05495	7.03650
1b	4-Bromophenyl	4-Bromophenyl	H	H	200	6.698970	7.39378	7.48988 [T]	7.45522	7.45406 [T]
1e	3-Cholorophenyl	3-Cholorophenyl	H	H	25	7.602060	7.26936 [T]	7.36286	7.33245	7.33140
1f	2-Cholorophenyl	2-Cholorophenyl	H	H	45	7.346780	7.20713	7.29673	7.26763	7.26850
1g	3,4-Dicholorophenyl	3,4-Dicholorophenyl	H	H	40	7.397940	7.61164	7.69528	7.63550	7.65015 [T]
1h	2,3-Dicholorophenyl	2,3-Dicholorophenyl	H	H	10	8.00000	7.50696	7.58392	7.52610	7.54398 [T]
1i	3,5-Dicholorophenyl	3,5-Dicholorophenyl	H	H	70	7.154900	7.58395	7.66602	7.60721	7.62273
1j	2,4-Dicholorophenyl	2,4-Dicholorophenyl	H	H	30	7.522870	7.55265 [T]	7.63265	7.57424	7.59073
1k	2,6-Dicholorophenyl	2,6-Dicholorophenyl	H	H	400	6.397940	7.45405	7.52786	7.47156	7.49108
1m	Penta -fluoro-phenyl	Penta -fluoro-phenyl	H	H	4	8.397940	8.43683	8.58029	8.43978	8.41437
1n	Phenyl	4-Cholorophenyl	H	H	90	7.045750	7.12959	7.22966	7.21390 [T]	7.20355 [T]
1o	4-Nitrophenyl	4-Nitrophenyl	H	H	150	6.823900	6.94474	7.04917 [T]	7.02692	7.01008
1q	4-Cyanophenyl	4-Cyanophenyl	H	H	100	7.000000	6.78718	6.88583 [T]	6.87279	6.85977
1r	4-Methoxyphenyl	4-Methoxyphenyl	H	H	30	7.522870	6.77226	6.8636	6.84920	6.84132
1s	3,4-Dimethoxyphenyl	3,4-Dimethoxyphenyl	H	H	700	6.154900	6.15870	6.20123	6.16513	6.17960
1v	4-S-methyl-phenyl	4-S-methyl-phenyl	H	H	150	6.823900	7.22109	7.30694	7.27347	7.27614
1w	4-Methyl Phenyl	4-Methyl Phenyl	H	H	20	7.698970	7.04026	7.13565	7.11467	7.10889 [T]
1y	1-Naphthalene	1-Naphthalene	H	H	80	7.096910	6.71713	6.72731	6.69929	6.71260
1z	CH_2-2-thiophene	CH_2-2-thiophene	H	H	500	6.301030	6.35182	6.40688	6.35031	6.37432

1aa	Cyclohex-3-ene	Cyclohex-3-ene	H	H	40	7.397940	7.56161	7.61942	7.62825	7.57924
1cc	Phenyl	Cyclohexyl	H	H	10	8.000000	7.78846	7.85379	7.85699	7.80371
1dd	CH_3	Cyclohexyl	H	H	100	7.000000	7.19492	7.25509	7.24364	7.19987
1ee	3,4-Dichlorophenyl	Cyclohexyl	H	H	0.8	9.096910	8.13311 T	8.18909 T	8.16286	8.12567
1ff	4-Chlorophenyl	Cyclohexyl	H	H	6	8.221840	7.96794 T	8.02913	8.01696	7.97173 T
1gg	Cyclohexyl	3,4-Dichlorophenyl	H	H	0.4	9.397940	8.12738	8.18353	8.15728	8.12035
1hh	Cyclohexyl	4-Chlorophenyl	H	H	8	8.096910	7.96222	8.02357	8.01138 T	7.96641
1ii	1-Adamantyl	2-Fluorophenyl	H	H	10	8.000000	8.50397	8.56463	8.58815	8.50549
1jj	1-Adamantyl	4-Fluorophenyl	H	H	10	8.000000	8.37538 T	8.42785	8.45451 T	8.37578
1kk	2-Pyridyl	1-Adamantyl	H	H	06	8.221840	8.12603	8.18014	8.22804	8.14171
1ll	3-Pyridyl	1-Adamantyl	H	H	20	7.698970	8.09549	8.14768	8.19628	8.11089
1mm	Cyclohexyl	Cyclohexyl	H	H	4	8.397940	8.63464	8.66087 T	8.66768	8.57916
1nn	1-Adamantyl	1-Adamantyl	H	H	4	8.397940	8.50762	8.40958	8.46532	8.35412
1oo	Cycloheptyl	Cycloheptyl	H	H	1.5	8.823900	8.74602	8.72504	8.65948	8.60142
1qq	Cyclobutyl	Cyclobutyl	H	H	400	6.397940	6.80740	6.83370	6.79702	6.77679
1rr	Cyclopropyl	Cyclopropyl	H	H	1000	6.000000	6.18966	6.23126	6.20342	6.19458
1ss	4-Methyl-cylcohexyl	4-Methyl-cylcohexyl	H	H	4	8.397940	8.23749	8.26764	8.25422	8.19765
1vv	Cinnamyl	Cinnamyl	H	H	70	7.154900	7.36547	7.44335	7.41227	7.40610
1xx	Phenyl	Phenyl	CH_3	H	800	6.096910	6.89021	6.16950	6.13048	6.97432 T
1yy	Cyclohexyl	Cyclohexyl	$COOCH_2CH_3$	H	7	8.154900	8.72411 T	8.14937	8.11390	8.65591
1zz	Cyclohexyl	Cyclohexyl	$COCH_3$	H	1.5	8.823900	8.65619 T	8.03398 T	8.00420 T	8.60085
1aaa	Cyclohexyl	Cyclohexyl	H	2-F	2	8.698970	8.89359	8.94174 T	8.93318 T	8.84247 T

T- Represents the test set molecule in the respected model.

Table 2. Unicolumn Statistics of Training and Test Sets.

Model 1 Uni-column Statistics for training set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.5247	9.3979	6.0000	0.8986	255.8386
Model 1 Uni-column Statistics for test set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.9037	8.8239	7.0000	0.5878	55.3256
Model 2 Uni-column Statistics for training set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.5800	9.3979	6.0000	0.8540	257.7205
Model 2 Uni-column Statistics for test set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.6348	8.8239	6.6990	0.9546	53.4437
Model 3 Uni-column Statistics for training set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.5731	9.3979	6.0000	0.8804	272.6326
Model 3 Uni-column Statistics for test set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.7063	8.8239	7.0000	0.7662	38.5316
Model 4 Uni-column Statistics for training set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.6153	9.3979	6.0000	0.8726	251.3048
Model 4 Uni-column Statistics for test set					
Column Name	Average	Max	Min	StdDev	Sum
pIC_{50}	7.4824	8.6990	6.0969	0.8513	59.8593

Table 3.Statistical parameters of developed models

Parameters	**Model-1**	**Model-2**	**Model-3**	**Model-4**
N	34	34	36	33
df	29	29	30	28
r^2	0.6964	0.7107	0.7102	0.7225
q^2	0.6133	0.6230	0.6090	0.6139
F test	16.6336	17.8071	14.7011	18.2295
r^2 se	0.5281	0.4900	0.4999	0.4914
q^2 se	0.5961	0.5594	0.5807	0.5796
pred_r^2	0.7703	0.7256	0.7526	0.6034
pred_r^2se	0.3433	0.5011	0.4922	0.5435
best_rand_r^2	0.36543	0.39054	0.45262	0.47192
best_rand_q^2	0.14050	0.02089	0.08555	0.17578
Z score_rand_r^2	6.40112	6.54517	6.22102	6.63809
Z score_rand_q^2	1.98570	1.81359	2.48490	1.98989
Z score_pred_r^2	2.02296	1.75580	1.58481	1.96859
α_rand_r^2	0.00000	0.00000	0.00000	0.00000
α_rand_q^2	0.05000	0.05000	0.01000	0.05000
α_rand_pred_r^2	0.05000	0.05000	0.10000	0.05000

Fig 2. Graph of Observed vs. Predicted activities for training and test set molecules **model-1**,
A) Training set (Red dots) B) Test set (Blue dots)

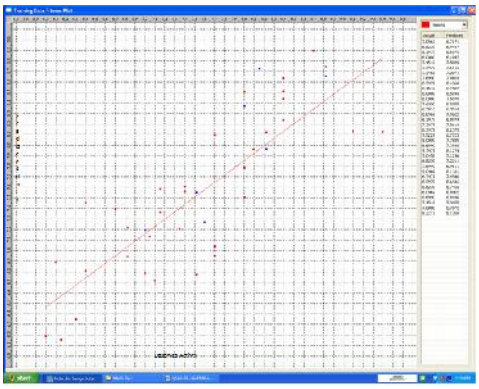
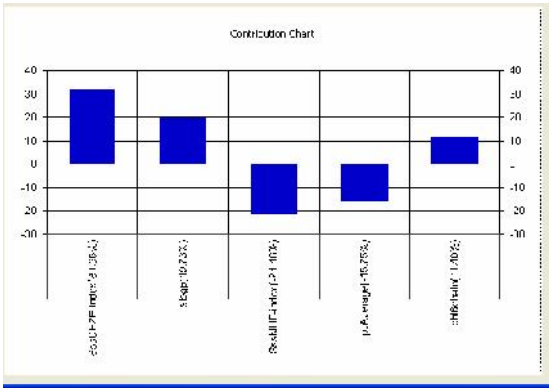


Fig.3. Plot of percentage contribution of each descriptor in developed **model-1**



In conclusion, above QSAR study reveals that the SssNHE- index, IPC Average contributes negatively, which indicate the negative impact of SssNHE- index, IPC Average on anti-allergic activity. This suggest that by decreasing number of NH group connected with two single bonds in a molecule will be helpful for designing of more potent anti-allergic agents.

The present QSAR model reveals that $SssCH_2E$-index descriptor has a major contribution (31.96%) in explaining variation in activities. The interpretation of the model suggests that descriptor $SssCH_2E$ -index (31.96%) is directly proportional to the activity. Hence, increasing the number of CH_2 group connected with two single bonds in a compound will help in designing more potent anti-allergic agents.

The definition for the descriptors that were found to be dominating in the developed QSAR models is given in Table-4.

MODEL – 2 [Random selection method (85 %)]

After further trials performed by random selection method, training sets of 85% were used to develop another statistically significant QSAR equation by Partial Least Squares with stepwise forward-backward variable selection method, the final QSAR equation developed as follows:

pIC_{50} = 0.167201 $SssCH_2E$ -index + 0.445863 slogp – 1.41567 SssNHE- index – 1.03507 e $^{-008}$ IPC Average + 2.79095 Chi6chain – 0.408382 SaasN (N oxide) E index + 11.7383

Values of different statistical parameters generated in model during 2D QSAR studies is shown in Table – 3.

The observed and predicted pIC_{50} values are shown in Table 1.

The plot of observed vs. predicted activity is shown in Fig. (4).From the plot it can be seen that the model is quite able to predict the activity of training set (maximum points are close to regression line) as well as external.

The descriptors which contribute for the pharmacological action are shown in **Fig. (5)**.The major group of descriptors involved sub groups like $SssCH_2E$ -index, slogp, SssNHE- index, IPC Average and Chi6chain and SaasN(Noxide)E-index, help in understanding the effect of substituent at different position of substituted 2-phenyl-benzimidazole.

In conclusion, above QSAR study reveals that SssNHE- index, IPC Average and SaasN(Noxide)E-index contributes negatively, which indicate the negative impact of SssNHE-index , IPC Average and SaasN(Noxide)E-index on anti-allergic potential. As SssNHE- index is having maximum negative contribution, hence, decreasing number of NH group connected with two single bonds in a molecule will be helpful for designing of more potent anti-allergic agents.

The above QSAR model reveals the major contribution of $SssCH_2E$ -index in explaining variation in activities. The interpretation of the model suggests that descriptor $SssCH_2E$-index (27.60%) is directly proportional to the activity. Hence, by increasing the number of CH_2 group connected with two single bonds in a compound will help in designing more potent anti-allergic agents.

The definition for the descriptors that were found to be dominating in the developed QSAR models is given in Table-4.

MODEL – 3 [Random selection method (90 %)]

Another statistically significant QSAR equation was obtained by Partial Least Squares with stepwise forward-backward variable selection using random selection method as selection criteria of training and test sets. Training sets of 90% were used to develop the final QSAR equation as follows:

pIC_{50} = 0.167723 $SssCH_2E$ + 0.429605 slogp – 1.38535 SssNHE- index – 1.05549 e $^{-008}$ IPC Average + 3.09541 Chi6chain – 0.428059 SaasN (N oxide) E index + 11.5476

Values of different statistical parameters generated in model during 2D QSAR studies is shown in Table – 3

The observed and predicted pIC_{50} values are shown in Table 1.

The plot of observed vs. predicted activity is shown in Fig. (6).From the plot it can be seen that the model is quite able to predict the activity of training set (maximum points are close to regression line) as well as external.

The descriptors which contribute for the pharmacological action are shown in Fig.(7).The major group of descriptors involved sub groups like $SssCH_2E$-index, slogp, SssNHE-index, IPC Average,Chi6chain and SaasN(Noxide)E-index, help in understanding the effect of substituent at different position of substituted 2-phenyl-benzimidazole.

In conclusion, above QSAR study reveals that SssNHE-index, IPC Average and SaasN(Noxide)E-index contributes negatively, which indicate the negative impact of SssNHE-index , IPC Average and SaasN(Noxide)E-index on anti-allergic potential. The descriptor SssNHE-index is having (-19.52) contribution. This suggest that, by decreasing number of NH group connected with two single bonds in a molecule will be helpful for designing of more potent anti-allergic agents.

The above QSAR model reveals that $SssCH_2E$ -index descriptor has a major contribution in explaining variation in activities. The interpretation of the model suggests that descriptor $SssCH_2E$-index (27.90%) is directly proportional to the activity. Hence, by increasing the number of CH_2 group connected with two single bonds in a compound will help in designing more potent anti-allergic agents.

The definition for the descriptors that were found to be dominating in the developed QSAR models is given in Table-4.

MODEL – 4 [Manual selection method]

In order to develop the significant model by manual selection method the whole range of biological activity were sorted in ascending order and every 3rd - 8th compound was assigned in test set. Significant QSAR equation depicted below was developed by Partial Least Squares with stepwise forward-backward variable selection method.

pIC_{50} = 0.159939 $SssCH_2E$ -index + 0.433064 slogp – 1.34414 SssNHE- index – 1.02748 e $^{-008}$ IPC Average + 2.88281 Chi6chain + 11.3453.

Values of different statistical parameters generated in model during 2D QSAR studies is shown in Table – 3.

The observed and predicted pIC_{50} values are shown in Table 1.

The plot of observed vs. predicted activity is shown in Fig.(8).From the plot it can be seen that the model is quite able to predict the activity of training set (maximum points are close to regression line) as well as external.

The descriptors which contribute for the pharmacological action are shown in Fig.(9).The major group of descriptors involved sub groups like $SssCH_2E$-index, slogp, SssNHE-index , IPC Average and Chi6chain , help in understanding the effect of substituent at different position of substituted 2-phenyl- benzimidazole.

In conclusion, above QSAR study reveals that SssNHE-index and IPC Average contributes negatively, which indicate the negative impact of SssNHE- index and IPC Average on anti-allergic potential. This suggest that by decreasing number of NH group connected with two single bonds in a molecule will be helpful for designing of more potent anti-allergic agents.

The above QSAR model reveals that $SssCH_2E$-index descriptor has a major contribution in explaining variation in activities. The interpretation of the model suggests that descriptor $SssCH_2E$-index (29.76%) is directly proportional to the activity. Hence, by increasing the number of CH_2 group connected with two single bonds in a compound will help in designing more potent anti-allergic agents.

The definition for the descriptors that were found to be dominating in the developed QSAR models is given in Table-4.

3.2 Validation of QSAR models

All the above QSAR models have shown good correlation between their corresponding descriptors and biological activity. Also large values of F indicate that the model fit in all cases was not a chance occurrence and all models were statistically significant [30]. However, a QSAR model is considered to be predictive, if the following conditions are satisfied: pred_$r^2 \geq 0.5$ and $q^2 \geq 0.6$ [31-34].

Table 4 Definition for the descriptors that were found to be dominating in the QSAR models

Descriptor	Definition
$SssCH_2E$ -index	The descriptor signifies the electrotopological state indices for number of CH_2 group connected with two single bonds.
slogp	The descriptor signifies the Octanol/water partition coefficient.
SssNHE- index	The descriptor signifies the electrotopological state indices for number of NH group connected with two single bonds.
IPC Average	This is a type of information theory based descriptor.
Chi6chain	The descriptor signifies a retention index of six membered ring**.**
SaasN(Noxide)E-index*	Electrotopological state indices for number of nitro-oxide group connected with two aromatic and one single bond.

*contributing in model 2 and 3 only, rest all are contributing in all four models

Fig 4. Graph of Observed vs. Predicted activities for training and test set molecules **model-2**,
A)Training set (Red dots) B) Test set (Blue dots)

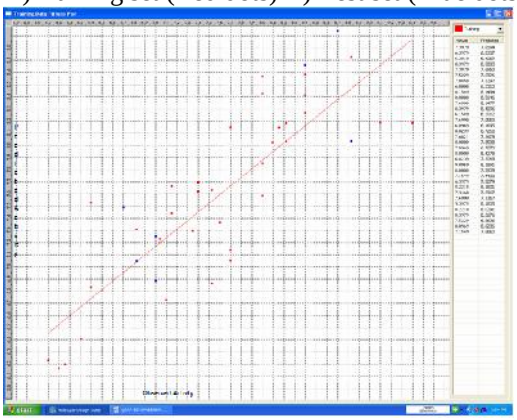


Fig.5. Plot of percentage contribution of each descriptor in developed **model-2**

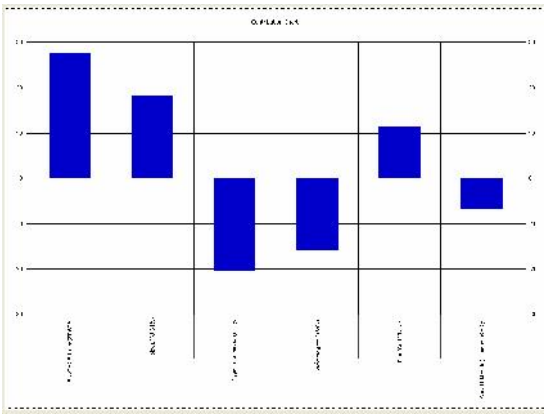


Fig 6. Graph of Observed vs. Predicted activities for training and test set molecules **model- 3**,
A) Training set (Red dots) B) Test set (Blue dots)

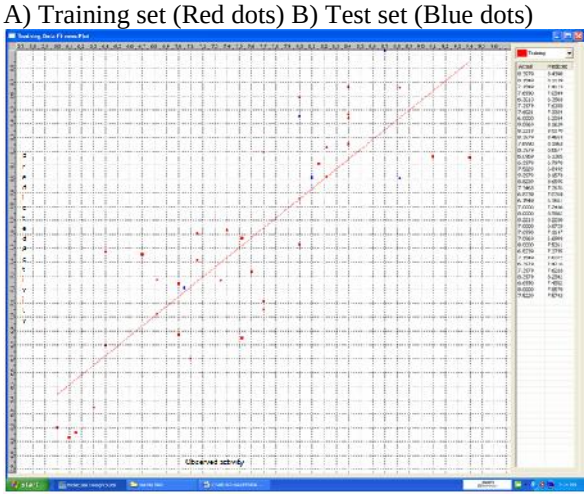


Fig.7. Plot of percentage contribution of each descriptor in developed **model-3**

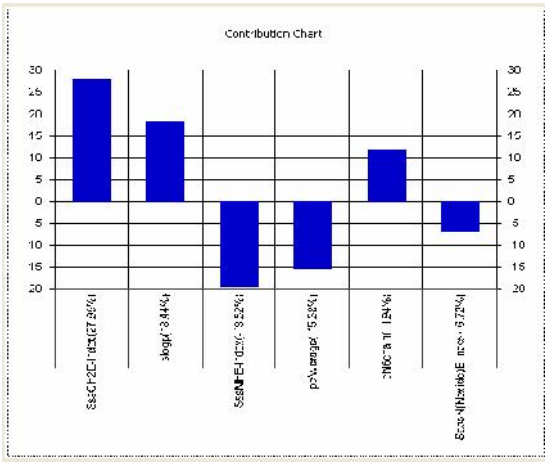


Fig 8. Graph of Observed vs. Predicted activities for training and test set molecules **model -4**,
A) Training set (Red dots) B) Test set (Blue dots)

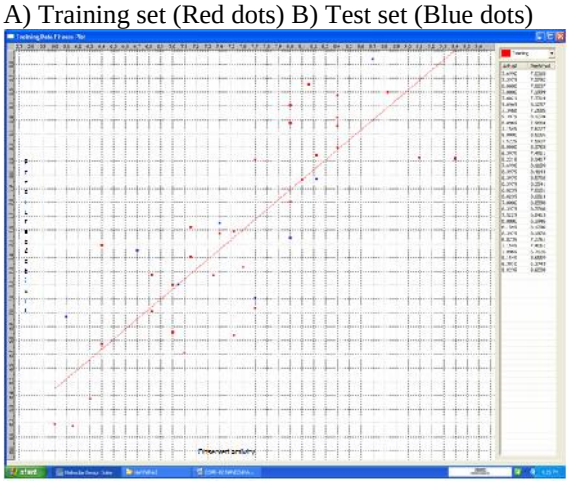
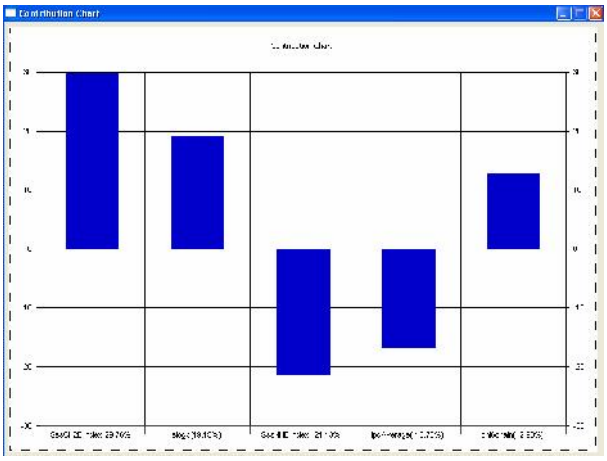


Fig.9. Plot of percentage contribution of each descriptor in developed **model-4**



4. CONCLUSION

Conventional quantitative structure activity relationship studies was performed on a series of substituted-2-phenyl-benzimidazole derivative by means of the PLS with stepwise forward-backward variable selection method that may be statistically treated to reveal the molecular characteristics which are essential for potency. The models so obtained were analyzed and validated for their statistical significance and external predictivity. The developed 2D- QSAR models revealed the importance of different physico- chemical properties of compounds contributing to the biological activity. In all four optimized models, the SWFB-PLS method is giving noteworthy results.

A close observation of contribution chart of the descriptors suggests positive contribution of $SssCH_2E$-index and was found to be predominating in all the above significant models. Hence, increasing the value of this descriptor will provide more potent anti allergic agent. Similarly in all the equations SssNHE- index was found to be negative, hence decreasing the value of this descriptor will provide potent compound. Thus, the awareness and understanding of the descriptors involved in the anti-allergic activity of these compounds could provide a great opportunity for the ligand structures design with appropriate features.

5. ACKNOWLEDGEMENT

The authors would like to thanks SLT Institute of Pharmaceutical Sciences,Guru Ghasidas University (Bilaspur) for providing the necessary facilities and Vlife Science Pune for providing the software.

6. REFERENCES

1. Goldsby R.A., Kindt T.J., Osborne B.A., Kuby J., Immunology, W.H. Freeman and Company, New York,2002, 5th Ed.,361–370.
2. Raj N.R., Jain S.K., Raj C.N., Panda A.B., Various screening methods for anti-allergic activity: An overview, Int. J. Pharm. Sci. Nanotech., 2010, 3,906-911.
3. Galli S.J, Kalesnikoff J., Grimbaldeston M.A., Piliponsky A.M., Williams C.M., Tsai M., Mast cells as “tunable” effector and immunoregulatory cells: recent advances, Annu. Rev. Immunol., 2005, 23, 749-86.
4. Persson C.G.A., Role of plasma exudation in asthmatic airways, Lancet 1986; ii: 1126-1128.
5. Hiroyuki Nakano, Tsutomu Inoue, Nobuhide Kawasaki, Hideki Miyataka, Hitoshi Matsumoto, Takeo Taguchi, *et al*., Synthesis of Benzimidazole Derivatives as Antiallergic Agents with 5- Lipoxygenase Inhibiting Action, Chem. Pharm. Bull., 1999, 47, 1573-1578.
6. Metzger H., Alcaraz G., Hohman R., Kinet J.P., Pribluda V., Quarto R.,The receptor with high affinity for immunoglobulin E., Annu. Rev. Immunol., 1986, 4, 419-470.
7. Leonard B., Bacharier M.D., Raif S., Geha M.D., Molecular mechanisms of IgE regulation, J. Allergy Clin. Immunol, 2000, 105, S547-558.
8. Tewtrakul Supinya, Subhadhirasakul Sanan,Anti-allergic activity of some selected plants in the Zingiberaceae family, J. Ethnopharmacol., 2007, 109, 535–538.
9. Gould H., Beavil R., Relijic R., Shi J., Ma C., Sutton B., *et al.*, IgE regulation: molecular mechanisms, New York: John Wiley & Sons Ltd ,1997, 37-59.

10. Tsicopoulos A., Joseph M., The role of CD23 in allergic disease, Clin. Exp.Allergy, 2000,30 ,602- 605.
11. Sarfati M., Delespesse G.,Possible role of human lymphocyte receptor for IgE (CD23) or its soluble fragments in the in vitro synthesis of human, J. Immunol., 1988,141,2195-2199.
12. Lawrence D.A., Weigle W.O., Spiegelberg H.L., Immunoglobulins cytophilic for human lymphocytes, monocytes, and neutrophils. , J. Clin. Invest, 1975, 55,368-387.
13. Backer V., Ulric C. S., Wendelboe D., Bach-Mortensen N., Hansen K.K, Laursen E. M., Dirksen A., Distributions of Serum IgE in Children and Adolescents Aged 7 to 16 years in Copenhagen, in Relation to Factors of Importance, Allergy, 1992, 47,484-489.
14. MacDonald S.M, Rafnar T., Langdon J., Lichtenstein L.M., Molecular Identification of an IgE- Dependent Histamine-Releasing Factor, Science, 1995, 269,688-690.
15. Raj Navin, Jain Sanmati K., 3D QSAR studies in conjunction with k-nearest neighbor molecular field analysis (k-NN-MFA)on a series of substituted 2-phenyl-benzimidazole derivatives as an anti allergic agents, Dig. J. Nanomater. Bios, 2011, 6, 1811-1821.
16. Habib N.S., Soliman R., Ashour F.A., Taiebi M., Synthesis and antimicrobial testing of novel oxadiazolyl benzimidazole derivatives., Pharmazie,1997,52,746-749.
17. Tuncbilek M., Goker H., Ertan R., Eryigit R., Kendi E., Altanlar E., Synthesis and antimicrobial activity of some new anilino benzimidazoles.,Arch. Pharm., 1997, 330, 372–376.
18. Pedini M., Alunni Bistochi G., Ricci A., Bastianini L., Lepri E., New heterocyclic derivatives of benzimidazole with germicidal activity – XII. Synthesis of N1-glycosyl-2-furyl benzimidazoles. Farmaco, 1994, 49, 823-827.
19. Lackner T.E., Clissold S.P., Bifonazole. A review of its antimicrobial activity and therapeutic use in superficial mycoses, Drugs, 1989, 38,204-225.
20. Richards Mark L., Lio Shirley Cruz, Sinha Anjana, Banie Homayon, Thomas Richard J., Major Michael, *et al*., Substituted 2-phenyl-benzimidazole derivatives:novel compounds that suppress key markers of allergy, Eur. J.Med. Chem.2006,41, 950–969.
21. VLifeMDS 3.5; Molecular Design Suite, Vlife Sciences Technologies Pvt. Ltd., Pune, India (2004), www.vlifesciences.com.
22. Williams D.A., Lemke T.L. Foye's Principles of Medicinal Chemistry, Lippincott Williams & Wilkins, Baltimore, 2002, 5th Ed, 81.
23. Halgren T.A.,Merck molecular force field-I. Basis, form, scope, parameterization, and performance of MMFF94, J. Comput. Chem, 1996, 17,490-519.
24. Halgren T.A., Merck molecular force field-II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, J. Comput. Chem, 1996, 17, 520-552.
25. Halgren T.A Merck molecular force field-III. Molecular geometries and vibrational frequencies for MMFF94, J. Comput. Chem, 1996,17, 553-586.
26. Halgren T.A Merck molecular force field-IV.conformational energies and geometries for MMFF94, J.Comput.Chem, 1996,17, 587-615.
27. Chakraborti A.K., Gopalakrishnan B., Sobhia M.E., Malde A 3D-QSAR Studies of Indole derivatives as Phosphodiesterase IV Inhibitors, Eur. J. Med.Chem., 2003, 38, 975-982.
28. Jain Sanmati K., Mallick S., Dubey R., Nag S., Yadav A., 2D-QSAR analysis on 4-Flouro-2-Cyanopyrrolidine derivatives as DPP-IV Inhibitors, J. Comput. Method Mol. Design, 2011, 1, 14- 25.
29. Ajmani S., Jadhav K., Kulkarni S.A., Three-dimensional QSAR using the k-nearest neighbor method and its interpretation, J. Chem. Inf. Model., 2006, 46, 24-31.
30. Puratchikody A., Nagalakshmi G.,Doble M., Experimental and QSAR studies on antimicrobial activity of benzimidazole derivatives.,Chem. Pharm. Bull. 2008, 56, 273-281.
31. Tropsha A., Gramatica P., Gombar V.K., The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models, QSAR Comb. Sci., 2003, 22, 69-77.
32. Golbraikh A., Tropsha A., Beware of q2!, J. Mol. Graph. Model, 2002, 20, 269-276.
33. Afantitis A., Melagraki G., Sarimveis H., Igglessi-Markopoulou O., Kollias G., A novel QSAR model for predicting the inhibition of CXCR3 receptor by 4-N-aryl-[1,4] diazepane ureas., Eur. J. Med. Chem., 2009, 44:877–884.
34. Ravichandran V., Mourya V.K., Agrawal R.K., QSAR prediction of HIV–1 reverse transcriptase inhibitory activity of benzoxazinone derivatives, Internet Electron. J. Mol. Des., 2007,6,363-374.

*****