QSAR MODELING OF BETA-LACTAM ANTIBIOTIC CEPHALOSPORIN AGAINST TRANSPEPTIDASE USING MLR METHOD

AJEET*

Department of Pharmaceutical Chemistry, S. D. College of Pharmacy and Vocational Studies, Muzaffarnagar 251001, India

ABSTRACT

The objective of the work was to develop a Quantitative Structure Activity Relationship (QSAR) model for understanding the effect of eccentric connective index, fragment complexity and topological polar surface area on the inhibitive activity of cephalosporins. For developing the model, Multiple Linear Regression (MLR) has been employed as a effective and efficient method and this model has been validated with statistical analysis such as fraction of variance, cross validation test, quality factor, fischers test, standard deviation, internal validation test (Y- randomization test), external validation test. A regression based QSAR model has been developed with cross validation test q2 = 0.9013 and fraction of variance r2 = 0.9014, i.e. >90% predictive efficiency and all the statistical tests have validated this model. The developed QSAR model reveals that the cephalosporin derivatives must have more eccentric connectivity index as well as topological polar surface area for enhanced transpeptidase inhibitory action at R1 and R2 substituents. A negative coefficient of fragment complexity containing R1 and R2 substituents decreases the activity of cephalosporin derivatives towards its inhibitory action of transpeptidase.

Keywords: Cephalosporin analogues; Eccentric Connectivity index; Fragment Complexity; Topological Polar Surface Area

INTRODUCTION

Cephalosporins belong to β-lactam antibiotics, which was originally derived from fungus *Cephalosporium Acremonium* by an Italian scientist Giuseppe Brotzu whose basic moiety is 7-Aminocephalosporanic acid. Cephalosporin belongs to the category of β-lactum antibiotics which has a basic moiety of 7-Aminocephalosporanic acid It was observed that the cultures of *Cephalosporium acremonium* inhibited *Salmonella typhi*. For the prophylaxis and treatment of bacterial infections, cephalosporins are frequently recommended. Being bactericidal in nature, 7-Aminocephalosporanic acid ruptures the peptidoglycan layer of cell wall, as it is important for structural integrity of cell wall. Peptidoglycan synthesis in bacterial cell wall is facilitated by transpeptidase. 7-Aminocephalosporanic acid acts by inhibition of transpeptidase through penicillin binding proteins-PBP's [1]. Cephalosporins are widely used as broad spectrum antibiotic, although it has many side effects.

The QSAR studies are perfect tool for understanding the drug design process in terms of their chemical and pharmacological activity interaction, along with this it is also used in toxicology and pesticide research [2-3]. QSAR studies can focus on mechanism of action of ligands with human, bacteria, virus, membranes, enzymes etc. It can also be used for the evaluation of metabolism, absorption, distribution and excretion phenomena. The QSAR methodology comprises of computationally derived descriptors to correlate with pharmacological activities. These descriptors are principally of four types such as electronic, stearic, hydrophobic and topological indices [4]. It can be examined in tandem with equations of a similar mechanistic genre to establish its authenthicity and reliability. Various QSAR studies have been done to correlate the pharmacodynamic and pharmacokinetic behavior of cephalosporin using different descriptors. Genetic algorithm, partial least squares method and their combination has been used to develop QSAR models. In the present study, we have developed a QSAR model on a series of cephalosporin analogues with respect to their inhibitory activity towards transpeptidase inhibition. The descriptors used are eccentric connectivity index (ECI)[5], fragment complexity (FC)[6] and topological

***Corresponding author:**
ajeet_pharma111@rediffmail.com

polar surface area (TPSA)[7]. It is for the first time that eccentric connectivity index (ECI), fragment complexity (FC) and topological polar surface area (TPSA) are used as descriptors to correlate the transpeptidase inhibitory activity of cephalosporins using multiple linear regression (MLR) analysis.

MATERIALS AND METHOD

Modeling parameters and structure optimization

The 2D structure construction, energy minimization and geometry optimization of the selected benzimidazole derivatives were carried out by using ChemDraw Ultra 7.0 and Chem3D Pro 7.0 (CambridgeSoft Corporation, 100 CambridgePark Drive, Cambridge MA, 02140 USA) on an Intel(R) Core(TM)2 Duo Central Processing Unit T6670 @ 2.20 GHz and 4.00 GB of RAM, running the Windows 7 Home Basic, 64-bit compatible operating system. The energy minimization was carried out to minimum RMS Gradient of 0.100, with step interval of 2.0 Fs and frame interval of 10. Descriptors calculated for the training set are given below-

ALOGP	Molecular linear free energy relation
Aromatic atoms count	Petitjean number
Aromatic bonds count	Rotatable bonds count
Atom count	Rule of five
Autocorrelation (arge)	TPSA
Autocorrelation (Mass)	VadjMa
Autocorrelation (Polarizability)	Weight
BCUT	Weighted path
Bond count	Wiener numbers
BPol	XlogP
Carbon types	Zagreb index
Chi chain	CPSA
Chi cluster	Gravitational index
Chi path cluster	Length over breadth
Chi path	Moment of inertia
Eccentric connectivity index	Petitjean shape index (topoShape, geomShape)
Atom type electro-topological state	WHIM
Fragment complexity	Largest Pi system
Hbond acceptor count	Longest aliphatic chain
Hbond donor count	Mannhold LogP
Kappa shape indices	McGowan volume
	MDE

Descriptor selection

The selection of descriptors among the calculated descriptors for the multiple linear regression analysis is based on the correlation matrix. This matrix is prepared and analyzed for the least correalated descriptors. The correlation matrix is given in Matrix 1.

Matrix 1. Correlation matrix

	ECI	FC	TPSA
ECI	1	0.9222	0.8456
FC	0.9222	1	0.6284
TPSA	0.8456	0.6284	1

All the bioactivity values and information about 2D structure of cephalosporin derivatives were taken from literature[1]. IC_{50} is referred as the molar concentration of a compound that inhibits 50% growth of bacteria[8]; -log1/IC_{50} is subsequent variable that comprises the bioactivity parameter for the QSAR model. In order to calculate the 2D molecular descriptors, PaDEL descriptor software[9, 10] which incorporates CDK library has been used. For the development of QSAR model, multiple linear regression analysis was employed.

Molecular descriptors

Eccentric connectivity index denoted by ξ is defined as the summation of the product of eccentricity and the degree of each vertex in the hydrogen suppressed molecular graph having n vertices. Eccentric connectivity index takes into consideration the eccentricity as well as valency of the vertices in a hydrogen suppressed graph[5].

Topological polar surface area is calculated from molecular bonding informations only. Hou et. al. has used the solvent-accessible surface area for calculating the topological polar surface area. Different procedures of surface calculations, even different calculation parameters may generate different TPSA. It may be calculated as-

$$TPSA = \sum n_i s_i$$

Where n_i is the frequency of fragment i in the molecule, s_i is the surface contribution of type i [7].

Statistical parameters

Table 1. Biological (-log 1/IC_{50} mol/L^{-1}), physiochemical and structural parameters of cephalosporin derivatives used to derive QSAR equation.

S. No.	Name	-log 1/IC_{50}			ECI[a]	FC[a]	TPSA[a]
		Observed	Predicted	Residuals			
1	Cefaclor	4.66	4.941319	-0.28132	484	1048.09	138.03
2	Cefadroxyl	4.96	4.904943	0.055057	540	1336.09	158.26
3	Cefazolin	5.96	5.977663	-0.01766	756	1304.15	227.65
4	Cefmenoxime	6.01	5.992645	0.017355	872	1753.17	262.9
5	Cefmetazole	6.28	5.958325	0.321675	705	1269.14	226.72
6	Cefodizime	5.77	6.061374	-0.29137	1132	2268.17	297.54
7	Cefotaxime	5.66	5.630366	0.029634	717	1531.14	223.58
8	Cefoxitin	5.63	5.475239	0.154761	579	1199.11	189.63
9	Cefsulodine	5.13	4.976178	0.153822	1043	2340.14	220.86
10	Ceftazidime	4.84	4.826629	0.013371	1035	2637.15	237.59
11	Ceftizoxime	5.58	5.720816	-0.14082	500	1000.12	197.28
12	Cephalexin	4.64	4.654501	-0.0145	484	1297.08	138.03

[a] ECI: eccentric connectivity index, FC: fragment complexity, TPSA: topological polar surface area

Table 2. Statistical results of model validation

$n/p \geq 4$	r^2	q^2	S	$r^2 - q^2 < 0.3$	Q	F
4	0.9014	0.9013	0.53144	0.0001	1.786	24.3786

Table 3. Y-Randomization Test results for QSAR model

No. of Y-randomization	First	Second	Third	Fourth	Fifth
r^2	0.2866492	0.4621214	0.0769552	0.0671725	0.1802311

Table 4. Biological (-log 1/IC_{50} mol/L^{-1}), physiochemical and structural parameters of positive controls used for external validation of QSAR equation.

No.	Name	-log 1/IC_{50}			ECI[a]	FC[a]	TPSA[a]
		Observed	Predicted	Residuals			
1	Cefamandole	5.97	4.909944	1.060056	891	2033.13	197.86
2	Cefipime	5.68	4.155008	1.524992	801	2608.13	197.28
3	Cefixime	5.05	5.952707	-0.90271	685	1339.14	234.58
4	Cefoperazone	4.91	4.318894	0.591106	1451	3733.19	267.58
5	Cefotetan	5.16	6.822363	-1.66236	962	1582.18	308.62
6	Cefotiam	5.72	5.089263	0.630737	1031	2478.16	244.55
7	Ceftriaxone	5.74	6.186724	-0.44672	1048	1989.18	287.84
8	Cephalothin	6.2	5.063668	1.136332	579	1286.1	163.61
9	Cephradine	4.94	4.451774	0.488226	484	1473.08	138.03

ECI: eccentric connectivity index, FC: fragment complexity, TPSA: topological polar surface area

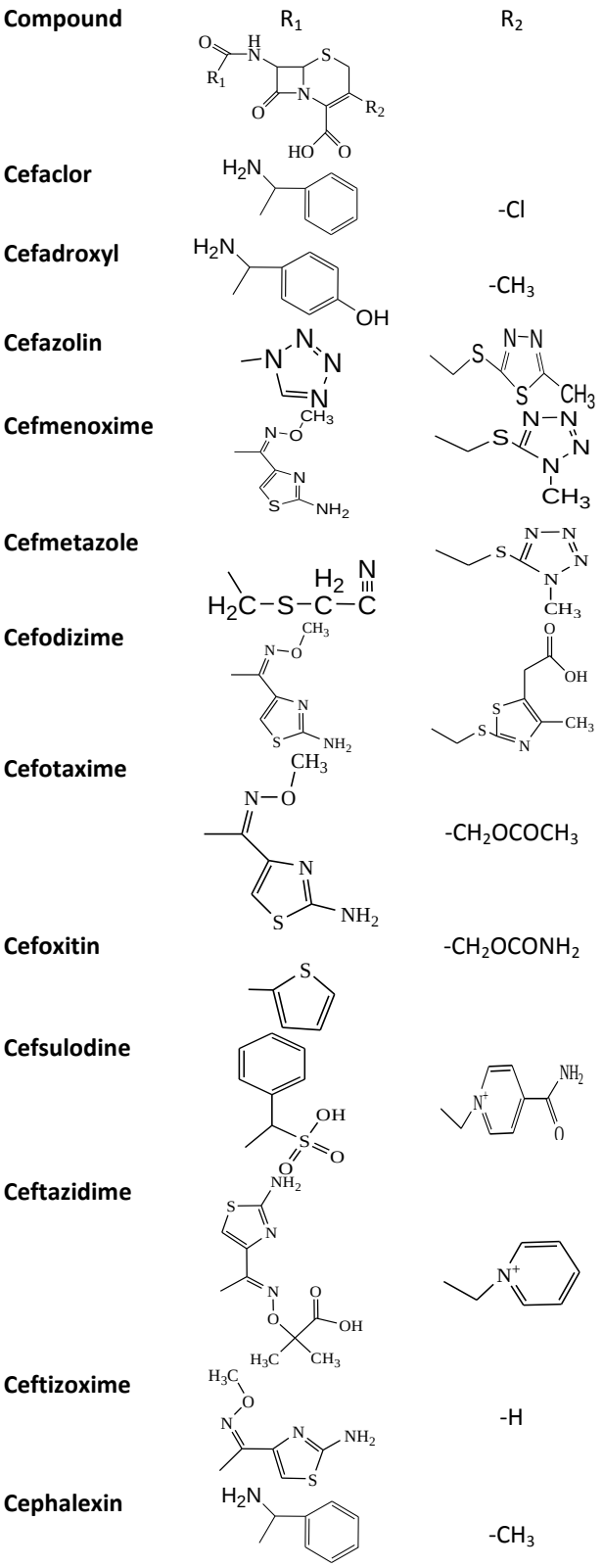


Fig. 1. Structure of cephalosporin pharmacophore and its derivatives for which the QSAR model is developed

(fraction of variance), cross-validated r^2 is denoted as q^2, s is standard deviation. Q is quality factor, where *Q = r/s* (here r is correlation coefficient and s is standard deviation). Fischer statistics is denoted by F.

Model validation

The QSAR model validation was carried with statistical analysis, internal validation and with external validation methods.

Results and Discussion

The 2D structure of cephalosporin pharmacophore and its derivatives for which the QSAR model has been developed is shown in Figure 1. From the data in Table 1, QSAR equation was developed where number of data point (n) is 12, is given below. Here 95% confidence intervals are given in parantheses.

$$-\log(1/IC_{50}) = 4.03627\ (\pm 0.8164189) + 0.0009517\ (\pm 0.0033934)\ (ECI) - 0.0011519\ (\pm 0.0010014)\ (FC) + 0.0119664\ (\pm 0.0081855)\ (TPSA)$$

A comparison (multiple linear regression curve) of observed values and predicted values of *–log (1/IC$_{50}$)* for cephalosporin derivatives used for development of QSAR equation is shown in Figure 2 and multiple linear graph is shown in Figure 3.

Validation of QSAR model

A quantitative assessment of model robustness has been performed through model validation. All the statistical results of model validation have been given in Table 2.

Statistical analysis

n/p ratio*: *n/p = ≥ 4, where *n* is the number of data points and *p* is the number of descriptors used in the QSAR model. The model obeys the condition.

***Fraction of variance (r^2)*:** The value of fraction of variance may vary between 0 (means model without explanatory power) and 1 (means perfect model). QSAR model having $r^2 > 0.6$ will only be considered for validation[8]. The value for this QSAR model is 0.9014.

***Cross-Validation Test (q^2)*:** A QSAR model must have $q^2 > 0.5$ for the predictive ability[8]. The value of q^2 for this QSAR model is 0.9013.

***Standard deviation (s)*:** The smaller s value is always required for the predictive QSAR model. The value of s for this QSAR model is 0.53144.

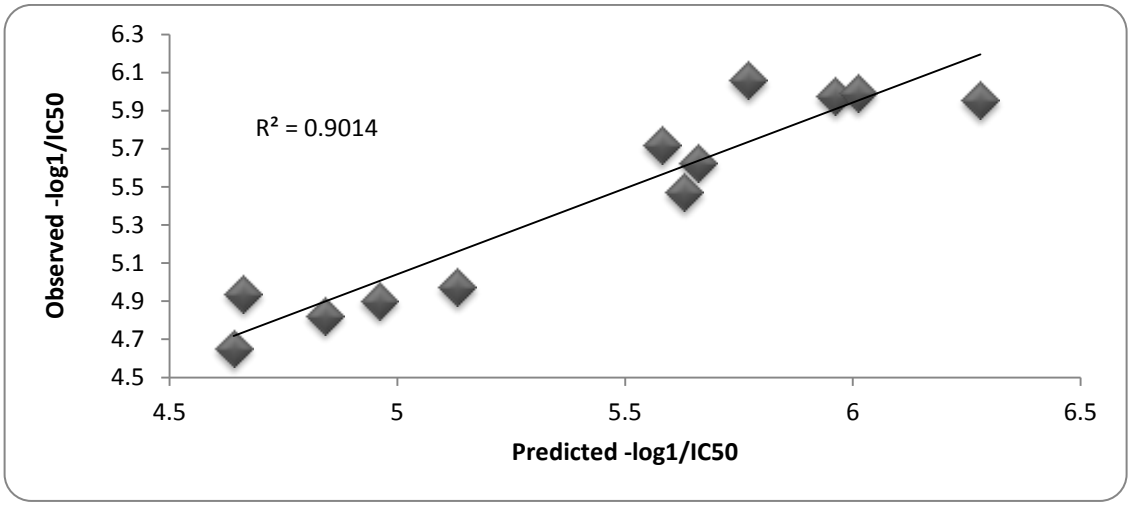


Fig. 2. A plot of observed values and predicted values of $-\log 1/IC_{50}$ for cephalosporin derivatives.

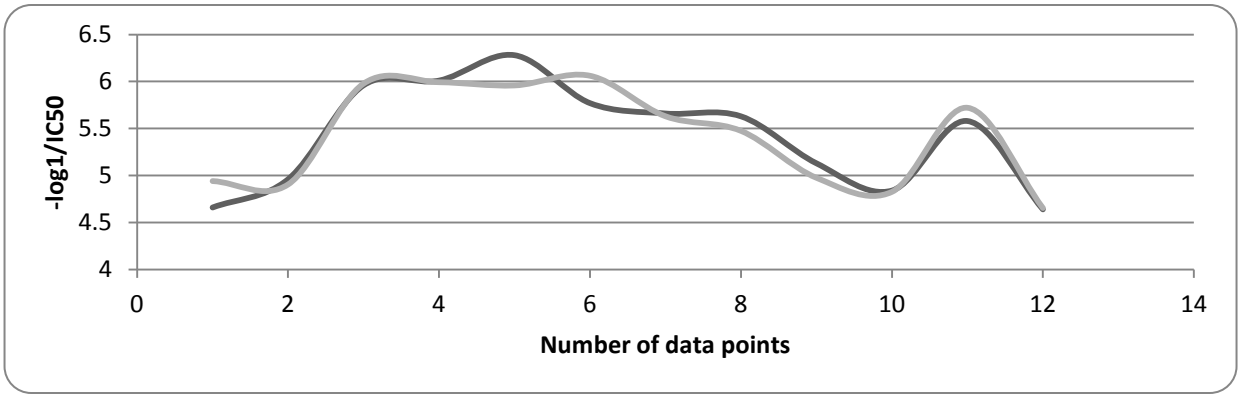


Fig. 3. Multiple linear graph of observed values and predicted values of $-\log(1/IC_{50})$ for cephalosporin derivatives used for development of QSAR equation. (Blue color series- observed bioactivity values, Red color series- predicted bioactivity values)

***r^2-q^2 < 0.3*:** The difference between r^2 and q^2 should never be exceeding by 0.3. A large difference suggests the following: presence of outliers, over-fitted model, and presence of irrelevant variables in data[8]. The value of r^2-q^2 for this QSAR model is 0.0001.

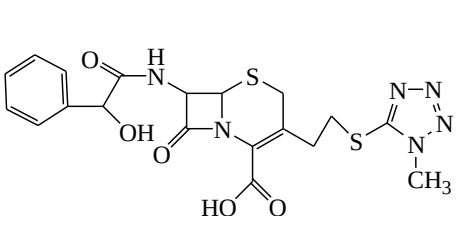
***Quality Factor (Q)*:** Overfitting and chance correlation, due to excess number of descriptors, can be detected by Q value. Positive value for this QSAR model suggests its high predictive power and lack of overfitting[8].

***Fischer Statistics (F)*:** The F value of QSAR model was compared with their literature value at 95% level. The F value of this QSAR model is 24.3786 (where $F > F_{lit}$) suggests that the QSAR model is statistically signifigant at 95% level.

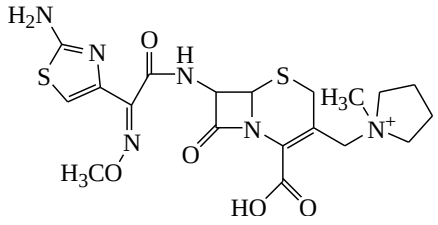
Internal Validation

Y-Randomization Test: To establish the QSAR model robustness, this technique is being used widely. For this test, the dependent variable vector is randomly shuffled, and a new QSAR model is developed using the unchanged independent variable. This process was repeated for five times. The statistical data of r^2 for five runs are given in Table 3. The values r^2<0.6 in Y-randomization test confirm the robustness of this QSAR model[8].

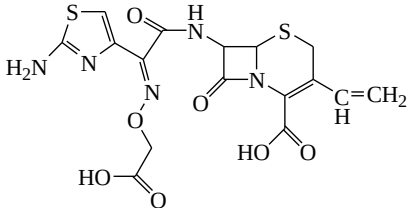
External Validation



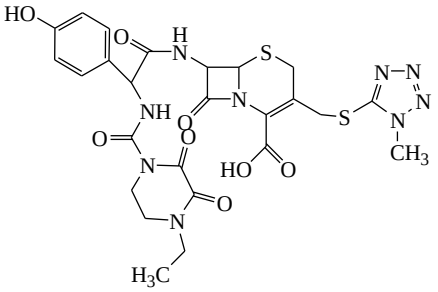
Cefamandole



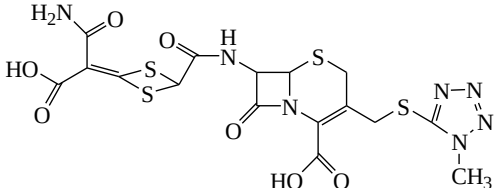
Cefipime



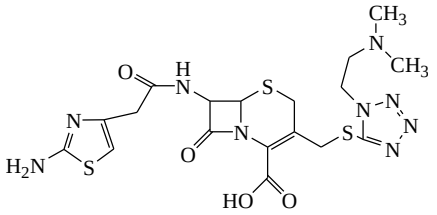
Cefixime



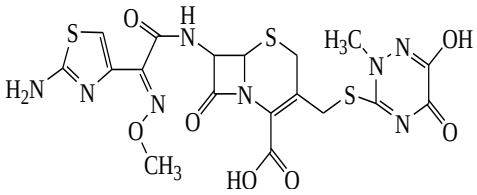
Cefoperazone



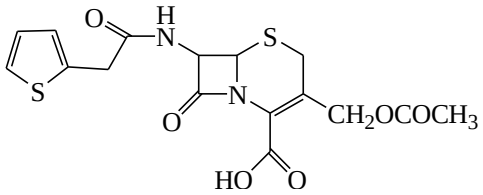
Cefotetan



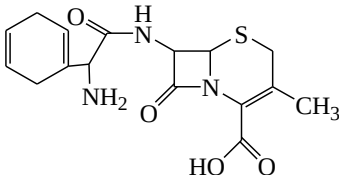
Cefotiam



Ceftriaxone



Cephalothin



Cephradine

Fig. 4. The chemical structures of positive controls used for external validation of QSAR model.

Prediction of bioactivity of positive control: For predicting the bioactivity value from this QSAR model, the positive control (Figure 4) i.e. the chemical structures with same pharmacophore or of

the same category as of training set (not included in training set) is being tested. The predicted values of positive control are given in Table 4. The low value of residuals confirms the robustness of this QSAR model.

CONCLUSION

An analysis of QSAR equation shows that ECI, FC and TPSA are the important determinants for the antibacterial activities of cephalosporin analogue. The inhibitory activity is mainly dependent on eccentric connectivity index (ECI) with a major contribution coming from fragment complexity (FC) and topological polar surface area (TPSA). The developed QSAR model reveals that the cephalosporin derivatives must have more eccentric connectivity index for enhanced transpeptidase inhibitory action at R_1 and R_2 substituents. A negative coefficient of fragment complexity containing R_1 and R_2 substituent decreases the activity of cephalosporin derivatives towards its inhibitory action of transpeptidase. Moving towards the effects of the topological polar surface area on the bioactivity of derivatives of cephalosporin, the developed QSAR model suggests that an increment in TPSA at substituents R_1 and R_2 will definitely be favorable to the activity.

REFERENCES

[1] Saha, R. N., Raman, M. Modeling of Biological Activity and Pharmacokinetics of Cephalosporin Derivatives. *Indian J. Pharm. Educ. Res.* 42(3), 207-14, (2008).

[2] Hansch, C., Maloney, P. P., Fujita, T., Muir, R. M. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature.* 194, 178-80, (1962).

[3] Bermudez-Saldana, J. M., Cronin, M. T. D. Quantitative structure–activity relationships for the toxicity of organophosphorus and carbamate pesticides to the Rainbow trout. Onchorhyncus mykiss. *Pest. Manag. Sci.* 62, 819-831, (2006).

[4] Verma, R. P., Hansch, C. Matrix metalloproteinases (MMPs): Chemical–biological functions and (Q)SARs. *Bioorg. Med. Chem.* 15, 2223-2268, (2007).

[5] Sardana, S., Madan, A. K. Topological models for prediction of antihypertensive activity of substituted benzylimidazoles. *J. Molecular Structure (Theochem).* 638, 41-49, (2003).

[6] Gregg, S., Eiso, A. B., Jan, S. Integration of fragment screening and library design. *Drug Discovery Today.* 12, 1032-1039, (2007).

[7] Hou, T. J., Xu, X. J. ADME Evaluation in Drug Discovery. 3. Modeling Blood-Brain Barrier Partitioning Using Simple Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* 43, 2137-2152, (2003).

[8] Verma, R. P., Hansch, C. QSAR modeling of taxane analogues against colon cancer. *Eur. J. Med. Chem.* 45, 1470-1477, (2010).

[9] Yap, C. W. PaDEL-descriptor: open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* 32, 1466-1474, (2011).

[10] Ajeet, Tripathi, L., Kumar P., Designing of Novel 6(H)-1,3,4-Thiadiazine Derivatives as MMP12 Inhibitors: A MLR and Docking Approach. *American Journal of Pharmacological Sciences,* 1(2), 29-34, (2013).