International Journal of PharmTech Research
CODEN (USA): IJPRIF ISSN : 0974-4304
Vol.4, No.2, pp 676-688, April-June 2012

Two Dimensional QSAR Study of Some Novel 2-Azetidinone Series for Their Antibacterial Activity against Bacillus subtilis.

Mehta Parul D*, A. K. Pathak

Department of Pharmacy, Barkatullah University, Bhopal, 462026, M.P.,India.

Corres.author: parulmehta1@rediffmail.com, parulsengar@gmail.com
Tel.:+91 09301145642, +91 0755 2491846; fax: +91 0755 2491857

Abstract : Two dimensional quantitative structure activity relationship (QSAR) study on series of substituted 2-azetidinone derivatives was performed by using V-LIFE MDS 3.0 software. Several statistical expression for 2D QSAR were developed using statistical methods like multiple regeression, principle component regression, partial least square regression etc. Out of several models, the best five 2D QSAR models having highest correlation coefficient and cross validated squared correlation coefficient were selected for further study (r^2>0.7, q^2>0.7). QSAR study revealed that Atomic valence connectivity index, element count, electro topological, estate contribution and alignment-independent descriptors are primarily responsible for biological activity. This approach showed that physicochemical descriptor **Hydrogencounts, SddsN(nitro)E-index** and alignment-independent descriptors **T_C_Cl_3** were found to show significant correlation with biological activity in 2-azetidinone derivatives. The information rendered by 2D QSAR models may lead to a better understanding of structural requirements of antibacterial activity and can help in the design of novel potent molecules.
Key words: 2D QSAR, Antibacterial, 2-Azetidinone.

INTRODUCTION:

The synthesis of heterocyclic compound has always drawn the attention of chemist over the years mainly because of their important biological properties. Particularly, the role of β-lactam which are endowed with unique structure and potent antibacterial activity. The 2-azetidinone (β-lactam) ring system is the common structural feature of a number of broad spectrum β-lactam antibiotics, including penicillins, cephalosporins, carbapenems, nocardicins, monobactams, clavulanic acid, sulbactams and tazobactams, which have been widely used as chemotherapeutic agents to treat bacterial infections and microbial diseases [1-9].

Most of the researches up to early 90s focused on synthesis of 2-azetidinones and study of their antibacterial property. In recent years, renewed interest has been focused on the synthesis and modification of β-lactam ring to obtain compounds with diverse pharmacological activities like cholesterol absorption inhibitory activity [10], human tryptase[11], thrombin[12] and chymase inhibitory activity[13], vasopressin V1a antagonist activity[14], antidiabetic[15] anti-inflammatory[16] antiparkinsonian[17] and anti-HIV activity[18].

They are also found to be a potent inhibitor of serine protease, human leukocyte elastase and human cytomegalovirus protease enzyme[19-22], and are effective on central nervous system; in recent past these derivatives are also found to be moderately active against several types of cancer[23]. Recently we publish a detail review article on "2-Azetidinone a new profile of various pharmacological activity" [24]. This review mainly focuses on the diverse pharmacological properties associated with monocyclic 2- azetidinone moiety.

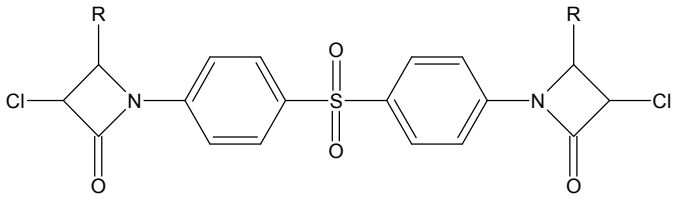
In the present study, 2D-QSAR analysis of some novel synthesized 2-Azetidinone compounds with antimicrobial activity was performed by using VLife MDS software. Regression method like multiple regression, principle component regression, partial least square regression etc were used to build a QSAR model in the form of mathematical equation. This equation explains variation of one or more dependant variables (usually activity) in terms of independent variables (descriptors) multiple regression was used as the standard method for multivariate data analysis. It is also called as ordinary least squares regression (OLS) and this method estimate the value of the regression coefficient by applying least squares curve fitting method[25-27]. stepwise multiple regression (SMR) was used as an approach to select a subset of variables, when the number of independent variable (descriptors) were much more than the number of data points (molecules). Each variable was added to the equation and a new regression was performed.

MATERIAL AND METHODS

Experimental

2D –QSAR study was done on a series of some newely synthesized 4,4'-bis [3-chloro-4-aryl-azetidin-2-one-1-yl] diphenyl sulphone derivatives[28]. A data set of 20 molecules with various substitutions was taken for the study; antibacterial activity against *B.subtilis* was expressed in term of minimum inhibitory concentration (MIC). The biological data was converted to negative log (-log) in moles to reduce the skewness of the data set. (Table-1). Structures and antimicrobial activity data of the compounds are given in Figure 1 and Table 1.

Figure 1. Structure of 4,4'-bis [3-chloro-4-aryl-azetidin-2-one-1-yl] diphenyl sulphone



R = Heterocyclic or substituted aryl ring

Table 1: biological activity data and structures of the compound in series

Compd.	R	MIC (μg/ml)	-log MIC
AZ01	C_6H_5-	49	4.30980
AZ02	P-CH_3-C_6H_4-	47	4.32790
AZ03	O-Cl-C_6H_4-	04	5.39794
AZ04	M-Cl-C_6H_4-	09	5.04575
AZ05	P-Cl-C_6H_4-	6.5	5.187086
AZ06	O-OH-C_6H_4-	20	4.69897
AZ07	M-OH-C_6H_4-	27	4.56863
AZ08	P-OH-C_6H_4-	21	4.67778
AZ09	O-NO_2-C_6H_4-	2.5	5.60205
AZ10	M-NO_2-C_6H_4-	1.5	5.82390
AZ11	P-NO_2-C_6H_4-	1.0	6.0000
AZ12	M-Br-C_6H_4-	13	4.88605
AZ13	P-OCH_3-C_6H_4-	31	4.50863
AZ14	3,4-Di-OCH_3-C_6H_3-	36	4.443697
AZ15	3,4,5-Tri-OCH_3-C_6H_2-	41	4.387216
AZ16	3-OCH_3,4-OH-C_6H_3-	32	4.494850
AZ17	P-N$(CH_3)_2$-C_6H_4-	42	4.37675
AZ18	3,5-Di-Br,2-OH-C_6H_3-	12	4.920818
AZ19	Furan	16	4.79588
AZ20	Thiophene	14	4.85387

Molecular modeling

The molecular structures of the compounds in series were sketched by using V-Life MDS module 2DD draw of VLife science molecular modeling software. The sketches molecules were then transferred to three-dimensional structures (3D) the geometries of generated 3D structures were optimized using MMFF force field as implemented in the V-Life MDS. The gradient norm 0.001 kcal/molA° was used to calculate electronic and geometric parameters of the molecules. Complete geometry optimization was performed taking the most extended conformations as starting geometries. The basis of energy minimization is that the drug binds to effectors/receptor in the most stable form i.e. minimum energy state form.

Descriptor generation:

The relationship between biological activities and various descriptors (Physiochemical and alignment-independent) were established by sequential multiple regression analysis (MLR) using MDS 3.0, in order to obtain QSAR models. The MDS 3.0 program was employed for the calculation of different quantum chemical descriptors including heat of formation, dipole moment, local charges, different topological descriptors[29-30]. elemental count and constitutional descriptors for each molecule. Chemical parameters including molar volume (V), molecular surface area (SA), hydrophobicity (log P), hydrogenacceptor count (HAC), hydration energy (HE) and molecular polarizability (MP) were also calculated by using software.

Model development:

The calculated descriptors were gathered in a data matrix. First, the descriptors were checked for constant or near constant values and those detected were discarded from the original data matrix. Then, the descriptors were correlated with each other and with the activity data. Variable selection for the QSAR modeling was carried out by stepwise partial least square method (PLS) using statistical program of V-Life MDS. The program employs a stepwise technique i.e. only one parameter at a time was added to a model and always in the order of most significant to least significant in terms of F-test values. Finally, different regression analysis with stepwise selection and elimination of variables was applied to the development of QSAR models using software.

The resulting models were validated by leaveone- out cross-validation procedures to check their predictivity and robustness. The antibacterial activity data and various parameters (Physiochemical and alignment independent) were taken as dependent and independent variables respectively and correlation were established between them by employing multiple sequential regressions (MLR), partial least square (PLS) and multiple component regression (PCR) method, using random selection. Statistical parameters were calculated subsequently for each steps in the process, so the significance of the added parameter could be verified. The goodness of the correlation was tested by the regression correlation coefficient (r^2), the F-test and the standard error of estimate (SEE). The correlation coefficient value closer to 1.0 represent the better fit of the model. High values of the F-test indicate that the model is statistically significant.

Finally the derived QSAR models were used for the prediction of the activity value of the compound in the test set and the external validation parameter, predictive r^2 (r^2-pred) was calculated for evaluating the predictive capacity of the QSAR model. A value of r^2 predictive greater than 0.5 indicates good predictive capacity of the QSAR model. Then different statistical model were developed by using this method. The models showed the better correlation between biological activity and physicochemical descriptor values. The values of the descriptor were used to create a 2D QSAR model for assuming the biological activity with the help of standard methods[31-34].

Table 2: Descriptors selected for the 2-azetidinone series for their antibacterial activity against *B.subtilis*

Model	Parameter-1	Parameter-2	Parameter-3	Parameter-4	Parameter-5
Model-1	SddsN(nitro)E-index (-26.88%)	T_C_Cl_3 (+21.55%)	Hydrogens Count (-21.12%)	SaaCHcount (-10.88%)	T_2_N_7 (+19.57%)
Model-2	SddsN(nitro)E-index (-36.35%)	T_C_Cl_3 (+23.12%)	Hydrogens Count (-18.14%)	chi4pathCluster (+13.72%)	T_2_T_6 (-8.65%)
Model-3	SddsN(nitro)E-index (-35.34%)	T_T_Cl_3 (+21.98%)	T_C_C_2 (-21.70%)	chi3Cluster (+14.90%)	T_N_Cl_5 (-6.08%)
Model-4	SddsN(nitro)E-index (-27.52%)	T_T_Cl_3 (+22.00%)	Hydrogens Count (-19.99%)	SaaCHcount (-11.15%)	T_2_N_7 (+19.34%)
Model-5	SddsN(nitro)E-index (-36.33%)	T_T_Cl_3 (+22.46%)	T_C_C_2 (-21.02%)	chi3Cluster (+14.31%)	T_N_Cl_5 (-5.87%)

Table 3: Statistical and validation parameters of five different models for their antibacterial activity against *B.subtilis*

Parameters	Model-1	Model-2	Model-3	Model-4	Model-5
Optimum component (n)	14	15	15	15	15
Degree of freedom	11	11	11	11	11
r^2	0.9676	0.9591	0.9792	0.9788	0.9888
q^2	0.8993	0.7501	0.9584	0.9342	0.9710
F- test	109.407	85.9325	172.986	169.068	322.316
r^2_se	0.1098	0.1143	0.0887	0.0877	0.0658
q^2_se	0.1936	0.2825	0.1256	0.1544	0.1056
pred_r^2	0.9641	0.9858	0.9239	0.9215	0.8710
pred_r^2 se	0.0938	0.0687	0.1260	0.1453	0.1519
ZScore R^2	3.97588	3.71816	3.86623	4.08098	4.45233
ZScore Q^2	2.48912	2.43661	2.56127	2.74686	2.71746
Best Rand R^2	0.76410	0.79964	0.89473	0.79019	0.83421
Best Rand Q^2	0.51297	0.39639	0.60491	0.54074	0.66293
Alpha Rand R^2	0.00100	0.00100	0.00100	0.00061	0.00013
Alpha Rand Q^2	0.01000	0.01000	0.01000	0.01000	0.01000
Z Score Pred R^2	1.58042	2.09369	1.26722	1.44144	1.09960
Best Rand Pred R^2	0.93509	0.91674	0.90741	0.95389	0.83825
Alpha Rand Pred R^2	0.10000	0.05000	0.00000	0.10000	0.00000

Table 4: Actual, predicted and residual value of different models of 2-azetidinone series for their antibacterial activity against *B.subtilis* by 2D QSAR method

Str. code	Actual activity -log MIC	Model-1		Model-2		Model-3		Model-4		Model-5	
		Predicted activity	Residual	Predicted activity	Residual	Predicted activity	Residual	Predicted activity	Residual	Predicted activity	Residual
AZ01	4.3098	4.431994	-0.12219	4.478209	-0.16841	4.398307	-0.08851	4.416683	-0.10688	4.404316	-0.09452
AZ02	4.3279	4.390948	-0.06305	4.449847	-0.12195	4.376225	-0.04833	4.38545	-0.05755	4.380837	-0.05294
AZ03	5.39794	5.34739	0.05055	5.414489	-0.01655	5.394931	0.003009	5.348279	0.049661	5.39426	0.00368
AZ04	5.04575	5.119058	-0.07331	5.131521	-0.08577	5.006605	0.039145	5.113987	-0.06824	5.008873	0.036877
AZ05	5.187086	5.119058	0.068028	5.187225	-0.00014	5.212807	-0.02572	5.113987	0.073099	5.205064	-0.01798
AZ06	4.69897	4.571912	0.127058	4.588839	0.110131	4.600933	0.098037	4.558752	0.140218	4.594388	0.104582
AZ07	4.56863	4.571912	-0.00328	4.554382	0.014248	4.683475	-0.11485	4.558752	0.009878	4.671816	-0.10319
AZ08	4.67778	4.571912	0.105868	4.610086	0.067694	4.683475	-0.00569	4.558752	0.119028	4.671816	0.005964
AZ09	5.60205	5.640983	-0.03893	5.661291	-0.05924	5.627339	-0.02529	5.641368	-0.03932	5.630168	-0.02812
AZ10	5.8239	5.811214	0.012686	5.828907	-0.00501	5.91834	-0.09444	5.811087	0.012813	5.907747	-0.08385
AZ11	6	5.964624	0.035376	5.933616	0.066384	5.917594	0.082406	5.963499	0.036501	5.906991	0.093009
AZ12	4.88605	4.662394	0.223656	4.634501	0.251549	4.683475	0.202575	4.645403	0.240647	4.671816	0.214234
AZ13	4.50863	4.390948	0.117682	4.47716	0.03147	4.446326	0.062304	4.38545	0.12318	4.447978	0.060652
AZ14	4.443697	4.349901	0.093796	4.443235	0.000462	4.42034	0.023357	4.354216	0.089481	4.422219	0.021478
AZ15	4.387216	4.308854	0.078362	4.386728	0.000488	4.401145	-0.01393	4.322983	0.064233	4.402831	-0.01562
AZ16	4.49485	4.530865	-0.03602	4.612711	-0.11786	4.642163	-0.14731	4.527519	-0.03267	4.63168	-0.13683
AZ17	4.37675	4.427321	-0.05057	4.413706	-0.03696	4.431358	-0.05461	4.431349	-0.0546	4.431171	-0.05442
AZ18	4.920818	5.03271	-0.11189	4.970508	-0.04969	5.076214	-0.1554	5.016193	-0.09538	5.040222	-0.1194
AZ19	4.79588	4.892793	-0.09691	4.784922	0.010958	4.859183	-0.0633	4.874124	-0.07824	4.840785	-0.04491
AZ20	4.85387	4.892793	-0.03892	4.784922	0.068948	4.859183	-0.00531	4.874124	-0.02025	4.840785	0.013085

Tabel 5: correlation matrix of different parameter used in the development of different models for their antibacterial activity against *B.subtilis*

Descriptors	chi3Cluster	chi4pC	HC	SaaCH count	SddsNE -i	T_2_T_6	T_2_N_7	T_T_Cl_3	T_C_C_2	T_C_Cl_3	T_N_Cl_5
chi3Cluster	**1**	0.880148	0.257845	-0.34469	-0.40652	0.692017	0.482511	-0.12989	0.552785	-0.12989	0.103551
chi4pC	0.880148	**1**	0.528579	-0.56111	-0.2362	0.575988	0.324296	-0.14898	0.657666	-0.14898	-0.01541
HC	0.257845	0.528579	**1**	-0.21653	0.256902	0.008702	-0.06398	-0.25097	0.861005	-0.25097	-0.20386
SaaCHcount	-0.34469	-0.56111	-0.21653	**1**	-0.20242	0.054855	0.228326	0.197746	-0.04168	0.197746	0.160633
SddsNE-i	-0.40652	-0.2362	0.256902	-0.20242	**1**	-0.63741	-0.88936	0.172368	0.084768	0.172368	-0.31937
T_2_T_6	0.692017	0.575988	0.008702	0.054855	-0.63741	**1**	0.454759	-0.01168	0.361799	-0.01168	0.464811
T_2_N_7	0.482511	0.324296	-0.06398	0.228326	-0.88936	0.454759	**1**	-0.19442	0.07823	-0.19442	0.129219
T_T_Cl_3	-0.12989	-0.14898	-0.25097	0.197746	0.172368	-0.01168	-0.19442	**1**	-0.08281	1	0.254024
T_C_C_2	0.552785	0.657666	0.861005	-0.04168	0.084768	0.361799	0.07823	-0.08281	**1**	-0.08281	-0.06727
T_C_Cl_3	-0.12989	-0.14898	-0.25097	0.197746	0.172368	-0.01168	-0.19442	1	-0.08281	**1**	0.254024
T_N_Cl_5	0.103551	-0.01541	-0.20386	0.160633	-0.31937	0.464811	0.129219	0.254024	-0.06727	0.254024	**1**

chi4pC:chi4pathCluster, HC:HydrogensCount, SddsNE-i: SddsN(nitro)E index

RESULT AND DISCUSSION

Among the generated QSAR models, five best model were selected on the basis of various statistical parameters such as squared correlation co-efficient (r^2) which is relative measure of quality of fit, fischer's value (F test)which represents F-ratio between the variance of calculated and observed activity, standard error ((r^2-Se) representing absolute measure of quality of fit, correlation co-efficient (q^2), standard error of cross-validated squared correlation co-efficient (q^2-Se), predicted squared regression (pred- r^2) to estimate the predictive potential of the models respectively.

Out of hundreds of equations generated, five best QSAR equations are given below:

Model:1:- Optimum Components = 3; n = 15; Degree of freedom = 11; r^2 = 0.9676; q2 = 0.8993; F test = 109.4074; r^2 se = 0.1098; q^2 se = 0.1936; pred_r^2 = 0.9641; pred_r^2se = 0.0938

log1_MIC = - 0.7200 SddsN(nitro)E-index + 0.1142 T_C_Cl_3 - 0.0452 HydrogensCount - 0.0700 SaaCHcount + 0.0770 T_2_N_7 + 6.1504

The model-1 fulfill the selection criteria's like correlation coefficient r>0.8 (r^2 = 0.9676) with low standard error of squared correlation co-efficient ((r^2-Se) showed the relative good fitness [**Graph-1]** of the model and F value 109.4074. the validation criteria for selection of the model are cross validated squared correlation co-efficient (q^2>0.7) for training set and (pred _ r^2>0.4) for test set. This model was developed by taking five compounds in test set and rest of the 15 compounds in training set. This model showed low standard error of cross validated squared correlation co-efficient (q^2-Se) and standard error of pred _ r^2Se, which show accuracy of the statistical calculation. The model correlates descriptors SddsN(nitro)E-index, T_C_Cl_3 HydrogensCount, T_2_N_7 and SaaCHcount with the biological activity. In this model Element count and electrotopological parameter showed negative contribution while alignment independent descriptor shows positive contribution. The element count descriptors, Hydrogen count(-21.12%) and SaaCHcount (-10.88%) showed negative contribution, electrotopological parameter, SddsN(nitro)E-index(-26.88%) showed negative contribution while the alignment independent descriptors T_C_Cl_3 (+21.55%) and T_2_N_7 (+19.57%) showed positive contribution [**chart-1].**

Graph -1: Fitness Plot between the experimental and Predicted Activities for model-1 [Training Set (red spots) and test set (blue spots)]

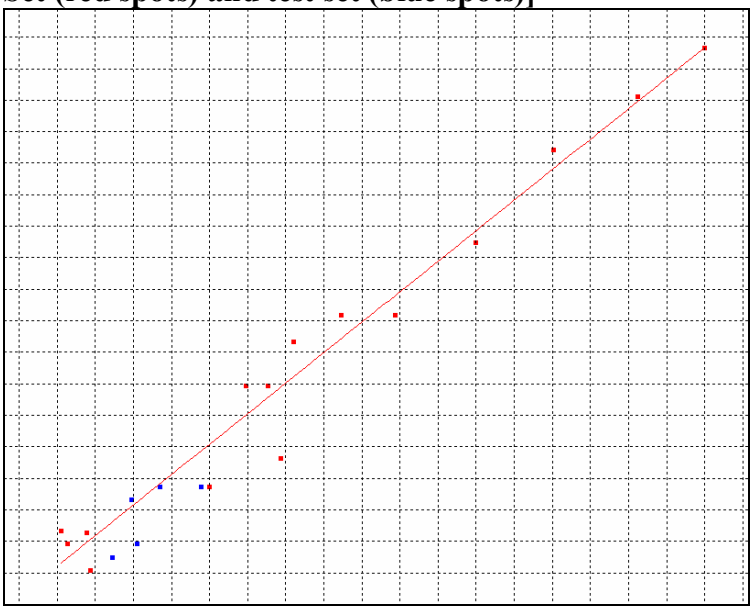
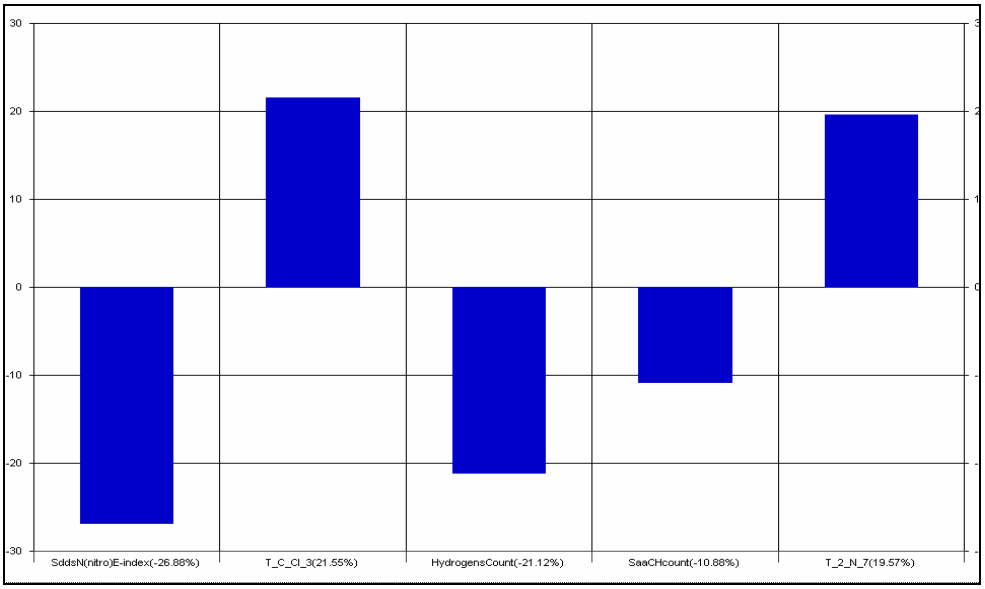


Chart: 1: Contribution chart of individual descriptors that are important for the biological activity in model -1.



Model-2:- Optimum Components = 3; n = 15; Degree of freedom = 11; r^2 = 0.9591; q^2 = 0.7501; F test = 85.9325; r^2 se = 0.1143; q^2 se = 0.2825; pred_r^2 = 0.9858; pred_r^2se = 0.0687

log1_MIC = - 1.1665 SddsN(nitro)E-index + 0.1243 T_C_Cl_3 - 0.0401 HydrogensCount + 0.1615 chi4pathCluster - 0.0092 T_2_T_6 + 4.8532

This model was generated by taking five compounds in test set and rest of the 15 compounds in training set. This model showed squared correlation co-efficient r^2 = 0.9591, q^2 = 0.7501 and pred_r^2 = 0.9858, having low standard error of cross validated squared correlation co-efficient (q^2-Se=0.2825). Five descriptors SddsN(nitro)E-index, HydrogensCount, T_C_Cl_3, T_2_T_6 and chi4pathCluster correlate with the biological activity. The fitness plot showed that all the compounds were near to the line [Graph-2]. In this model electrotopological descriptor SddsN(nitro)E-index (-32.00%), element count descriptor HydrogensCount (-16.88%) and alignment independent descriptor T_2_T_6 (-8.65%) showed negative contribution while alignment independent descriptor T_C_Cl_3 (+23.12%) and path cluster descriptor chi4pathCluster (+13.72%) showed positive correlation. [**chart-2]**

Graph -2: Fitness Plot between the experimental and Predicted Activities for model-2 [Training Set (red spots) and test set (blue spots)]

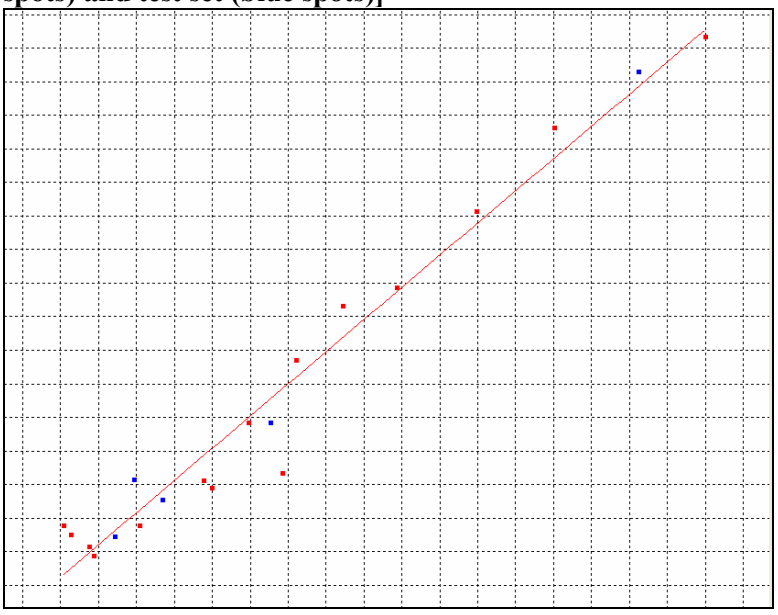
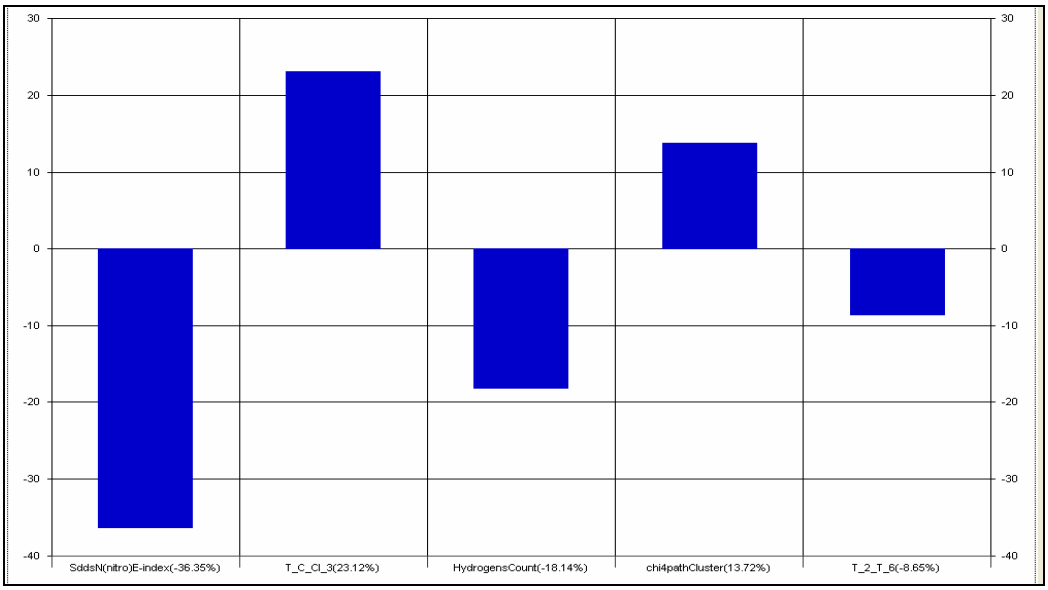


Chart: 2: Contribution chart of individual descriptors that are important for the biological activity in model -2



Model-3:- Optimum Components = 3; n = 15; Degree of freedom = 11; r^2 = 0.9792; q^2 = 0.9584; F test = 172.9860; r^2 se = 0.0887; q^2 se = 0.1256; pred_r^2 = 0.9239; pred_r^2se = 0.1260

log1_MIC = - 1.0752 SddsN(nitro)E-index + 0.1323 T_T_Cl_3 - 0.0768 T_C_C_2 + 0.4939 chi3Cluster - 0.1031 T_N_Cl_5 + 5.3748

This model was generated by taking five compounds in test set and rest of the 15 compounds in training set. The model-3 fulfill the selection criteria's with r^2 >0.8 (r^2= 0.9792) with low standard error of squared correlation co-efficient (r^2_Se) showed the relative good fitness [Graph-3] of the model and F value showed good statistical significance of the regression model. The model has cross validated squared correlation co-efficient (q^2 = 0.9584) for training set and pred_r^2 = 0.9239 for test set. This model showed standard error of cross validated squared correlation co-efficient (q^2-Se=0.1256). Five descriptors SddsN(nitro)E-index, T_T_Cl_3, T_C_C_2, chi3Cluster and T_N_Cl_5 correlate with the biological activity. In this model electrotopological descriptor SddsN(nitro)E-index (-37.26%), alignment independent descriptor T_N_Cl_5 (-6.08%) and T_C_C_2 (-21.70%) showed negative contribution while alignment independent descriptor T_T_Cl_3 (+21.98%) and cluster descriptor chi3Cluster (+14.90%) showed positive correlation. [**chart-3]**

Graph -3: Fitness Plot between the experimental and Predicted Activities for model-3 [Training Set (red spots) and test set (blue spots)]

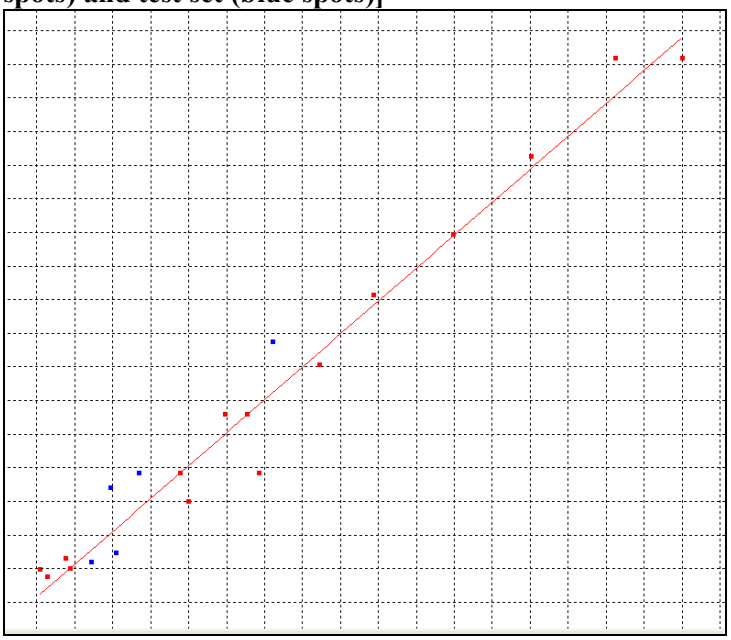
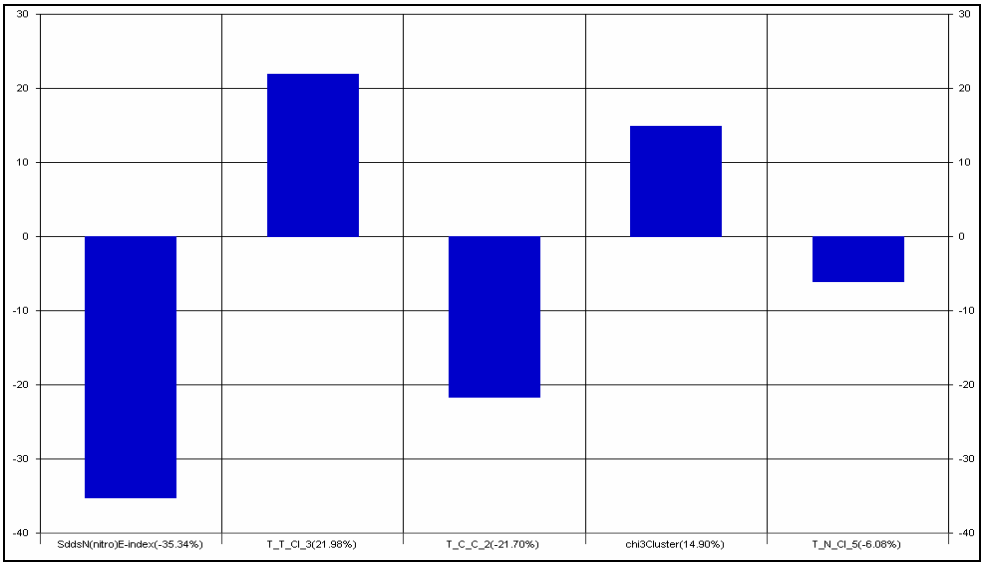


Chart: 3: Contribution chart of individual descriptors that are important for the biological activity in model -3



Model-4:- Optimum Components = 3; n = 15; Degree of freedom = 11; r^2 = 0.9788; q^2 = 0.9342; F test = 169.0687; r^2 se = 0.0877; q^2 se = 0.1544; pred_r^2 = 0.9215; pred_r^2se = 0.1453

log1_MIC = - 0.7407 SddsN(nitro)E-index + 0.1171 T_T_Cl_3 - 0.0433 HydrogensCount - 0.0710 SaaCHcount + 0.0765 T_2_N_7 + 5.6397

This model also showed low standard error of cross validated squared correlation co-efficient q^2-se and standard error of pred_r^2se, which show accuracy of the statistical calculation having r^2 = 0.9788 and q^2 = 0.9342 and also a good fitness plot [graph-4]. This model was generated by taking five compounds in test set and rest of the 15 compounds in training set. In this model five descriptor Hydrogen count(-19.99%), SaaCHcount (-11.55%), SddsN(nitro)E-index(-27.52%) showed negative contribution while the alignment independent descriptors T_T_Cl_3 (+22.00%) and T_2_N_7 (+19.334%) showed positive correlation. [**chart-4]**

Graph -4: Fitness Plot between the experimental and Predicted Activities for model-4 [Training Set (red spots) and test set (blue spots)]

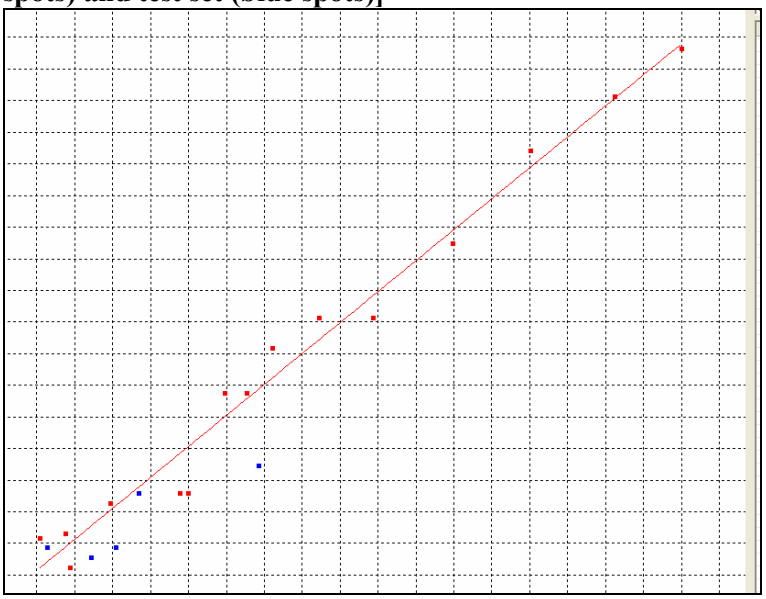
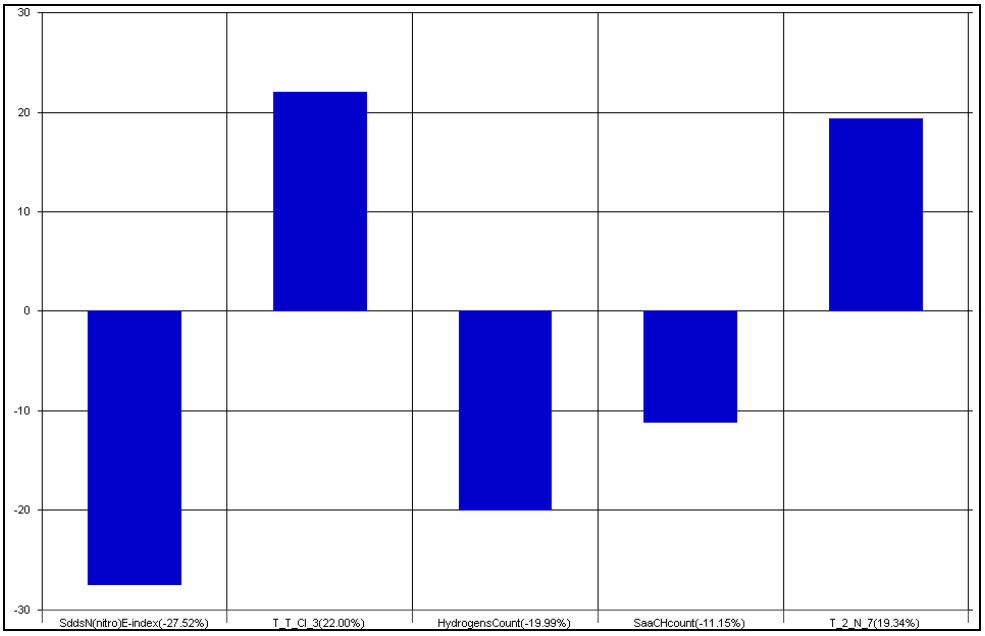


Chart: 4: Contribution chart of individual descriptors that are important for the biological activity in model -4



Model-5:- Optimum Components = 3; n = 15; Degree of freedom = 11; r^2 = 0.9888; q^2 = 0.9710; F test = 322.3163; r^2 se = 0.0658; q^2 se = 0.1056; pred_r^2 = 0.8710; pred_r^2se = 0.1519

log1_MIC = - 1.0898 SddsN(nitro)E-index + 0.1333 T_T_Cl_3 - 0.0727 T_C_C_2 + 0.4633 chi3Cluster - 0.0981 T_N_Cl_5 + 5.2908

This model was generated by taking five compounds in test set and rest of the 15 compounds in training set. The model-3 fulfill the selection criteria's with r^2 >0.8 (r^2=0.9769) with low standard error of squared correlation co-efficient (r^2_Se) showed the relative good fitness [Graph-5] of the model and F test = 131.2410 of the regression model. The validated model has cross validated squared correlation coefficient q^2 = 0.8787 for training set and pred_r^2 = 0.9255. This model showed low standard error of cross validated squared correlation co-efficient (q^2_Se) and standard error of pred_r^2se = 0.1551. The contribution chart showed that five descriptor SddsN(nitro)E-index, T_T_Cl_3, T_C_C_2, chi3Cluster and T_N_Cl_5 correlate with the biological activity. In this model SddsN(nitro)E-index (-36.33%), T_N_Cl_5 (-5.87%) and T_C_C_2 (-21.02%) showed negative contribution while alignment independent descriptor T_T_Cl_3 (+22.46%) and cluster descriptor chi3Cluster (+14.31%) showed positive correlation. [**chart-5].**

Graph -5: Fitness Plot between the experimental and Predicted Activities for model-5 [Training Set (red spots) and test set (blue spots)]

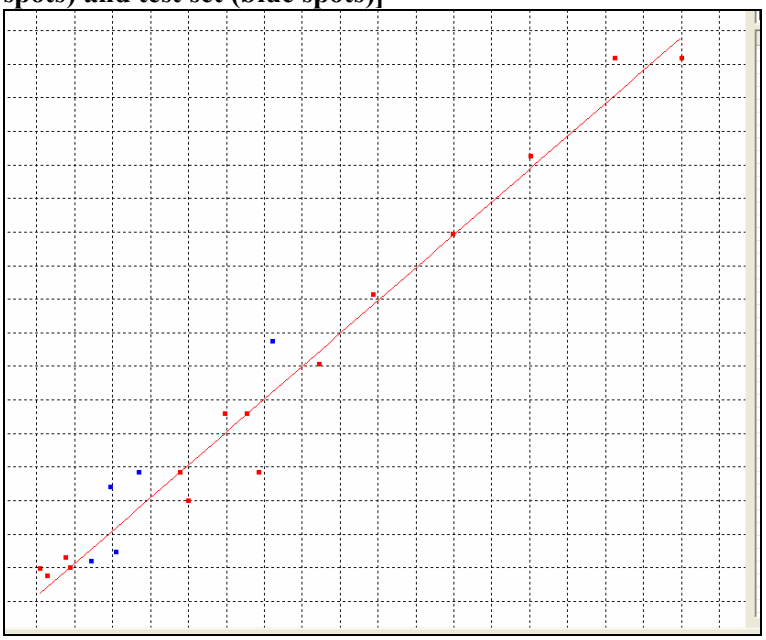
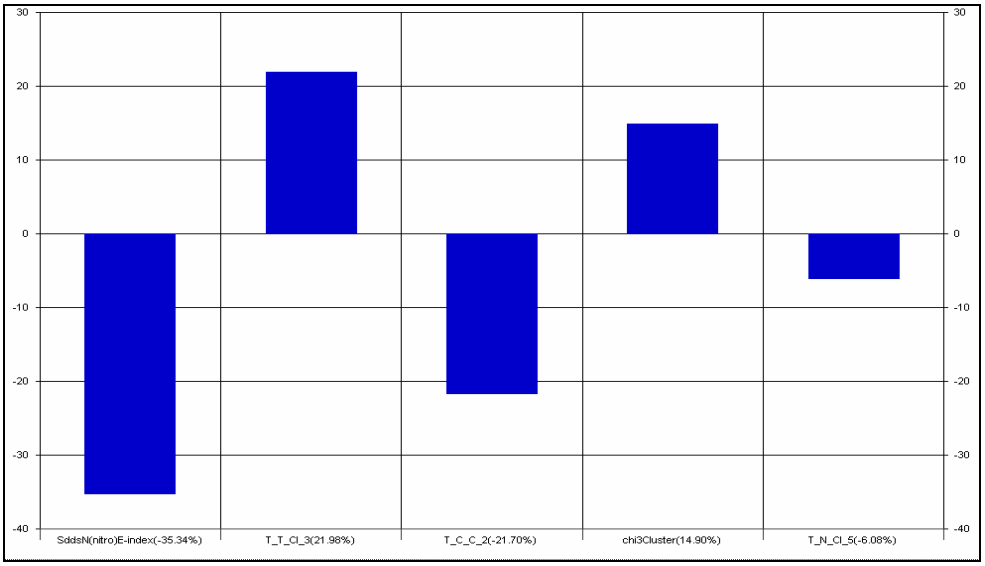


Chart: 5: Contribution chart of individual descriptors that are important for the biological activity in model -5



The developed models were analyzed to find common descriptors for the 2-azetidinone series, which are important for the antibacterial activity. The positive and negative contributions of the individual descriptors for the biological activity were studied and the model is used for the predictivity of the newly design compound. Descriptors selected for the 2-azetidinone series for their antibacterial activity against *B.subtilis* are given in table 2. Statistical parameters of the models are given in table 3 and actual and predicted activities of the model are given in table 4.

In this sequence, **SddsN(nitro)E-index, HydrogensCount** and **T_T_Cl_3** are common physico-chemical parameter in five models. **HydrogensCount and SddsN(nitro)E-index** shows negative contribution while **T_T_Cl_3** shows positive contribution for the activity. **SddsN(nitro)E-index**, Which is Electrotopological state indices for number of –nitro group connected with two double and one single bond. This descriptor signifies a retention index for number of –nitro group connected with two double and one single bond. So by decreasing the number of hydrogen atom in a compound and increasing the number of carbon atoms separated from chlorine atom by 3 bonds is desirable for designing of new compound.

All the parameters and their importance, which contributed to the antibacterial activity in the generated models, are discussed here:-

HydrogensCount: This descriptor signifies number of hydrogen atoms in a compound.

SddsN(nitro)E-index: Electrotopological state indices for number of –nitro group connected with two double and one single bond.

SaaCHcount: This descriptor defines the total number of carbon atoms connected with a hydrogen along with two aromatic bonds.

chi3Cluster: This descriptor signifies simple 3rd order cluster chi index in a compound.

chi4pathCluster: This descriptor signifies molecular connectivity index of 4th order pathcluster.

T_C_Cl_3: This is the count of number of Carbon atoms (single double or triple bonded) separated from any other Chlorine atom (single double or triple bonded) by 3 bonds in a molecule.

T_T_Cl_3: This is the count of number of any atoms (single double or triple bonded) separated from any other Chlorine atom (single double or triple bonded) by 3 bonds in a molecule.

T_C_C_2: This is the count of number of Carbon atoms (single double or triple bonded) separated from any other Carbon atom (single double or triple bonded) by 2 bonds in a molecule.

T_2_N_7: This is the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from nitrogen atom by 7 bonds.

T_2_T_6: This is the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from any atom by 6 bonds.

T_N_Cl_5: This is the count of number of Nitrogen atoms (single double or triple bonded) separated from any other Chlorine atom (single double or triple bonded) by 5 bonds in a molecule.

CONCLUSION:

From the detail study of 2D QSAR it was observed that the descriptors which are highly correlated with the biological activity of 2-azetidinone series are **Hydrogencounts, SddsN(nitro)E-index and T_T_Cl_3.** Hydrogencounts and SddsN(nitro)E-index showed negative contribution to the biological activity while T_T_Cl_3 showed positive contribution to the biological activity. It can be concluded that antibacterial activity of 4,4'-bis [3-chloro-4-aryl-azetidin-2-one-1-yl] diphenyl sulphone derivatives against B.subtilis is strongly influenced by the electrotopological parameter, element count and alignment independent descriptors. Consequently this study helped us to decide about the electronic and steric nature of substitution pattern around the selected 2-azetidinone nucleus. At various positions on the common template the substitution pattern was carried out and the same data was used for the design of new antibacterial agents. The regression equation obtained was used for prediction of activity of designed compounds in silico. The overall outcomes of these studies have provided great help to optimize the pharmacophore and to design the potent, selective antibacterial agent.

REFERENCES:

1. Holden K. G. and Morin R. B., Chemistry and Biology of β- lactam Antibiotics, Acadamemic press, London, 1982, 114
2. George G. I., The Organic Chemistry of *β*-Lactams, VCH Publications, New York, 1993.
3. Delpiccolo, C. M. L.; Fraga, M. A.; Mata, E. G. An Efficient, Stereoselective Solid-Phase Synthesis of β-Lactams Using Mukaiyama's Salt for the Staudinger Reaction. J. Comb. Chem. 2003, 5, 208-210.

4. Gootz T. D. Discovery and development of new antimicrobial agents. Clin. Microbiol. Rev. 1990, 3,13-31.
5. Maiti S. N, Babu R. P. K, Shan R. Overcoming bacterial resistance: Role of β-lactamase inhibitors. Top. Heterocycl. Chem. 2006, 2, 207-46.
6. Singh G. S. Beta-lactams in the new millennium. Part-1: monobactams and carbapenems. Mini. Rev. Med. Chem. 2004, 4, 69-92.
7. Singh GS. Recent progress in the synthesis and chemistry of azetidinones. Tetrahedron, 2003, 59,7631-49.
8. Risi C. D, Pollini G. P, Veronese A. C., Bertolasi V. A new simple route for the synthesis of (±)-2-azetidinones starting from β-enaminoketoesters. Tetrahedron, 2001, 57,10155- 161.
9. Durckheimer W, Blumbach J, Lattrell R, Scheunemann K. H. Recent developments in the field of β-lactam antibiotics. Angew. Chem. Int. Ed. Engl. 1985, 24,180-202.
10. Burnett D. A, Caplen M. A, Jr. Davis H. R, Burrier R. E, Clader J.W. 2-Azetidinones as inhibitors of cholesterol absorption. J. Med. Chem. 1994, 37,1733-6.
11. Slusarchyk W. A, Bolton S. A, Hartl K. S, Huang M. H, Jacobs G, Meng W, Ogletree M. L, Pi Z, Schumacher W. A, Seiler S. M, Sutton J. C, Treuner U, Zahler R, Zhao G, Bisacchi G. S. Synthesis of potent and highly selective inhibitors of human tryptase. Bioorg. Med. Chem. Lett. 2002, 12, 3235-8.
12. Han W. T, Trehan A. K, Wright J. J. K, Federici M. E, Seiler S. M, Meanwell N. A. Azetidin-2-one derivatives as inhibitors of thrombin. Bioorg. Med. Chem. 1995, 3, 1123-43.
13. Aoyama Y, Uenaka M, Kii M, Tanaka M, Konoike T, Hayasaki- Kajiwara Y, Naya N, Nakajima M. Design, synthesis and pharmacological evaluation of 3-benzylazetidine-2-onebased human chymase inhibitors. Bioorg. Med. Chem. 2001, 9, 3065-75.
14. Guillon C. D, Koppel G. A, Brownstein M. J, Chaney M. O, Ferris C. F, Lu S-f, Fabio K. M, Miller M. J, Heindel N. D, Hunden D. C, Cooper R. D. G, Kaldor S. W, Skelton J. J, Dressman B. A, Clay M. P, Steinberg M. I, Bruns R. F, Simon N. G. Azetidinones as vasopressin V1a antagonists. Bioorg. Med. Chem. 2007, 15, 2054-80.
15. Goel R. K, Mahajan M. P, Kulkarni S. K. Evaluation of antihyperglycemic activity of some novel monocyclic beta lactams. J. Pharm. Pharm. Sci. 2004, 7, 80-3.
16. Kumar A, Rajput CS. Synthesis and anti-inflammatory activity of newer quinazolin-4-one derivatives. Eur. J. Med. Chem. 2009,44, 83-90.
17. Srivastava S. K, Srivastava S, Srivastava S. D. Synthesis of new carbazoyl-thiazol-2-oxo-azetidines antimicrobial, anticonvulsant and anti-inflammatory agents. Indian J. Chem. B, 1999, 38,183-7.
18. Sperka T, Pitlik J, Bagossi P, Tozser J. Beta-lactam compounds as apparently uncompetitive inhibitors of HIV-1 protease. Bioorg. Med. Chem. Lett. 2005, 15, 3086-90.
19. Knight W. B, Green B. G, Chabin R. M, Gale P, Maycock A. L, Weston H, Kuo D. W, Westler W. M, Dorn C. P. Specificity, stability and potency of monocyclic beta-lactam inhibitorsof human leukocyte elastase. Biochemistry, 1992, 31, 8160-70.
20. Firestone R. A, Barker P. L, Pisano J. M, Ashe B. M, Dahlgren M. E. Monocyclic β-lactam inhibitors of human leukocyte elastase. Tetrahedron, 1990, 46, 2255-62.
21. Vergely I, Laugaa P, Reboud-Ravaux M. Interaction of human leukocyte elastase with a *N*-aryl azetidinone suicide substrate: Conformational analyses based on the mechanism of action of serine proteinases. J. Mol. Graph. 1996, 14, 158-67.
22. Singh R, Micetich RG. Monobactams as enzyme inhibitors. IDrugs 2000, 3, 512-7.
23. Veinberg G, Bokaldere R, Dikovskaya K, Vorona M, Mucel D, Kazhoka Kh, Turovskis I, Shestakova I, Kanepe I, Domrachova I, Lukevics E. Synthesis of penicillin derivatives and study of their cytotoxic properties. Chem. Het. Compd. 1998, 34, 1266-75.
24. Mehta P. D, Sengar N. P. S, Pathak A. K. 2-Azetidinone--a new profile of various pharmacological activities. Eur. J. Med. Chem. 2010, 45, 5541-60.
25. Randic M, Encyclopedia of Computational Chemistry, P.V.R Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, P. R. Schreiner Eds, John Wiley & Sons, Chinchester,1998, 3018-3032.
26. Geladi P, Kowalski B. R, Partial Least Squares: A Tutorial. Anal Chem Acta, 1986, 185, 1-17
27. Norinder U, Single Domain Mode Variable Selection in 3D QSAR Applications, J of Chemometrics, 1996, 10, 95-105
28. Mehta P. D, Pathak A. K, Snthesis, characterization and in vitro antimicrobial activity of novel 4,4'-bis [3-chloro-4-aryl-azetidin-2-one-1-yl] diphenyl sulphones, Bull. Pharm. Rearch, 2011, 1, in press..

29. MDS 1.0: "Molecular Design Suite, VLife Sciences Technologies » Pvt. Ltd. Pune, India, 2003. See www.vlifesciences.com
30. K. Baumann, An alignment-independent versatile structure descriptor for QSAR and QSPR based on the distribution of molecular features, Journal of Chemical Inf. Computer Science, 2002, 42, 26-35.
31. Cho S. J, Tropscha A, Crossvalidated R2-guided Region Selection for Comparative Molecular field Analyses: A Simple Method to Achieve Consistent Results, J Med Chem, 1995, 38, 1060-1066.
32. Cramer RD, Bunce JD, Patterson DE, Frank IE, Cross-Validation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies, Quant Struct-Act Relat, 1988, 7, 18-25
33. Lowrey AH, Cramer CJ, Urban JJ, Famini GR, Quantum-chemical descriptors for linear salvation energy relationships, Comput Chem, 1995, 199, 209-215
34. VLifeMDS Molecular Design Suite, Vlife Sciences Technologies Pvt. Ltd., Pune, India, (www.vlifesciences.com).

*****