iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 1 **Issue:** IV **Month of publication:** November 2013

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

MLR Study of IC50 values for Cephalosporin derivatives by MS Excel

*Rajendra Kumar Sharma[1], Arun Sikarwar1

1 Department of Chemistry, Government Home Science College, Hoshangabad-461001 (M.P.) INDIA

rksharma210375@gmail.com

Abstract:

In this paper there is a study of multiple linear regression analysis (MLR) to develop a regression equation for prediction of IC_{50} values of cephalosporin derivatives. IC_{50} values are considered in their –log IC_{50} i.e. pIC_{50} terms. MLR is applied in stepwise regression analysis by means of forward selection method. 22 compounds of cephalosporin were selected for present study. 3D structures of these compounds were prepared with the help of chemsketch and saved as a mol file. Various indices like winer, detour, randic, harary, balban, schultz molecular topological indices, topological distance indices T(N-N), T(N-S), T(N-O), 3D Morse indices, Sum of Keir-Hall electro topological state (S_s), various combinations of 3D Morse and Keir Hall topological state, $(N_S)^X$, α_{XX} were calculated with the help of Dragon software. The IC_{50} values in terms of pIC_{50} were collected from literature. Then a stepwise regression analysis between pIC_{50} and selected indices was carried out by Microsoft excel software. A multiple linear regression equation was developed. There occurs strong correlation (R^2 = 0.642, pearson product moment correlation coefficient, r^2 = 0.801) between observed and predicted values of pIC_{50} by developed regression equation. The developed equation/ model can be employed for suggesting some other cephalosporin molecules with improved IC_{50} which can be proposed for further practical study/ verification.

Keywords: MLR, topological distances, β-lactum antibiotic, 2D and 3D indices, pIC_{50}.

INTRODUCTION

Cephalosporins are broad-spectrum antibiotics, applied for the treatment of bacterial infections caused by susceptible organisms. such as respiratory tract infections (pneumonia, strep throat, tonsillitis, and bronchitis), skin infections and urinary tract infections, surgical prophylaxis prevention of bacterial infection before, during, and after surgery, open, laparoscopic or endoscopic surgery intravenous prophylaxis etc.'s. They are sometimes given with other antibiotics. Cephalosporins [1] belong to class of β-lactum type antibiotic. β-Lactam antibiotics (beta-lactam antibiotics) are a broad class of antibiotics. Most β-lactam antibiotics work by inhibiting cell wall biosynthesis in the bacterial organism and are the most widely used group of antibiotics. Up to year 2003, in the view of sales, more than half of all commercially available antibiotics in use were β-lactam compounds. Penicillin and most other β-lactam antibiotics [2] act by inhibiting penicillin-binding proteins, which normally catalyze cross-linking of bacterial cell walls. In the absence of β-lactam antibiotics, the bacterial cell wall plays an important role in bacterial reproduction. Adding β-lactam antibiotics to the cell medium while bacteria are dividing will cause them to shed their cell walls and fail to divide, forming large, fragile spheroplasts.

β-lactam antibiotics are bacteriocidal [3], and act by inhibiting the synthesis of the peptidoglycan layer of bacterial cell walls. The peptidoglycan layer is important for cell wall structural integrity, especially in Gram-positive organisms, being the

outermost and primary component of the wall. The final transpeptidation step in the synthesis of the peptidoglycan is facilitated by transpeptidases known as penicillin-binding proteins (PBPs). PBPs vary in their affinity for binding penicillin or other β-lactam antibiotics.

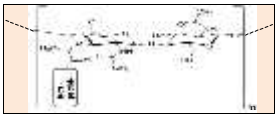


Figure 1: Chemical structure of peptidoglycan molecule.

Core structure of cephalosporin and related structure property relationship can be shown by following diagram:

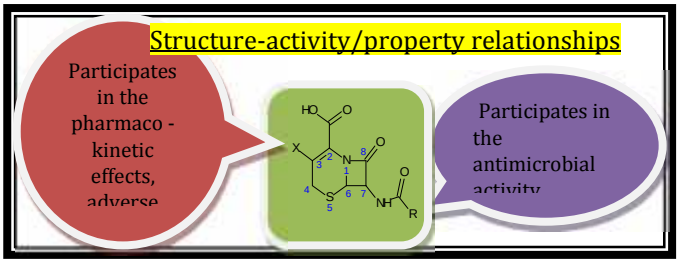


Figure 2: Structure-function relationships for cephalosporin core molecule

In the structure of cephalosporin as given in figure no. 2, nature of group –R and –X affects [4] activity as well as property of cephalosporin. Kinetic studies combined with absorption and nuclear magnetic resonance spectroscopy have shown the structure of the opening of the β-lactam ring of cephalosporyl. Cephalosporyl is the speices formed by the aminolysis of β-lactum ring of cephalosporin due to reaction with amino acid's amino group present in cell wall of bacterial cell wall. Opening of β-lactam ring takes place either in a concerted fashion or in some stages. The opening of the β-lactam ring leads to elimination of the group -X when this is configured as a leaving group. The process is well documented chemically and this property has been used as a strategy to obtain cephalosporins that can apply in a double action way. When the -X is conformed as the inactive form of the drug, the action of the β-lactam in the cephalosporin implies the release of the drug in situ. The cephalosporin nucleus can be modified to gain different properties by insertion of suitable –R group and –X group to obtain desirable activity/property. This task can be completed by QSAR/QSPR.

Material and methods:

QSAR/ QSPR i.e. Quantitative Structure Activity Relationship/ Quantitative Structure Property Relationship Studies for series of cephalosporin compounds consist the following steps:

1. To select accurate experimental physicochemical properties from compilations and studies

(i.e. from experimental or journal.)

2. To implement these properties and cephalosporin molecular descriptors to produce various QSAR/ QSPR equations in terms of single or multiple regression equations [5], [6].

$$Y = aX_1 + bX_2 + cX_3 + \ldots\ldots + nX_n$$

3. To determine the best QSPR equation on the basis of best prediction for each Physico-chemical property.

4. To propose a model i.e. MLRM (multi linear regression model).

Example of a QSAR/ QSPR model:

$$Y = aX_1 + bX_2 + cX_3 + \ldots\ldots + nX_n + \varepsilon$$

Where, Y = Physicochemical property as a dependent variable,

X_1, X_2, X_3, ... X_n = Specific molecular descriptor as independent variables.

a, b, c,……n = Regression constants., ε = Error

When correlation is not strong between dependent and independent variable then any single independent variable can not be used for satisfactory prediction of dependent variable. So multiple linear regression analysis is to be applied to increase correlation by adding suitable one or more indices in various steps. This can be done in two ways [7], [8] –

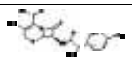
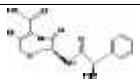
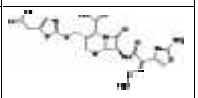
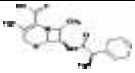
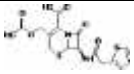
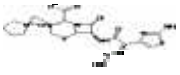
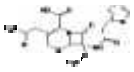
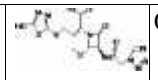
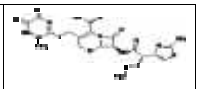
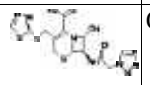
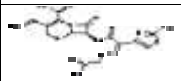
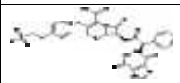
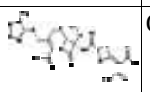
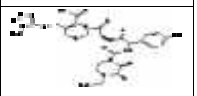
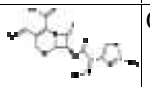
(a) By forward selection method, in which each variable index is added in various steps one by one in regression analysis. In each step variable index is selected on the basis of lowest P-value. Regression ansalysis is carried out till the step at which all variable appears significant ($P < 0.05$) in any one or more regression equation.

(b) By backward selection method, in which initially all variable indices are taken in regression analysis and variable indices are removed in step by step. Removal is done on the basis of P-value. In each step the variable index, whose removal results in highest significant regression equation, is removed.

Various molecular descriptors choosed for selected set of molecular structure of cephalosporin type drugs and calculated

with the help of Dragon software [9]. For this purpose 2D and 3D structures prepared by Chem sketch [10] and saved in the form of mol. files. Then some physico-chemical properties calculated by above software. The pharmacological property (activity) collected from reliable cheminformatics sources. Stepwise regression analysis was carried out for modeling the physico-chemical and pharmacological properties (activity) with the help of selected descriptors by Micro soft office 2007 software [11]. A regression equation and hence a regression model was developed. Then each physico-chemical and pharmacological property were predicted with the help of respective developed regression equations. Importance of regression equation was tested by calculating the correlation between observed and predicted values. Correlation between observed and predicted values were expressed by means of simple R^2, Adjusted-R^2, pearson product moment correlation coefficient $r^2 \equiv r$, PRESS. The detailed information of selected molecules for present study is given in table no.1 as follows:

Table no. 1 Structure of selected molecules for the study

Molecular Set	Name	Structure of molecule	Molecular Set	Name	Structure of molecule
C_1	Cefadroxil		C_{12}	Ceftobiprole	
C_2	Cefalexin		C_{13}	Cefminoxime	
C_3	Cefaclore		C_{14}	Cefpodizime	
C_4	Cefradine		C_{15}	Cefotaxime	
C_5	Cefalothin		C_{16}	Cefepime	
C_6	Cefoxitin		C_{17}	Ceftizoxime	
C_7	Cefazoline		C_{18}	Ceftriaxone	
C_8	Ceftezole		C_{19}	Cefexime	
C_9	Cefmetazole		C_{20}	Cefpimizole	
C_{10}	Cefotetam		C_{21}	Cefoperazone	
C_{11}	Cefdinir		C_{22}	Ceftazidime	

A description of selected pharmacological property (activity) are as follows:

Half maximal Inhibitory Concentration (IC_{50}):

It is a measure of the effectiveness of a compound in inhibiting biological or biochemical function. Often, the compound in question is a drug candidate. This quantitative measure indicates how much of a particular drug or other substance (inhibitor) is needed to inhibit a given biological process (or component of a process, i.e. an enzyme, cell, cell receptor or microorganism) by half. The IC_{50} [12] of a drug can be determined by constructing a dose-response curve and examining the effect of different concentrations of antagonist on reversing agonist activity. IC_{50} values can be calculated for a given antagonist by determining the concentration needed to inhibit half of the maximum biological response of the agonist.

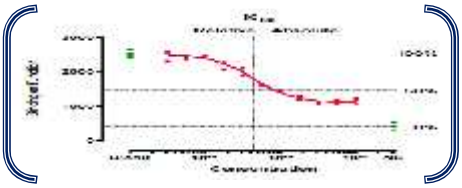


Figure 3: Illustration of IC_{50} by graph.

IC_{50} is not a direct indicator of affinity although the two can be related at least for competitive agonists and antagonists by the following Cheng-Prusoff equation [13]:

$$K_i = IC_{50} / \{1 + ([S]/K_m)\}$$

Where, K_i is the binding affinity of the inhibitor, IC_{50} is the functional strength of the inhibitor, [S] is substrate concentration and K_m is the affinity of the substrate for the enzyme. Whereas the IC_{50} value for a compound may vary between experiments depending on radioligand concentration, the K_i is an absolute value. K_i is the inhibition constant for a drug. The concentration of competing ligand in a competition assay which would occupy 50% of the receptors if no radioligand were present. IC_{50} values can be determined through journal searching and cheminformatics as wel as estimated also by various softwares calculations available online or in purchased version. These softwares require 3D structures of considered molecules.

Brief descriptions of selected descriptors [14], [15], [16] are as follows:

Wiener Index (W): Wiener index (also Wiener number, introduced by H.Wiener in 1947) is a topological index of a molecule, defined as half the sum of the numbers of edges in the shortest path in a chemical graph between all pairs of non-hydrogen atoms in a molecule. Chemical graph or molecular graph is the collection of points and lines connecting these points. This index was introduced by H. Wiener in 1947 Wiener index may be calculated using distance matrices.

Harrary Index (H):Harrary Index H is defined as half sum of the element of the inverse distance matrix.

$$H = \frac{1}{2}\sum_{i=1}^{j=1} (d^{-1})_{ij}$$

Where, $(d^{-1})_{ij}$ = ijth element of inverse distance matrix.

Schultz Molecular Topological Index (SMTI): This index was given by Harry P. Schultz,

SMTI is defined as the summation of all the terms obtained by multiplying valence row matrix (V) to sum of adjacency matrix (A) to distance matrix (D).

$$SMTI = \sum_{i=1}^{N} [V(A + D)]$$

Randic Index (X): This Index was proposed by Randic in 1975 and shows the effect of vertices i.e. shape of molecule over properties and activity. This is defined as summation of all terms obtained by taking inverse root of each term's derived by multiplication of coloumn and horizon element of matrix for connectivity. $X = \sum [d_i\ d_j]^{-\frac{1}{2}}$

d_i d_j are valences of vertices i and j which is equal to number of bonds connected to atom i and j.

Balban Index (J):

$$J = [\frac{M}{(\mu+1)}] \sum (D_i D_j)^{-1}.$$

M= No. of bonds (edges) .

μ = Cyclometric no.

D_i D_j =Sum of row i coloumn j of a distance matrix respectively.

Detour Index (W′)

It is half the sum of all matrix elements present in Detour matrix (DD). Entries of the Detour matrix (DD) represent the maximum topological distances between two given vertices.

$$W' = \frac{1}{2}\sum_{i,j}^{N_{SA}} dd_{ij}$$

dd_{ij} – the number of bonds in the longest path connecting the pair of atoms i and j

N_{SA} – the number of non-hydrogen atom in the molecule. This was first introduced by Harary in 1969, in the context of Graph theory. When there is no cycle present in the considered molecular structure, the distance and the detour matrixes are identical. This can characterize one or more rings in the molecule. Detour matrix has been proposed as a tool to characterize cyclic structures.

T(N-O): Sum of Topological Distances between Nitrogen and Oxygen atom.

T(N-N): Sum of Topological Distances between Nitrogen and Nitrogen atom.

T(N-S): Sum of Topological Distances between Nitrogen and Sulphur atom.

Topological Distances used in T(N-O), T(N-N), T(N-S):

This was given by Bonchev (1991) Trinajstic (1992). Let $\{x_1 \ldots x_k\}$ be the atoms of a particular chemical element x in the molecule. Similarly, $\{y_1 \ldots y_m\}$ are the atoms of a

particular chemical element y in the molecule. We permit the case of x = y as well. Then the atom pair topological distance between the chemical elements x and y is given by:

$$\frac{1}{2}\left(\sum_{i=1}^{k}\sum_{j=1}^{m} d_{ij}\right)$$

where, d_{ij} is the topological distance between x_i and y_i .

3D MoRSE Descriptors:

3D MoRSE descriptors (3Dimensional Molecule Representation of Structures based on Electron diffraction) are derived from infrared spectra simulation using a generalized scattering function. A typical MoRSE descriptor is denoted by Mor (sw), where s and w take the values $1 \leq s \leq 32$ and

$$w \varepsilon (u, m, v, e, p)$$

where, u is unweighted

m is weighted by mass

v is weighted by van der Waals volume

e is weighted by electronegativity

p is weighted by polarizability

The MoRSE descriptor is defined as follows:

$$\mathrm{Mor}(s, w) = I(s, w) = \sum_{i=2}^{n}\sum_{j=1}^{i-1} w_i w_j \sin(sr_{ij}) / (sr_{ij})$$

where, r_{ij} is the Euclidean distance between the atoms *i* and *j* ,

and w_i and w_j are the weights of the atoms i and j respectively.

Keir and Hall valence connectivity indices: The molecular connectivity valence delta-values have been defined in terms of the count of no hydrogen. Non valence electrons on a valence-state atom as screened from the nucleus by the core electrons.

Z_k - Total number of electrons present on atom k.

Z_k^v - Number of valence electrons in the k^{th} atom

H_k - Number of hydrogen atoms directly attached to the k^{th} non-hydrogen atom

m = 0 - Atomic valence connectivity indices

m = 1 - One bond path valence connectivity indices

m = 2 - Two bond fragment valence connectivity indices

m = 3 Three contiguous bond fragment valence connectivity indices etc.

Valence connectivity for the k^{th} atom in the molecular graph,

$$\delta_k^v = \frac{(z_k^v - H_k)}{(Z_k - Z_k^v - 1)}$$

Above all descriptors [17], [18] some other indicator indices representing the no. of particular atom in particular group or structure like no. of S atoms in group –X etc.s are also counted and applied in MLR. The combination of 3D Morse and sum of Keir hall topological distances are applied shown in table no. 4 Other indices like $(EN_{orb.})^{*O=C<}$, Donar sites and α_{XX} calculated by Dragon shown in table no. 5 considered for MLR analysis. For determining microscopic contribution of group –X the common indices like W, H, X^1, W', T(N-N), T(N-S), T(N-O) also calculated for selected set of molecules by removing –R group as shown in table no. 7

Experimental:

IC_{50} values in terms of pIC_{50} values searched out from literature. Selected molecules 2D and 3D structures were prepared by chem sketch software and saved in mol file formate. The various molecular descriptors were calculated by Dragon software (given in table no. 2, 3, 4, 5, 6 and 7)

In order to derive relationship of dependent variable (activity/ property) and independent variables (descriptors/ indices) different approaches for QSAR have been offered in the literature. These approaches range from simple linear regression (LR) and multiple linear regression trained by a variety of methods. MLR techniques can be used to determine the most relevant features for a given model.

Multiple regression carried out by MS excel 2007 through the method (a) out of above discussed methods.

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Table no. 2 Topological, connectivity and geometrical indices of selected set of molecular structures for QSAR study.

Indices → Molecular Set ↓	pIC_{50}	Weiner (W)	Randic (X^1)	Balban (J)	Detour (W')	Harary (H)	T(N-N)	T(N-S)	T(N-O)	SMTI
C_1	4.96	1570	11.81	1.57	2690	46.77	12	9	76	6763
C_2	4.64	1242	11.00	1.56	2222	42.42	12	9	53	5319
C_3	4.66	1383	11.41	1.58	2443	44.73	12	9	53	5882
C_4	4.94	1383	11.41	1.58	2443	44.73	12	9	44	5882
C_5	6.2	1790	12.38	1.56	2946	48.00	3	14	52	7526
C_6	5.63	1837	12.79	1.7	3045	51.36	16	31	79	7670
C_7	5.96	2468	13.94	1.3	4075	55.01	191	172	224	10736
C_8	0.89	2223	13.54	1.31	3727	52.95	191	172	175	9732
C_9	6.28	2591	14.35	1.64	4105	56.63	144	130	254	10706
C_{10}	5.16	4350	17.01	1.37	6438	70.53	150	196	516	18294
C_{11}	6.69	1701	12.38	1.62	2835	48.94	46	42	144	7128
C_{12}	6.06	4333	17.29	1.19	7008	71.76	222	48	365	18914
C_{13}	6.01	3413	15.81	1.35	5431	63.76	276	177	338	14566
C_{14}	5.77	4530	17.19	1.31	6928	69.30	101	162	350	19112
C_{15}	5.66	2584	14.24	1.62	4020	56.56	46	42	244	10666
C_{16}	5.68	3029	15.24	1.40	4863	62.51	91	58	180	12948
C_{17}	5.58	1531	11.94	1.61	2562	46.52	46	42	144	6456
C_{18}	5.74	4386	17.12	1.35	6872	70.17	219	152	435	18532
C_{19}	5.05	2560	14.24	1.63	3930	56.47	46	42	218	10576
C_{20}	4.8	8803	21.84	1.17	12869	91.97	99	116	554	37294
C_{21}	4.91	7383	21.01	1.23	11123	88.41	294	108	646	31520
C_{22}	4.84	4573	17.49	1.4	6924	72.45	91	58	28	19281

Table no. 3: Values of 3D indices (3D Morse and Keir Hall) for selected set of molecular 3D- structures.

Molecular Set	$(S_s)^{wox}$	$(Mor_{01u})^{wox}$	$(Mor_{01am})^{wox}$	$(Mor_{01v})^{wox}$	$Mor_{01en})^{wox}$	$(Mor_{01p})^{wox}$	$(S_s)^{wor}$	$(Mor_{01u})^{wor}$	$(Mor_{01am})^{wor}$	$(Mor_{01v})^{wor}$	$(Mor_{01en})^{wor}$	$(Mor_{01p})^{wor}$
C_1	68.8	741	405.6	307.8	791.7	304.2	48.2	325	190.5	129.8	353.5	146.1
C_2	63.2	703	368.6	295.1	739.9	328.4	46.5	253	168.1	105.3	280.8	117.5
C_3	63.2	703	368.6	295.1	739.9	328.4	50.3	253.0	222.4	115.5	288.3	130.7
C_4	62.2	780	368.6	295.1	739.9	328.4	48.2	325	190.5	129.8	353.5	146.1
C_5	57.2	528	346.2	237.2	562.9	275.5	61.8	496	296.8	194.8	548.6	215.8
C_6	62.6	666	415.5	285.5	713.1	330.2	65.8	561	328.2	215.0	620.4	237.4
C_7	60.7	465	317.6	197.9	519.2	212.3	62.7	595	459.9	270.5	646.1	324.4
C_8	60.7	465	317.6	197.9	519.2	212.3	61.0	496	422.9	234.6	546.5	278.2
C_9	64.9	595	389.2	248.6	647.2	286.7	69.6	741	496.1	304.8	816.0	343.6
C_{10}	87.2	820	667.3	362.3	925.6	418.8	69.6	741	496.1	304.8	816.0	343.6
C_{11}	71.0	595	452.1	266.5	662.8	298.5	51.2	351	210.6	146.3	380.6	163.7
C_{12}	71.8	630	412.3	264.6	700.7	283.8	71.8	1081	514.6	420.7	1136	468.9
C_{13}	70.6	703	488.2	304.8	772.1	343.2	64.2	595	420.1	254.8	654.8	287.8
C_{14}	70.6	703	488.2	304.8	772.1	343.2	75.8	741	573.4	335.1	815.3	392.7
C_{15}	70.6	703	488.2	304.8	772.1	343.2	61.8	496	296.8	495.8	548.5	215.8
C_{16}	71.5	666	485.5	297.4	735.5	333.2	58.2	820	350.9	300.4	850.4	342.4

C_{17}	70.6	703	488.2	304.8	772.1	343.2	46.5	253	168.1	105.3	280.8	117.5
C_{18}	70.5	703	488.2	304.8	772.1	343.2	78.0	741	530.7	317.6	825.0	355.4
C_{19}	84.6	820	609.6	356.8	922.1	395.9	51.2	351	210.6	146.3	380.6	163.7
C_{20}	95.8	1275	747.4	538.5	1334	578.3	84.2	990	610.1	406.7	1076	461.8
C_{21}	86.5	1225	666.1	509.1	1311	555.1	64.2	595	420.1	254.8	654.8	287.8
C_{22}	89.4	1035	691.3	438.4	1141	488.6	60.7	595	337.4	257.5	633.0	284.7

> $(Mor_{1u})^{wox}$, and $(Mor_{1am})^{wox}$, $(Mor_{1v})^{wox}$, $(Mor_{1en})^{wox}$ $(Mor_{1p})^{wox}$ = 3D MoRSE descriptors for Cep training set molecules with out -x group unweighted and weighted by atomic mass, vander waal volume, electronegativity, polarizability respectively.

> $Mor_{1u})^{wor}$, and $(Mor_{1am})^{wor}$, $(Mor_{1v})^{wor}$, $(Mor_{1en})^{wor}$ $(Mor_{1p})^{wor}$ = 3D MoRSE descriptors for Cep training set molecules with out -x group unweighted and weighted by atomic mass, vander waal volume, electronegativity, polarizability respectively

> $(S_s)^{wox}$, $(S_s)^{wor}$ = Keir-Hall electro topological state for Cep training set molecules with out -x group and –R group respectively.

(Superscript 'wox' and 'wor' shows values for structures with out –x group and with -R group respectively.)

C_3	113.45	12.89	2435.05	4167.10	6602.15	-1732.1
C_4	111.34	15.0	2439.73	4176.46	6616.19	-1736.7
C_5	119.0	-4.66	1949.90	3371.80	5321.7	-1421.9
C_6	128.33	-3.17	2410.38	4154.76	6565.14	-1744.4
C_7	123.34	-2.0	1711.97	2958.94	4670.91	-1247.0
C_8	121.67	-0.33	1711.97	2958.94	4670.91	-1247.0
C_9	134.50	-4.66	2166.76	3738.52	5905.28	-1571.8
C_{10}	156.83	17.67	3194.12	5568.24	8762.36	-2374.1
C_{11}	122.17	19.83	2274.93	3954.86	6229.79	-1679.9
C_{12}	143.66	0.0	2291.44	3952.88	6244.32	-1661.4
C_{13}	134.72	6.38	2611.31	4519.62	7130.93	-1908.3
C_{14}	146.38	-5.28	2611.31	4519.62	7130.93	-1908.3
C_{15}	132.38	8.72	2611.31	4519.62	7130.93	-1908.3
C_{16}	129.67	13.33	2517.60	4369.20	6886.80	-1851.6
C_{17}	117.05	24.05	2611.31	4519.62	7130.93	-1908.3
C_{18}	148.55	-7.45	2611.31	4519.62	7130.93	-1908.3
C_{19}	135.84	33.50	3104.48	5388.96	8493.44	-2284.5
C_{20}	180.08	11.58	4472.80	7670.60	12143.40	-3197.8
C_{21}	150.67	22.33	4266.44	7307.88	11574.30	-3041.4
C_{22}	150.09	28.75	3794.64	6554.28	10348.90	-2759.6

Where, $(S_s)^{wox}$, $(S_s)^{wor}$ = Keir-Hall electro topological state for Cep training set molecules with out -x group and –R group respectively.

Sum of all 3D MoRSE descripters unweighted and v atomic by mass, vander waal volume, electronegativ x group = $\sum_{u,am,v,en,p}$(MoRSE)wox

Sum of all 3D MoRSE descripters unweighted and weighted atomic by mass, vander waal volume, elec les without − R group = $\sum_{u,am,v,en,p}$(MoRSE)wor

Table no. 4: Descripters generated by combination of Keir-Hall electro topological state and 3D MoRSE

Molecular Set	$(S_s)^{wox} + (S_s)^{wor}$	$(S_s)^{wox} - (S_s)^{wor}$	$\sum$(MoRSE)wox		$\sum$(MoRSE)wox + wor	wor
			A	B	C	D
C_1	117.0	20.66	2550.27	4359.54	6909.81	-1809.3
C_2	109.67	16.67	2435.05	4167.10	6602.15	-1732.1

Table no. 5: Values of $(EN_{orb.})^{*O=C<}$, Donar sites and α_{XX}

Molecular Set	$(EN_{orb.})^{*O=C<}$	(Sites)	(Sites)	(3D-Surfacearea)$^{V.W.}$	α_{XX}	Training Set	$(EN_{orb.})^{*O=C<}$	(Sites)	(Sites)	(3D-Surfacearea)$^{V.W.}$	α_{XX}
C_1	12.5	9	5	438.6	36.2	C_{12}	11.4	15	6	590.5	49.5
C_2	12.3	8	4	396.6	34.2	C_{13}	12.9	13	4	561.9	74.6
C_3	12.3	9	4	409.9	50.1	C_{14}	12.9	13	5	614.7	88.4

C_4	12.3	8	4	424.2	50.3	C_{15}	12.9	12	4	522.3	62.5
C_5	12.7	8	2	467.2	37.4	C_{16}	12.8	10	3	592.9	70.3
C_6	12.8	9	4	436.5	42.2	C_{17}	12.8	10	4	418.9	48.6
C_7	12.9	11	2	490.9	70.4	C_{18}	12.9	16	5	603.8	80.6
C_8	12.9	11	2	457.6	65.8	C_{19}	12.9	13	5	485.6	59.1
C_9	12.9	12	2	535.9	48.4	C_{20}	12.9	18	5	788.9	70.8
C_{10}	13.1	16	5	590.6	590.6	C_{21}	12.9	17	4	759.2	73.3
C_{11}	12.9	10	5	408.1	54.0	C_{22}	12.9	13	4	623.1	57.9

Table no. 6: Index showing the no. of different types atoms present in particular structure.

Molecule set	Total no. of atoms in group-X, $(N_T)^X$	No. of C atoms in group X, $(N_C)^X$	No. of N atoms in group- X, $(N_N)^X$	No. of S atoms in group X, $(N_S)^X$	No. of O atoms in group X, $(N_O)^X$	Total no. of atoms in group R, $(N_T)^R$	No. of C atoms in group –R, $(N_C)^R$	No. of N atoms in group –R, $(N_N)^R$	No. of S atoms in group –R, $(N_S)^R$	No. of O atoms in group –R, $(N_O)^R$
C_1	1	0	0	0	0	9	7	1	0	0
C_2	0	0	0	0	0	8	7	1	0	0
C_3	1	0	0	0	0	8	7	1	0	0
C_4	1	1	0	0	0	8	7	1	0	0
C_5	5	3	0	0	2	6	5	0	1	0
C_6	4	1	1	0	1	6	5	0	1	0
C_7	6	2	2	2	0	6	2	4	0	0
C_8	8	3	2	3	0	6	2	4	0	0
C_9	8	3	4	1	0	5	3	1	1	0
C_{10}	8	2	4	1	0	11	5	1	2	3
C_{11}	2	2	0	0	0	9	4	3	1	1
C_{12}	12	9	2	0	1	9	4	4	0	1
C_{13}	8	3	4	1	0	9	4	3	1	1
C_{14}	11	6	1	2	2	9	4	3	1	1
C_{15}	5	3	0	0	2	9	4	3	1	1
C_{16}	7	6	1	0	0	9	4	3	1	1
C_{17}	0	0	0	0	0	9	4	3	1	1
C_{18}	10	4	3	1	2	9	4	3	1	1
C_{19}	2	2	0	0	0	12	5	3	1	3
C_{20}	13	8	1	1	3	17	12	3	3	3
C_{21}	8	3	4	1	0	14	7	3	0	4
C_{22}	7	6	1	0	0	15	8	3	1	3

Table no. 7: Topological, connectivity and geometrical indices of selected set of molecular structures without –R group for QSAR

Molecule Set	$(W)^{wor}$	$(X_1)^{wor}$	$(J)^{wor}$	$(W')^{wor}$	$(H)^{wor}$	$\{T(N\text{-}N)\}^{wor}$	$\{T(N\text{-}S)\}^{wor}$	$\{T(N\text{-}O)\}^{wor}$	$(SMTI)^{wor}$
C_1	415	7.58	2.07	839	28.01	3	3	26	1718
C_2	1242	11.00	1.55	2222	42.42	12	9	53	5319
C_3	415	7.58	2.07	839	28.01	3	3	26	1718
C_4	415	7.58	2.07	839	20.01	3	3	26	1718
C_5	812	9.47	2.02	1444	35.44	3	3	12	3284
C_6	866	9.89	2.19	1554	38.48	16	8	79	3468
C_7	1079	10.63	1.6	1971	40.33	35	45	96	4624
C_8	1486	11.96	1.68	2635	42.27	67	46	207	6224
C_9	1486	11.96	1.68	2635	42.27	67	46	207	6244
C_{10}	492	8.11	2.09	968	29.94	3	3	26	2018
C_{11}	1873	13.02	1.39	3404	52.26	42	18	124	8160
C_{12}	1218	11.04	1.60	2221	42.49	67	46	166	5180
C_{13}	1822	12.42	1.52	3022	47.93	18	35	115	7626

C_{14}	812	9.47	2.022	1444	35.44	3	3	52	3284
C_{15}	1015	10.47	1.71	1907	41.26	14	7	51	4324
C_{16}	354	7.16	2.01	726	25.78	3	3	26	1482
C_{17}	1728	12.34	1.57	3056	48.77	53	36	182	7246
C_{18}	492	8.11	2.09	968	29.94	3	3	26	2018
C_{19}	2276	13.21	1.51	3624	52.12	14	36	147	9460
C_{20}	1218	11.04	1.61	2221	42.49	67	46	166	5180
C_{21}	1048	10.63	1.65	2012	40.65	14	7	51	4480
C_{22}	492	8.11	2.09	968	29.94	3	3	26	2018

Table no. 8: Correlation of some selected indices with pIC_{50} values

Indices	Correlation	Indices	Correlation	Indices	Correlation	Indices	Correlation
W	-0.01	T(N-S)	-0.14	$(Mor_{1v})^{wox}$	-0.03	$(Mor_{1v})^{wor}$	0.15
χ1	0.02	T(N-O)	0.05	$(Mor_{1en})^{wox}$	-0.05	$(Mor_{1en})^{wor}$	0.18
J	0.20	SMTI	-0.01	$(Mor_{1p})_{wox}$	0.001	$(Mor_{1p})^{wor}$	0.10
W'	-0.01	$(S_s)^{wox}$	-0.004	$(S_s)^{wor}$	0.14	$(N_S)^X$	-0.47
H	0.03	$(Mor_{1u})^{wor}$	-0.64	$(Mor_{1u})^{wor}$	0.19	$(S_s)^{wox}$ - $(S_s)^{wor}$	-0.12
T(N-N)	-0.06	$(Mor_{1am})^{wox}$	0.05	$(Mor_{1am})^{wor}$	0.03	α_{XX}	-0.03

Table no. 9: Summary of stepwise multinear regression for developing a model equation for prediction of pIC_{50}

	Step no.	Developed MLR equation	R^2	PRESS	SE
1	Step -1	pIC_{50}= 5.6473-1.6499 (±0.2638) (N_S)	0.219	21.327	1.03
2	Step -2	pIC_{50}= 4.9708 - 1.6499(±0.4288) $(N_S)^X$ + 0.0157(± 0.0056) T(N-S)	0.449	15.041	0.89
3	Step -3	pIC_{50}= 5.5383 - 1.9456(±0.3916) $(N_S)^X$ + 0.0160(±0.0049) T(N-S) - 0.0413(±0.0158) $\{(S_s)^{wox} - (S_s)^{wor}\}$	0.601	10.894	0.78
4	Step -4	pIC_{50}= 5.4085 - 2.1864(±0.4190) $(N_S)^X$ + 0.0213 (±0.0061) T(N-S) - 0.0353(±0.0159) $\{(S_s)^{wox} - (S_s)^{wor}\}$ – 0.0027(±0.0019) α_{XX}	0.642	9.778	0.76
5	Step -5 (Final)	No variable can be added satisfactory.	-	-	-

Results and discussions: By multiple linear regression equation analysis regression equation developed for prediction of IC_{50} can be represented by the following equation:

$$pIC_{50} = 5.4085 - 2.1864(\pm 0.4190)\ (N_S)^X + 0.02133\ (\pm 0.0061)\ T(N\text{-}S) - 0.0353(\pm 0.0159)\ \{(S_s)^{wox} - (S_s)^{wor}\} - 0.0027(\pm 0.0019)\ \alpha_{XX}$$

Statistics of the developed regression equation is as follows:

n	R^2-Adj.	R^2	Pearson's R	F-ratio	Overall significance-F	SE	PRESS
22	0.558	0.642	0.801	7.624	0.001	0.76	9.778

Predicted value releated to observed value by the following equation:

$$pIC_{50})_{pred.} = 0.642(pIC_{50})_{obs.} - 1.889$$

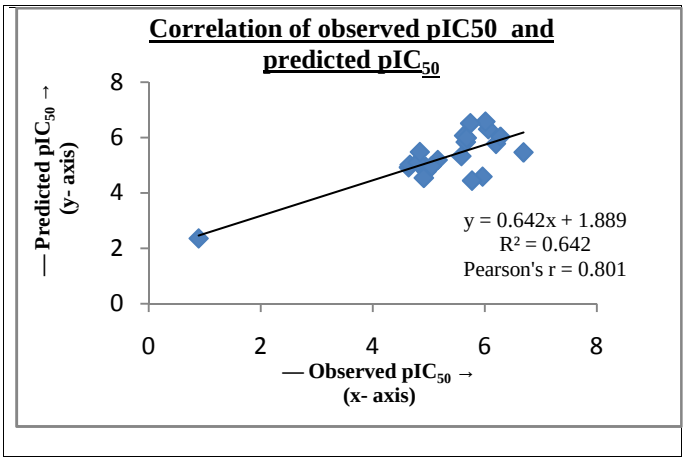


Figure no. 4 Graph showing predicted and observed pIC_{50} for selected set of molecules

Table no. 10: Predicted and observed values of pIC_{50}

Moleculear Set	Observed pIC_{50}	Predicted pIC_{50}	Molecular Set	Observed pIC_{50}	Predicted pIC_{50}
C_1	4.96	4.77	C_{12}	6.06	6.29
C_2	4.64	4.91	C_{13}	6.01	6.57
C_3	4.66	5.01	C_{14}	5.77	4.44
C_4	4.94	4.94	C_{15}	5.66	5.83
C_5	6.2	5.77	C_{16}	5.68	5.98
C_6	5.63	6.07	C_{17}	5.58	5.32
C_7	5.96	4.58	C_{18}	5.74	6.51
C_8	0.89	2.35	C_{19}	5.05	4.96
C_9	6.28	6.03	C_{20}	4.8	5.09
C_{10}	5.16	5.18	C_{21}	4.91	4.54
C_{11}	6.69	5.46	C_{22}	4.84	5.47

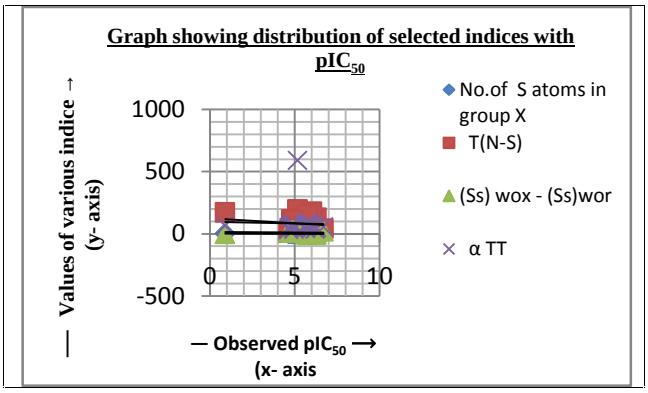


Figure 5: Distribution of selected indices with pIC_{50}

CONCLUSIONS:

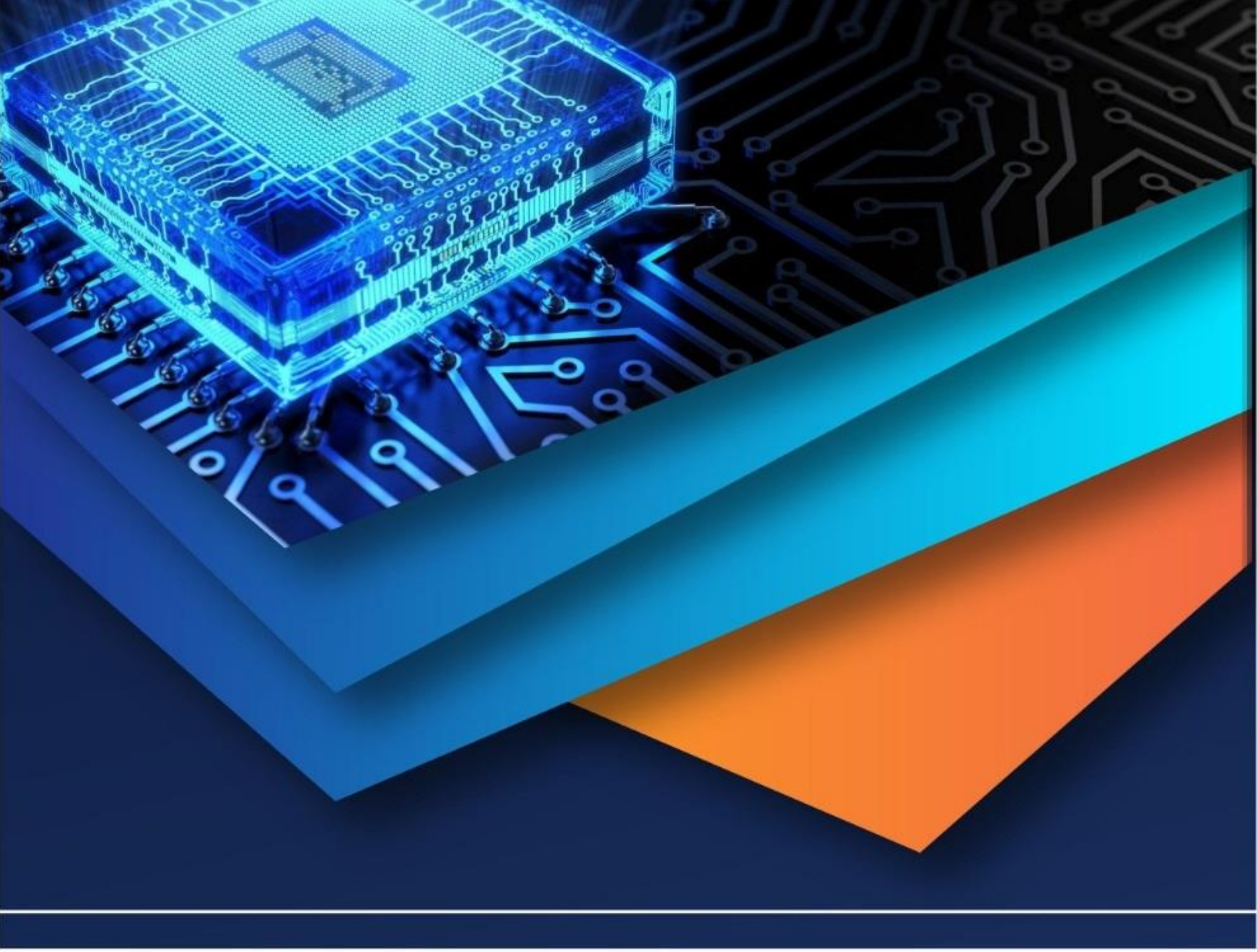
From the developed equation this can be concluded that IC_{50} value is predicted by $(N_S)^X$ in larger proportion than others. Since, $(N_S)^X$ appears in negative contribution so to decrease pIC_{50} no. of S atoms in group –X of cephalosporin core structure should be minimized and vice versa. However, T(N-S) appears in positive factor so to increase pIC_{50} sum of topological distances between nitrogen and sulphur should be maximized. Another fact which should be considered during designing of new molecule may be that for larger pIC_{50} new molecule. However, polarizibility in XX direction also helpful till little extent to propose new molecules with desired IC_{50} value.

REFERENCES:

1. R. B. Morin, M. Gorman, Chemistry and Biology of β-lactum antibiotics, Academic Press, New-York, 1, pp 1-10, 1982 .
2. C. E. Newall, P. D. Hallam, Comprehensive Medicinal Chemistry, 2, pp 609-653, Pergamon Press, 1990.
3. E. B. Grant, D. Guiadeen, E. Z. Baum, B. D. Foleno, H. Jin, D. A. Montenegro, E. A. Nelson, K. Bush, D. J. Hlasta, Bioorganic & Medicinal Chemistry Letters , Pergamon, Science direct, 10, pp 2179- 2182, 2000.
4. J. G. Hardman, L. E. Limbird, In: Goodman and Gilman's The Pharmacological Basis of Therapeutics, McGraw-Hill Publication, New York, 10th ed., 2001.
5. R. A. Johnson, D. W. Wichern, Applied Multivariate Stastical Analysis, Prentice- Hall International, USA, 1998.
6. M. Kuanar, S. K. Kuanar, B. K. Mishra, “Correlation of line graph parameters with physicochemical Properties”

Indian Journal of Chemistry, ,vol. 38A, pp 525-528, 1999.
7. K. T. Fang, S. Kotz, K. W. Ng, Symmetric Multivariate and Related Distributions, Chapman and Hall , London, 1990.
8. W. Stute, S. Thies, L. X. Zhu, "Model Checks for regression", in An Innovation process approach, Ann. Statist. Publications, 26, pp 1916- 1934, 1998.
9. Dragon evaluation verson, Talete srl. Milano chemometrics, Italy, verson 5.5, 2007.
10. ACDLAB product, ACD/ChemSketch, version 12.0, for Microsoft, 2007.
11. P. P. Drake, Online notes on "Regression step-by-step, using Microsoft Excel," James Madison University, Acessed on 20th Oct. 2012.
12 M. J. Stewart, I. D. Watson, British Journal of Clinical Pharmacology, 16 (1), pp 3-7, 1983.
13 Y. Cheng, W. H. Prusoff, Biochem Pharmacol., 22 (23), pp 3099–108, 1973.
14. L. B. Keir, L. H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New-York, 1, pp 1-110, 1976.
15. L. S. Urra, M. P. Gonzalez, M. Teijeira, 3D-MoRSE descriptors and statistical considerations about variable selection, Bioorg. Med. Chem., 14(21), pp 7347-7358, 2006
16. J. Schuur, J. Gasteiger, "Software Development in Chemistry", Fachgruppe Chemie-Information-Computer (CIC), Frankfurt am Main (Germany), 10, pp 67-8, 1996.
17 R. K. Sharma, A. Sikarwar, N. Gupta, "Correlation Studies of Topological Indices for Cephalosporin type Antibiotic Drug Derivatives", Asian Journal of Biochemical and Pharmaceutical Research, 1(2), pp 443-452, 2011.
18 R. K. Sharma, A. Sikarwar, R. Sharma, P. Sharma, "QSAR and QSPR Modeling of Parachore and Molar Volume (V_m) for Cephalosporin type drug derivatives containing β-lactum type structure with the help of some topological, connectivity, geometrical and 3D indices", AJBPR, 2(4), pp 180-194, 2012.

-*-



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY